The Relation Between Analysts' Forecasts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings

by

PATRICIA M. DECHOW
University of Michigan Business School
Ann Arbor, MI 48109-1234

AMY P. HUTTON
Harvard Business School, Harvard University
Boston, MA 02163

and

RICHARD G. SLOAN*
University of Michigan Business School
Ann Arbor, MI 48109-1234

First Version: July, 1996
This Version: June, 1999

---

* We thank Chris Allen, Sarah Eriksen and Sarah Woolverton for their research assistance. We are grateful for the comments and suggestions of two anonymous referees, S.P. Kothari (the discussant), Gordon Richardson (the editor), and participants of the 1998 Contemporary Accounting Research Conference. We also are grateful for the comments of participants in the summer research colloquiums at Harvard University and Stanford University and workshops at the University of California at Berkeley, Emory University, the University of Georgia, the University of Illinois at Urbana-Champaign, the University of Michigan, University of North Carolina at Chapel Hill, Washington University, and University of Waterloo. We thank I/B/E/S/ for providing earnings forecast data.

The Relation Between Analysts' Forecasts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings

Abstract

We evaluate the role of sell-side analysts' long-term earnings growth forecasts in the pricing of common equity offerings. We find that, in general, sell-side analysts' long-term growth forecasts are systematically overly optimistic around equity offerings and that analysts employed by the lead managers of the offerings make the most optimistic growth forecasts. Additionally, we find a positive relation between the fees paid to the affiliated analysts' employers and the level of the affiliated analysts' growth forecasts. We also document that the post-offering under performance is most pronounced for firms with the highest growth forecasts made by affiliated analysts. Finally, we demonstrate that the post-offering under performance disappears once we control for the over optimism in earnings growth expectations. Thus, the evidence presented in this paper is consistent with the 'equity issue puzzle' arising from overly optimistic earnings growth expectations held at the time of the offerings.

I. INTRODUCTION

This paper evaluates the role of sell-side analysts' long-term earnings growth forecasts in the pricing of common equity offerings. While it is well documented that firms experience unusually low stock returns in the five years following equity offerings (Loughran and Ritter 1995 and Spiess and Affleck-Graves 1995), the reason for this under performance is not well understood. Competing explanations include mismeasured risk-premia, research design biases, and overly optimistic expectations about future firm performance.[1] In this paper we examine the overly optimistic expectations explanation for the 'equity issue puzzle.' We investigate whether sell-side analysts' long-term growth forecasts are overly optimistic at the time of equity offerings and whether these overly optimistic expectations are reflected in stock prices. We also investigate whether the over optimism in analysts' forecasts and the corresponding overpricing of equity offerings is greatest for offers covered by analysts affiliated with the lead investment bank underwriting the offering.

The concern that sell-side analysts compromise their objectivity and independence in order to win investment-banking business is often discussed in the financial press.[2] This concern arises because analysts' employers, the investment banks, provide both brokerage services to investor clients and underwriting services to client firms. A conflict of interests arises when an analyst issues a negative recommendation for a stock that is simultaneously being solicited for underwriting business. This conflict of interest is

---

[1]Brav and Gompers (1997), Barber and Lyon (1997) and Kothari and Warner (1997) discuss potential problems with measurement of risk premia and stock returns. Loughran and Ritter (1997) provide evidence suggesting that pricing multiples around the equity offerings are consistent with investors having overly optimistic expectations about future profitability.

[2]Examples include: "Who is pulling the strings?" by M. Celarier, *Euromoney*, April 1996; "Today's analyst often wears two hats," by R Lowenstein, *Wall Street Journal*, May 2, 1996; "Some analysts enter land of big bucks," by M. Baker, *Wall Street Journal*, July 2, 1996; and "Today, delivering good news is a way to ensure good business relationships," by M. Siconolfi, *Wall Street Journal*, May 18, 1998.

intensified by the fact that analysts earn large bonuses for bringing investment banking clients to his or her firm.[3]

Demonstrating the pressures on analysts, the *Wall Street Journal* has reported several alleged incidents of top executives withholding underwriting business from investment banks whose analysts reduce earnings forecasts or downgrade their firms' stock ratings. Most recently on March 23, 1999, the *Wall Street Journal* reported that after an analyst at Salomon Smith Barney, Colin Devine, cut his target price for Conseco from 36 to 32 and downgraded his rating of its stock from 'outperform' to 'neutral,'

> … people close to Salomon say Mr Hilbert [*CEO of Conseco*] called the head of research at Salomon to complain, and said Conseco would withhold business unless Mr Devine recanted. … Last year, Salomon was lead or co-manager on more than $7 billion of Conseco … securities. Salomon didn't participate in Conseco's two offerings this year, and won't be in on the next one, Conseco confirms.[4]

This conflict of interest has potentially costly consequences for investors purchasing underwritten securities.[5] For example, *Fortune Magazine* blames the rise and fall in Boston Chicken's stock price not on management's failure to disclose losses at the franchises, but rather on the aggressive pushing of the stock by the brokerage firms who underwrote Boston Chicken's many security offerings:

> Truth be told, the trouble is less with Boston Chicken and more with the folks who pushed its stock in spite of the warnings. Those red flags, for instance, didn't prevent analysts from Merrill Lynch, Alex Brown, and Morgan Stanley—Boston Chicken's recent underwriters—from strongly recommending the stock. Indeed, even as the share plunged in April [*1997*] amid reports of slowing sales, these three firms pushed through a mammoth $287.5 million bond offering. [6]

---

[3] See "Some analysts enter land of big bucks," by M. Baker, *Wall Street Journal*, July 2, 1996, page C1.

[4] "Are there cracks in Conseco's house of acquisitions?" by S.J. Paltrow, *Wall Street Journal*, March 23, 1999, page C1.

[5] See also "Wild week for PIA Merchandising includes plummeting stock price, duel of underwriters," by E.S. Browning, *Wall Street Journal*, May 30, 1996, page C2.

[6] "The Boston Chicken Problem - The restaurant chain's rise and fall has been breathtaking. Who is to blame for the mess? Try all those brokerage firms that have been flacking the chichen peddler's puffed-up stock – even as problems mounted," by N. D. Schwartz, *Fortune*, July 7, 1997.

The objective of our research is to provide empirical evidence on whether analyst affiliation affects forecast optimism and in turn whether analysts' optimism is reflected in the stock prices of firms issuing equity. We focus our analysis on analysts' long-term earnings growth forecasts and directly relate the over optimism in these forecasts to the post-offering under performance, the 'equity issue puzzle.'

Compared to buy/sell recommendations and annual earnings forecasts, long-term growth forecasts provide a more powerful measure of market expectations useful for explaining the post-offering under performance. Since stock recommendations fall into only five categories, their ability to explain cross-sectional variation in post-offering returns is limited. The use of annual earnings forecasts as a measure of expectations is also limited because the long-run under performance in stock prices does not begin until several (usually six) months after the offerings and then continues for up to five years (see Loughran and Ritter 1995). Revisions in expectations about currently reported annual earnings are therefore not likely to explain the long-run under performance. The use of long-term growth forecasts also increases the power of our tests since analysts are frequently evaluated on the accuracy of their buy/sell recommendations and annual earnings forecasts, but not their long-term growth forecasts.[7] Thus, reputation effects are less likely to deter analysts from issuing overly optimistic long-term earnings growth forecasts. Finally, recent evidence in Stickel (1998) and Bradshaw (1999) indicate that forecast of long-term growth is an important factor in formulating the recommendation made by analysts. Thus, long-term growth is a number that is followed and used by the investment community.

The evidence presented in this paper is consistent with analysts biasing their forecasts of firms' long-term earnings growth around new equity offerings. The over-optimism in

[7]For example, the *Institutional Investor's* evaluation criteria for ranking analysts for their All-American Research Team specifically mentions short-term price performance and annual earnings forecast accuracy. Long-term growth forecast accuracy is not listed as a criteria (see, Stickel 1992).

these forecasts is most pronounced when the forecasting analyst is affiliated with the lead manager underwriting the offering. Additionally, the level of the growth forecast is positively related to the fees paid to the affiliated analysts' employers. We also document that the post-offering under performance is most pronounced for firms with the highest growth forecasts made by affiliated analysts. Our empirical tests demonstrate that once we control for the over optimism in earnings growth expectations the unusually low post-offering returns disappear. Thus, one interpretation of the evidence is that the 'equity issue puzzle' results from investors' naive reliance on overly optimistic long-term earnings growth forecasts made by analysts at the time of the equity offerings.[8]

The paper proceeds in five sections. The next section discusses existing research. The third section develops our empirical predictions. The fourth section describes our sample and data. The fifth section presents the results and the sixth section provides our conclusions.

II. EXISTING RESEARCH

Prior and concurrent research investigates various aspects of analysts' optimism around equity offerings. This research can be partitioned into research focusing on short-term forecasts, research investigating long-term forecasts and recommendations, research investigating analyst affiliation, and research investigating the stock market response to analysts' forecasts. Below we briefly describe the existing research and our contribution to this literature.

Existing research provides no evidence that analysts' near-term (annual) earnings forecasts are more optimistic around equity offerings, initial or seasoned. Hansen and

[8]Recent research demonstrates that managers manipulate earnings upward around new equity offerings [see Teoh, Welch and Wong (1998) and Rangan (1998)]. This research suggests that managers play a role in facilitating the markets' (and analysts') overly optimistic growth expectations around equity offerings. However, the large negative forecast errors (documented below) indicate that analysts are unable or unwilling to undo the manipulation of expectations by managers.

Sarin (1996) document that in general analysts' annual forecast errors around initial and seasoned equity offerings are not different than their forecast errors at other times [see also Ali (1996)]. They also find no difference in the near-term forecasts of affiliated and unaffiliated analysts [see also Lin and McNichols (1998a) who confirm these results].[9] Hansen and Sarin conclude that analysts are disciplined by reputation forces and consequently forecast credibly around equity offerings.

Noting that studies focusing solely on near-term earnings forecasts cannot resolve the question of whether concern for reputation is sufficient to offset pressures from investment banking relationships, Lin and McNichols (1998a) include an examination of analysts' long-term growth forecasts and stock recommendations. They document that affiliated analysts issue more optimistic long-term growth forecasts and stock recommendations than unaffiliated analysts around seasoned equity offerings. Michaely and Womack (1996) and Lin and McNichols (1998b) provide similar evidence for initial public offerings. Finally, without distinguishing between affiliated and unaffiliated analysts, Rajan and Servaes (1997) also document over optimism in analysts' long-term growth forecasts around initial public offerings (IPOs) and find that the firms with the highest projected growth experience the greatest post-IPO under performance. However, Rajan and Servaes do not attempt to explain the post-IPO under performance with the over optimism in analysts' growth forecasts.

Existing evidence on the effects of analysts' forecasts on the pricing of equities is indirect and mixed. Several papers document that stock prices react to the release of analysts' forecasts and stock recommendations, including Lin and McNichols and Michaely and Womack, who find a significant difference in the stock price reaction to affiliated versus unaffiliated analysts' recommendations around equity offerings. On the other hand, when

[9]Only Dugar and Nathan (1995) provide evidence that investment-banking affiliation affects the level of optimism in analysts' annual earnings forecasts. However, Dugar and Nathan's examination is not conditioned on an equity offering.

the examination is not conditioned on an equity offering, Dugar and Nathan (1995) find no significant difference in the stock price reactions to investment banking and non-investment banking analysts' stock recommendations. However, Dugar and Nathan present evidence consistent with the hypothesis that investors rely less on investment banking analysts' forecasts in forming their annual earnings expectations. In particular, they find that the strength of the relation between analysts' forecast errors and abnormal returns cumulated from the release of analysts' research reports to the next earnings announcement is stronger for non-investment banking analysts.

Overall the existing evidence is mixed and indirect regarding the extent to which investors rely on analysts' forecasts in forming the earnings expectations reflected in stock prices. None of the existing research directly links the over-optimism in analysts' forecasts around equity offerings to the post-offering under performance. Our contribution is providing a direct link. In addition, we also provide evidence that the over-optimism in analysts' long-term growth forecasts and corresponding overpricing of equity offerings is greatest for high growth firms covered by affiliated analysts. Finally, we are the first (to our knowledge) to document a systematic, positive relation between the magnitudes of affiliated analysts' growth forecasts and the underwriting fees paid to their employers.

III. HYPOTHESIS DEVELOPMENT

Below, we first discuss our predictions concerning analysts' earnings growth forecast errors and the biases in these forecasts. We then develop hypotheses concerning the possible ways in which the stock market incorporates information about these biases into stock prices.

Analysts' Forecast Errors

Previous research indicates that analysts tend to be overly optimistic in their forecasts of firms' earnings prospects (Abarbanell 1991; Brown, Foster, and Noreen 1985). The financial press suggests that the objectivity and independence of the analyst community steadily eroded during the 1980s because analysts abandoned primary research as a result of declining commission fees to pursue investment banking fees. The pursuit of investment banking fees gives analysts incentives to provide overly optimistic forecasts for firms with whom they have or wish to have underwriting relationships.

> When commissions on stock trading fell, investment research (which generated trading) no longer paid the freight. Today, analysts are supported partly by their corporate finance departments. And much of what they do -- marketing and preparing IPOs, for instance -- has little to do with pure research, and much to do with investment banking. In the U.S. in particular, investment banks have persuaded clients to hire underwriters on the basis of their analysts' selling power. ... In turn, the analyst's worth is increasingly dependent on his or her ability to bring in deals.[10]
>
> Of course, some money managers grumble that big emphasis on new-issue fees taints research results if the analysts try to avoid saying anything negative about their underwriting clients.[11]

In this paper we hypothesize that sell-side analysts, in general, provide overly optimistic forecasts of issuing firms' long-term earnings growth in order to attract and retain underwriting business. In other words, we hypothesize that α_0 is less than zero in the following equation:

$$FE_{t+1} = \alpha_0 + \varepsilon_{t+1} \tag{1}$$

The dependent variable, FE_{t+1}, is the analysts' forecast error, measured as realized long-term growth in earnings minus the analysts' forecast of long-term growth in earnings. Further, we hypothesize that analysts employed by the investment bank acting as the lead underwriter of the offering have even stronger incentives to make overly optimistic forecasts to lowering the offering firm's cost of capital. Alternatively, managers of the

[10]"Today's analyst often wears two hats," by R Lowenstein, *Wall Street Journal*, C1, May 2, 1996.

[11]"Some analysts enter land of big bucks," by M. Baker, *Wall Street Journal*, C1, July 2, 1996..

issuing firms may systematically select as their lead underwriter the investment bank employing the most optimistic analysts. Either way, we expect analysts employed by the lead underwriter to have the most optimistic forecasts. We refer to such analysts as affiliated and predict that α_0 will be more negative for these analysts.

Prior empirical evidence demonstrates that the optimism in analysts' long-term growth forecasts is increasing in the level of forecast growth [see Dechow and Sloan (1997), La Porta (1996), and Rajan and Servaes (1997)]. Firms receiving the highest long-term earnings growth forecasts, on average, also have larger forecast errors. Thus, the upward bias in analysts' forecasts appears to be driven primarily by the high growth forecasts given to the so-called 'glamour' stocks.[12] Following Dechow and Sloan, we model this phenomenon using a simple linear form:

$$FE_{t+1} = \alpha_0 + \alpha_1 \, Growth_{t+1} + \varepsilon_{t+1} \tag{2}$$

where $Growth_{t+1}$ is the analysts' forecasts of long-term earnings growth, and empirically, $1 < \alpha_1 < 0$.[13]

Use of this more detailed model of analysts' forecast errors, enables us to capture more of the predictable variation in the forecast errors. This, in turn, allows us to conduct more powerful tests of our stock price hypotheses, developed below. Equation (2) can also be used to investigate the nature of the incremental bias in affiliated analysts' long-term earnings growth forecasts. If the bias in affiliated analysts' long-term growth forecasts is unrelated to the level of forecast growth, then α_0 will be more negative for the affiliated analysts than for the unaffiliated analysts, and α_1 will be the same for the two groups.

[12] Labeling these firms 'glamour' stocks, Lakonishok, Shleifer and Vishny (1994) argue that investors over-estimate the future profitability of high growth potential firms.

[13] This regression is identical to regressing realized growth on forecast growth. We use the specification in equation (2) to focus attention on analysts' forecast errors, which we use as our measure of unexpected earnings growth in the stock price tests developed below.

However, if the incremental bias in the affiliated analysts' forecasts is related to the level of forecast growth, then α_1 will be more negative for the affiliated analysts than for the unaffiliated analysts.

Finally, if analysts' overly optimistic forecasts are motivated by their desire to generate underwriting business, then we expect their forecasts of long-term earnings growth to be positively related to the fees paid to their employers, the lead managers underwriting the equity offerings. Thus, we hypothesize that after controlling for realized growth in earnings, the level of affiliated analysts' growth forecasts is higher the greater the fee basis paid to their employers. We also expect analysts' over optimism to be positively related to the fees paid to their employers.

Stock Prices

We develop predictions concerning stock price behavior under two competing hypotheses: 1) the efficient market hypothesis and 2) a naïve expectations hypothesis. Under both hypotheses, investors use information in analysts' long-term earnings growth forecasts to form expectations of future dividends. The competing hypotheses differ with respect to *how* investors use information in analysts' forecasts to form their expectations of future dividends. Under the efficient market hypothesis, investors fully anticipate, and therefore stock prices fully reflect, the predictable bias in analysts' long-term earnings growth forecasts. Under the second hypothesis, investors naively rely on analysts' long-term growth forecasts, neglecting to adjust for the predictable bias in these forecasts when forming their expectations of future dividends. Thus, under the second hypothesis, stock prices fail to reflect the predictable bias in analysts' long-term growth forecasts.

Following Collins, Kothari, Shanken and Sloan (1994), a simple model for testing these competing hypotheses is obtained within the framework of Campbell (1991). Campbell shows that using the traditional dividend discounting valuation model, abnormal stock returns can be approximated as a linear function of the unexpected growth in current dividends and the change in the expected growth of future dividends. By further invoking the common assumption that revisions in dividend expectations are correlated with revisions in earnings expectations, we can express abnormal returns as a linear function of unexpected growth in earnings:[14]

$$AR_{t+1} = \beta_1(\varepsilon_{t+1}{}^*) + \upsilon_{t+1} \tag{3}$$

The dependent variable, AR_{t+1}, measures the abnormal stock return in the five years following the equity offering. $\varepsilon_{t+1}{}^*$ represents the market's assessment of the unexpected earnings growth in the five years following the equity offering. Finally, β_1 represents the valuation multiplier the market applies to unexpected earnings growth.[15]

Substituting the unexpected earnings growth implied by the model of analysts' forecast errors in equation (1) for ε_t^* in equation (3) gives:

$$AR_{t+1} = \beta_1 [FE_{t+1} - \alpha_0^*] + \upsilon_{t+1} \tag{4}$$

In this equation, α_0^* represents the market's assessment of the average bias in analysts' long-term growth forecasts. The efficient market hypothesis predicts that α_0^* will correspond to its counterpart in the equation (1), α_0. In other words, investors'

---

[14] Note that our empirical tests do not involve specific predictions about the magnitude of the response coefficient, β_1. Instead, our tests simply require that abnormal returns are positively associated with unexpected earnings growth. Given this positive relation, we test whether the abnormal stock returns following an equity offering: (1) rationally respond to the unpredictable portion of the deviation between realized growth and analysts' growth forecast, $FE_{t+1} - \alpha_0$ or (2) naively respond to the total deviation between realized growth and analysts' growth forecasts, FE_{t+1}.

[15]In Campbell's model, the theoretical value of β_1 is one. Because we regress five-year returns on annualized growth rates, the theoretical value of β_1 in our specification is five. However, we expect β_1 to deviate from its theoretical value for two reasons. First, we use earnings growth rates in place of dividend growth rates. Second, our specification omits changes in growth expectations beyond the five-year forecast period (since they are not available). However, as indicated in footnote 14, our empirical tests are not based on predictions about the value of β_1. Rather, our tests simply require the relation between stock returns and unexpected earnings growth to be positive.

expectations of future earnings growth rationally anticipate the average bias in analysts' long-term growth forecasts. Thus, stock prices respond only to the unpredictable portion of the analysts' forecast error, ε_{t+1}, which is equal to $FE_{t+1} - \alpha_0$. The naïve expectations hypothesis predicts that α_0^* in equation (4) will equal zero, since investors naively believe that analysts' long-term growth forecasts are unbiased. Under this hypothesis, investors' expectations of future earnings growth equal the analysts' growth forecast. Thus, stock prices respond to the entire forecast error, FE_{t+1}.

The regression specification in equation (4) is non-linear in the regression coefficients β_1 and α_0^*. Hence, we conduct our statistical analysis using non-linear least squares. Specifically, we jointly estimate the following two equations using non-linear weighted least squares (see Mishkin 1983):

$$FE_{t+1} = \alpha_0 + \varepsilon_{t+1}$$

$$AR_{t+1} = \beta_1 [FE_{t+1} - \alpha_0^*] + \upsilon_{t+1} \tag{5}$$

The market efficiency hypothesis is then evaluated by testing the cross-equation restriction that $\alpha_0^* = \alpha_0$, while the naïve expectations hypothesis is evaluated by testing the restriction that $\alpha_0^* = 0$.

While non-linear least squares is the appropriate statistical technique for our tests, we also provide parallel tests using ordinary least squares (OLS) to illustrate the intuition behind our tests for readers who feel more comfortable with OLS. Our OLS tests are conducted by estimating the two equations in (5) using OLS.

$$FE_{t+1} = \alpha_0 + \varepsilon_{t+1}$$

$$AR_{t+1} = \beta_0 + \beta_1 FE_{t+1} + \upsilon_{t+1} \tag{5-OLS}$$

Comparing the abnormal return regression in equation (5-OLS) to the model in equation (4), we see that $\beta_0 = -\alpha_0^* \beta_1$. Hence, the market efficiency hypothesis implies that $\beta_0 = -\alpha_0 \beta_1$ (i.e., abnormal returns only respond to the unpredictable portion of the forecast error), while the naïve reliance hypothesis implies that $\beta_0 = 0$ (i.e., abnormal returns respond to the entire forecast error). Note that we cannot test the market efficiency restriction using OLS, because it is a non-linear cross-equation restriction (hence our original use of non-linear least squares). However, we can report the magnitudes of the OLS coefficients to illustrate the intuition behind the non-linear testing procedure.

Our second set of stock price tests examines the extent to which prices reflect information in the level of forecast growth about future forecast errors. Equation (2) above and the associated discussion indicate that forecast errors tend to be greater for firms with higher forecast growth. Substituting the forecast error prediction model in equation (2) for ε_t^* in equation (3) gives:

$$AR_{t+1} = \beta_1 \,[FE_{t+1} - \alpha_0^* - \alpha_1^* Growth_{t+1}] + \upsilon_{t+1} \qquad (6)$$

In this equation ($\alpha_0^* + \alpha_1^*$ $Growth_{t+1}$) represents the market's assessment of the average bias in analysts' long-term growth forecasts. The efficient market hypothesis predicts that α_0^* and α_1^* will correspond to their counterparts in the forecasting equation, α_0 and α_1 in equation (2). In other words, investors' expectations of future earnings growth, while based on analysts' forecast of future growth, rationally anticipate the average bias in analysts' long-term growth forecasts. Thus, stock prices respond only to the unpredictable portion of the analysts' forecast error, ε_{t+1}, which is equal to ($FE_{t+1} - \alpha_0 - \alpha_1$ $Growth_{t+1}$). The naïve reliance hypothesis predicts that α_0^* and α_1^* in equation (6) equal zero since investors believe that analysts' long-term growth forecasts are without bias. Under this hypothesis, investors' expectation of future earnings growth is equal to analysts' growth forecast. Thus, stock prices respond to the entire forecast error, FE_{t+1}.

We again conduct our statistical tests by estimating equations (2) and (6) simultaneously using non-linear least squares.

$$FE_{t+1} = \alpha_0 + \alpha_1 Growth_{t+1} + \varepsilon_{t+1}$$

$$AR_{t+1} = \beta_1 [FE_{t+1} - \alpha_0^* - \alpha_1^* Growth_{t+1}] + \upsilon_{t+1} \qquad (7)$$

The market efficiency hypothesis is then evaluated by testing the cross-equation restrictions that $\alpha_0^* = \alpha_0$ and $\alpha_1^* = \alpha_1$, while the naïve expectations hypothesis is evaluated by testing the restrictions that $\alpha_0^* = 0$ and $\alpha_1^* = 0$. We also report results using OLS by estimating the two equations in (7) using OLS.

$$FE_{t+1} = \alpha_0 + \alpha_1 Growth_{t+1} + \varepsilon_{t+1}$$

$$AR_{t+1} = \beta_0 + \beta_1 FE_{t+1} + \beta_2 Growth_{t+1} + \upsilon_{t+1} \qquad \text{(7-OLS)}$$

Comparing the abnormal return regression in equation (7-OLS) to the model in equation (6), we see that $\beta_0 = -\alpha_0^* \beta_1$ and $\beta_2 = -\alpha_1^* \beta_1$. Hence, the market efficiency hypothesis implies that $\beta_0 = -\alpha_0 \beta_1$ and $\beta_2 = -\alpha_1 \beta_1$ (i.e., abnormal returns only respond to the unpredictable portion of the forecast error), while the naïve reliance hypothesis implies that $\beta_0 = 0$ and $\beta_2 = 0$ (i.e., abnormal returns respond to the entire forecast error). Note again that the market efficiency restrictions cannot be tested using OLS, because they are non-linear cross-equation restrictions. However, we report the magnitudes of the OLS regression coefficients to highlight the intuition behind our non-linear tests.

IV. SAMPLE FORMATION AND VARIABLE MEASUREMENT

We require the following information to test our predictions: data on common stock offerings including the names of the lead managers of the offerings; analysts' long-term forecasts of earnings growth and the names of the firms for whom the analysts work; realized earnings growth; and stock returns. Details concerning the common stock offerings are obtained from the Securities Data Company, Inc. (SDC). Analysts' long-term forecasts of earnings growth and the names of their employers are obtained from Institutional-Broker-Estimates-System (I/B/E/S). Realized earnings growth rates are calculated using earnings data from Compustat. Monthly stock returns are obtained from the Center for Research in Security Prices (CRSP).

Table 1 summarizes our sample formation. We extract from SDC a total of 7,636 common stock underwritten offerings made between 1981-1990. This sample period is chosen for two reasons. First, 1981 is the first year in which I/B/E/S consistently provides analysts' estimates of long-term earnings growth forecasts. Second, to calculate analysts' forecast errors, we require five years of future realized growth in earnings. Thus, the final year in the sample is 1990.

We require firms to be covered on CRSP, Compustat and I/B/E/S and to have sufficient stock return and earnings data to examine their post-offering performance. We also require at least one long-term forecast within the 12 month window (-9 to +3) surrounding the issue date of the equity offering. As detailed in table 1, we lose 1,723 firm-offerings because the issuing firm is not covered on CRSP or Compustat. An additional 218 observations are lost because the issuing firm is not covered by I/B/E/S. These observations tend to be initial public offerings by small market capitalization firms not listed on major exchanges (stocks trading on pink sheets). An additional 3,165 firm-offerings are lost because of insufficient stock return or earnings data on CRSP and Compustat. Exclusion of these observations is likely to create a survivorship bias, which may explain the less dramatic post-offering under performance for our sample compared to the under performance documented in prior research. Finally, long-term growth forecasts are unavailable within our window for 1,351 firm-offerings. These restrictions result in a final sample of 1,179 firm-offerings made by 1,006 firms, only one-fifth of the total number of equity offerings made in the sample period 1981-1990.[16] However, the offerings we examine account for 30 percent of the total dollar value of all equity issued during this time period. Further, for each calendar year the median asset value of firms in our sample falls in the top two to four size deciles on Compustat. Thus, the sample examined is of economic significance.

For our final sample of 1,179 firm-offerings, we have 7,169 analysts' long-term earnings growth forecasts within the 12 months (-9 to +3) surrounding the issue dates. Using the names of lead managers obtained from SDC and the names of analysts' employers obtained from I/B/E/S, we categorize individual analysts as either *affiliated* or *unaffiliated* with a particular firm offering. If the analyst is employed by the investment bank acting as the lead manager for the offering (or if the analysts is employed by a subsidiary or the parent of the investment bank), then the analyst is classified as affiliated. We classify 622 analysts' forecasts as affiliated and 6,547 as unaffiliated.

SDC also provides information on the fee paid to the underwriters of each equity offering. The underwriting fee is shared by the lead manager, the co-managers, and the syndicate or selling group of the offering. Since we define affiliated analysts as those analysts employed by the lead manager, we examine the portion of the underwriting fee that is paid to the lead manager. The fee basis (Fee) is calculated as the fee paid to the lead underwriter divided by the total dollar value of the equity offering.

We measure post-offering stock price performance using five-year market-adjusted buy-hold stock returns. To ensure that all analysts' forecasts are known prior to the stock return cumulation period, we begin the cumulation period three months after the equity offering. The existence of negative abnormal stock returns following equity offerings has

[16] These constraints eliminate all but 86 initial public offerings from the final sample. The tenor of the results is unchanged if the 86 IPOs are excluded from the analysis.

been shown to be robust with respect to a wide variety of CAPM-based models for measuring abnormal returns (see for example Loughran and Ritter, 1995). We therefore expect to learn little from repeating our analysis for a variety of abnormal return measures. We do note, however, that we explicitly avoid using empirically motivated pricing models, such as the three-factor model suggested in Fama and French (1993). We avoid such models because their ability to predict future stock returns may be attributable to naive expectations about future profitability. In other words, while the 'size' and 'market-to-book' factors may be systematically associated with stock returns following equity offerings, we seek to determine whether the lower stock returns can be <u>explained</u> by naive earnings expectations. [17] Since these factors are empirically motivated, they do not, in and of itself, provide a satisfactory explanation for the size and market-to-book effects in stock prices.

We follow the I/B/E/S procedure for computing five-year annualized growth rates in earnings. This consists of fitting a least squares growth line to the logarithms of six annual earnings observations, beginning with the earnings observation immediately preceding the equity offering. We chose not to use a discrete annualized geometric growth rates because these rates can be extremely volatile when the base year is close to zero or when the base year or final year in the series contains significant nonrecurring items. Fitting a least squares regression line avoids placing excessive weight on the first and last observations in the growth period, resulting in less volatile growth estimates especially when these years include substantial nonrecurring items.[18] Negative earnings

[17] Brav and Gompers (1997) question whether the long-run under performance of initial public offerings is a unique anomaly or simply another manifestation of the Fama and French (1992, 1993) market-to-book, size anomaly. They document that the IPO anomaly is most pronounced for small firms with high market-to-book ratios. In this paper, we attempt to provide empirical evidence concerning <u>why</u> the anomaly exists. It is useful to note, however, that small firms with high-market-to-book ratios have the highest long-term growth forecasts and the largest analyst forecast errors.

[18] We use Compustat data item 18, earnings before extraordinary items, to minimize the effect of nonrecurring items. The results of this paper were also replicated using (i) operating income before special items after taxes (compustat data items #178 - #15 – #16 + #17) and (ii) I/B/E/S historical EPS growth reported in the fifth year following the equity offerings. The tenor of the results does not change using

values are set to missing, and if earnings are missing for either the first or last year of the six-year series, then we set the growth measure to missing.

V. RESULTS

Descriptive Statistics

Tables 2 and 3 provide descriptive details of our sample and an overview of our results. Formal statistical tests of our hypotheses are provided in tables 4 through 8. Panel A of table 2 provides means of analysts' forecasts and realized performance for our full sample of 7,169 analysts' long-term earnings growth forecasts. The mean abnormal stock return for the entire sample is -12.7 percent for the five years following the offering. This is a substantially less negative post-offering return than the -41.6 percent and -32.9 percent reported by Loughran and Ritter (1995) for initial public offerings and seasoned equity offerings, respectively.[19] One reason for this difference is the survivorship bias introduced by requiring our sample firms to have five years of earnings and stock return data following the equity offerings. Additionally, only firms followed by analysts are in our final sample and firms followed by multiple analysts are represented multiple times in the computation of the means.[20] Analyst following tends to be positively correlated with firm size (Bhushan, 1989) and smaller firms have the lowest post-offering abnormal stock price performance (Spiess and Affleck-Graves, 1995). Nevertheless, the long-run under performance of stock prices following equity offerings is clearly present in our sample.

---

these alternative measures of realized growth. Thus, our results are not driven by analysts' failure to anticipate nonrecurring, special items.

[19] Loughran and Ritter (1995) note that the measurement of the long-run under performance of issuing firms is sensitive to the benchmark employed. If the NASDAQ value-weighted index is used instead of the CRSP NYSE-AMEX value-weighted index, they report post-offerings returns of -29.0% and -19.5% for initial public offerings and seasoned equity offerings, respectively.

[20] The total number of offerings represented in the sample is 1,179. If each offering receives equal weighting in the mean, the mean abnormal return declines to –18 percent.

The mean realized growth in earnings for the full sample over the five years following the offering is 5.7 percent. The corresponding mean forecast growth in earnings at the time of the offering is 16.2 percent. On average, the forecast error in the five-year earnings growth forecasts is -10.6 percent. Analysts tend to over-estimate earnings growth by greater than 10 percent per year in the five years following equity offerings.[21] The negative abnormal returns in the five years following the offering are consistent with investors having overly optimistic expectations of earnings growth. Later in the paper, we demonstrate that the magnitudes of the earnings growth expectations implicit in stock prices are similar to the growth forecasts issued by analysts.

Panel B of table 2 stratifies the sample by analyst affiliation. All analyst forecasts fall into one of four categories:

(i) Affiliated Analysts - Pure Deals, the forecast is made by an analyst who is affiliated with the lead underwriter of the offering and there are no long-term forecasts made by unaffiliated analysts;

(ii) Affiliated Analysts - Mixed Deals, the forecast is made by an analyst who is affiliated with the lead underwriter of the offering and there are also long-term forecasts made by unaffiliated analysts;

(iii) Unaffiliated Analysts - Mixed Deals, the forecast is made by an analyst who is unaffiliated with the lead underwriter of the offering and there are also long-term forecasts made by affiliated analysts; and

[21] To assess whether a systematic bias exists in analysts' long-term growth forecasts during our sample period that is not associated with new issues, we collect all long-term growth forecasts found on I/B/E/S between the years 1981-1990. Comparing offer and non-offer years, the mean forecast growth is significantly higher for offering years, while the realized five-year earnings growth is significantly lower. Thus, while in general analysts over estimate growth rates for all firm-years on I/B/E/S during the period 1981-1990, analysts are significantly more overly optimistic in years in which firms issue equity. It is interesting to note, however, that pooling across offer and non-offer years, analysts' optimism does not differ significantly for issuing versus non-issuing firms.

(iii) Unaffiliated Analysts - Pure Deals, the forecast is made by an analyst who is unaffiliated with the lead underwriter of the offering and there are no long-term forecasts made by affiliated analysts.

The first category of Affiliated Analysts - Pure Deals consists of only 131 forecasts. The mean abnormal return for this sample is -32.3 percent, which is much more negative than the average returns for the entire sample, -12.7 percent. The forecast errors are also larger for this sample. The mean forecast error for Affiliated Analyst - Pure Deals is -14.8 percent, while the forecast error for the entire sample is -10.6 percent. These results are consistent with the affiliated analysts issuing more overly optimistic earnings growth forecasts and with investors sharing these overly optimistic earnings expectations. The statistics show a similar pattern for the 491 forecasts in the Affiliated Analysts - Mixed Deals category. The mean abnormal return is -21.3 percent, which is more negative than the average for the entire sample, and the mean forecast error is -14.3 percent, which is also more negative than the average for the entire sample.

The deals followed by unaffiliated analysts have the least negative abnormal returns and the least biased forecasts. For the 2,938 forecasts in the Unaffiliated Analysts - Mixed Deals category, the mean abnormal return is -12.3 percent and the mean forecast error is -10.5 percent.[22] For the 3,609 deals in the Unaffiliated Analysts - Pure Deals category, the mean abnormal return is -11.3 percent and the mean forecast error is -10.0 percent. This is consistent with the unaffiliated analysts issuing relatively less overly optimistic earnings growth forecasts and with investors sharing these less overly optimistic earnings expectations.

[22] The Affiliated Analysts-Mixed Deals and Unaffiliated Analysts-Mixed Deals represent the same underlying set of deals. The reason stock returns are more negative for the affiliated analysts is that the ratio of affiliated to unaffiliated analysts tends to be larger for the deals with more negative abnormal stock returns.

Table 3 reports the number of observations, mean abnormal returns, and mean forecast errors for the sample stratified by both analyst affiliation and forecast growth. Forecast errors tend to be larger for firms with higher forecast growth. Stratifying the sample by forecast growth therefore provides a further opportunity to examine the relation between variation in forecast errors and variation in abnormal returns.

Growth portfolios are formed by ranking all analysts' long-term growth forecasts and assigning observations in equal numbers to three portfolios (low, medium, and high) based on these rankings. If long-term growth forecasts are correlated with analyst affiliation, then the number of observations in each forecast growth portfolio will not necessarily be equally proportioned across sub-samples. This is illustrated in panel A of table 3. The affiliated analysts tend to be concentrated in the high forecast growth portfolio, with between 47 to 60 percent of observations being in this portfolio. The unaffiliated analysts tend to be more evenly distributed across the three forecast growth portfolios with between 31 to 33 percent being in the high growth portfolio.[23]

Panel B of table 3 reports the mean forecast errors for the affiliation and forecast growth sub-samples. Within analyst affiliation categories, the forecast errors are consistently more negative in the high forecast growth portfolio. Within the high forecast growth portfolios, the forecast errors are also consistently more negative for the affiliated analysts than for the unaffiliated analysts. Thus, analysts' over-optimism is most pronounced for the high growth portfolios, and within the high growth portfolio, affiliated analysts make the most overly optimistic forecasts. These regularities are mirrored in the mean abnormal returns reported in panel C of table 3. Firms in the high forecast growth portfolios experience the greatest long-run under performance, and within the high growth portfolios, the abnormal stock returns are consistently more negative for

[23] A 2x3 chi-square test comparing the distribution of affiliated analysts to the distribution of unaffiliated analysts across the forecast growth portfolios rejects the null that portfolio assignment is unrelated to analyst affiliation at the 0.0001 level.

affiliated analysts deals than for the unaffiliated analysts deals. Thus, firms' long-term stock price under performance is greatest when affiliated analysts project high earnings growth.

Overall, the descriptive evidence presented in table 3 indicates that analysts' long-term growth forecasts are the most overly optimistic when they are high and when they are made by affiliated analysts. The earnings expectations embedded in stock prices incorporate a similar pattern of forecast errors. We provide more formal statistical tests of these propositions later in the paper.

Tests of Bias in Analysts' Long-Term Earnings Growth Forecasts

Table 4 provides statistical tests of the differences in the forecast errors for the affiliated and unaffiliated analysts. We have no specific predictions concerning differential biases for the pure and mixed deals. We therefore combine forecasts for pure and mixed deals for both the affiliated and unaffiliated categories to increase the power of our statistical tests. Panel A of table 4 presents the distribution of forecast errors. Recall from table 2 that the mean forecast error for the entire sample is -10.6 percent. Panel A reveals that the forecast errors for affiliated analysts are consistently more negative than for the unaffiliated analysts. The mean forecast error for the affiliated analysts is -14.4 percent, while the mean forecast error for the unaffiliated analysts is -10.3 percent. A t-test for difference in means rejects the null of equality (p-value = 0.003), confirming our prediction that affiliated analysts tend to issue more optimistic long-term earnings growth forecasts. Panel A also reveals that the larger negative mean forecast error for affiliated (versus unaffiliated) analysts is driven by their over-optimistic forecasts of growth (p-value of 0.000) and not by lower growth realizations for firms they follow (p-value of 0.956).

To examine the sensitivity of the forecast errors to the growth expectation, in panel B of table 4 we estimate the regression of forecast errors on forecast growth in earnings. For the entire sample the regression results are similar to those reported by Dechow and Sloan (1997). The intercept is close to zero and the coefficient on forecast growth in earnings is -0.678. These coefficients indicate that realized growth in earnings is only about one-third of forecast growth in earnings. This, in turn, indicates that analysts' over-optimism is greater for firms with greater growth prospects. The results for the unaffiliated analysts are similar to the results for the entire sample. However, the results for the affiliated analysts indicate that while the intercept remains indistinguishable from zero, the coefficient on forecast growth in earnings falls to -0.832. This coefficient indicates that realized growth in earnings is only about one-sixth of forecast growth in earnings for forecasts issued by affiliated analysts. A Chow-test rejects the null hypothesis that the coefficient on forecast earnings growth is the same in the affiliated and unaffiliated regressions (p-value = 0.048). Thus, the over-optimism in affiliated analysts' growth forecasts, relative to unaffiliated analysts' growth forecasts, is more severe for 'glamour stocks' with high growth prospects.[24]

In table 5 we investigate whether the level of the affiliated analysts' growth forecasts, as well as the optimistic bias in their forecasts, is positively related to the fees paid to their employers (the lead underwriters of the equity offerings). Panel A documents a positive relation between the affiliated analysts' growth forecasts and the fee basis paid to their employers. Recall that the fee basis is the percentage of the dollar value of the offering paid to the lead manager. For each 100 basis points paid to the lead manager, analysts' growth forecasts increase by 650 basis points (6.5 percentage points).[25] Including

[24]To control for firm size, we also included the log of total assets as an additional explanatory variable in the regressions presented in tables 4 as well as the regressions presented below in tables 6 and 7. The tenor of the results remains unchanged.

[25]Although we do not report these results, it is interesting to note that there is a significantly higher fee basis paid for the affiliated versus the unaffiliated deals. In particular, when an analyst working for the lead manager provides a forecast of long-term earnings growth, the average fee basis paid to the lead

realized growth as an additional explanatory variable demonstrates that affiliated analysts' growth forecasts are unrelated to the level of the realized growth in earnings, while their forecasts remain positively related to the fees basis paid to their employers.

Panel B of table 5 examines whether the optimistic bias in affiliated analysts' growth forecasts is significantly related to the fee basis paid to their employers. The first regression shows that analysts' forecast errors are a function of the fee basis - affiliated analysts are more optimistic in their growth forecasts, the higher the fee basis. For each 100 basis points paid to the lead manager, affiliated analysts' over-estimate earnings growth by 470 basis points (4.7 percentage points). However, the fee basis explains very little of the cross-sectional variation in analysts' forecast errors (the adjusted R^2 is only 0.4 percent). Further, the coefficient on the fee basis falls to less than 100 basis points and becomes insignificant after controlling for the level of analysts' growth forecasts, as indicated in regression 2 in panel B of table 5. Comparing the R^2s reported in the B panels of tables 4 and 5, we see that the fee basis variable adds little to our understanding of analysts' forecast errors. In fact, the adjusted R^2 falls from 10.52 in table 4 to 10.38 in table 5 when Fee is included as an additional explanatory variable. Thus, in our stock price tests, presented in the next section, we do not consider the predictive power of the fee basis in explaining stock returns.

Tests of the Pricing of Bias in Analysts' Long-Term Earnings Growth Forecasts

In this section, we investigate whether the systematic bias in analysts' forecasts of earnings growth is reflected in stock prices. We first present our non-linear weighted least squares regressions, since these regressions allow us to conduct statistical tests of

manager is one percent of the total dollar-value of the offering. However, when only unaffiliated analysts provide growth forecasts, the average fee basis paid to the lead underwriter is 0.71 percent. The mean fee bases are significantly different at the 0.001 level.

the non-linear restrictions implied by our hypotheses. We then present the corresponding OLS regression results.

Since multiple analysts' forecasts can relate to a single equity offering, we conduct our pricing analysis using a single 'consensus' observation for each offering, in order to avoid cross-sectional dependence.[26] The forecast of growth in earnings used for each observation is the mean of the forecasts relating to the offer. We conduct the pricing tests for three samples. The first sample consists of all firm-offerings represented by the entire sample of analysts' forecasts. This sample provides a check for consistency between our results and the results in Dechow and Sloan (1997) and also provides a benchmark for our subsequent samples. The second sample includes all firm-offerings for which we have an affiliated analyst forecast, and includes both Affiliated Analysts - Pure Deals and Affiliated Analysts - Mixed Deals from table 2. In computing the mean forecast for the mixed deals, we exclude forecasts made by unaffiliated analysts. This sample allows us to examine whether affiliated analysts' forecasts are priced, irrespective of the availability of unaffiliated forecasts. The third sample includes all firm-offerings for which we have unaffiliated forecasts, and includes both Unaffiliated Analysts - Pure Deals and Unaffiliated Analysts - Mixed Deals. The computation of mean forecasts for the mixed deals excludes forecasts made by affiliated analysts. This sample allows us to examine whether unaffiliated analysts' forecasts are priced, irrespective of the availability of affiliated forecasts.[27]

Table 6 investigates the pricing of the consensus forecast errors by estimating the system of equations developed in section three:

---

[26] In addition, cross-sectional dependence may arise if some firms make more than one offering within five years. To address this concern, we re-estimated our results using only the first offering made by each firm. The tenor of the results is unchanged.

[27] Ideally, we would like to be able to take the sample of mixed deals and test whether the affiliated or the unaffiliated analysts' forecasts are priced. Unfortunately, because of the relatively small differences between the affiliated and unaffiliated forecasts on the mixed deals, we cannot statistically discriminate between these two alternatives.

$$FE_{t+1} = \alpha_0 + \varepsilon_{t+1}$$

$$AR_{t+1} = \beta_1 \, [FE_{t+1} - \alpha_0^*] + \upsilon_{t+1} \qquad (5)$$

Recall that α_0 represents the mean forecast error in analysts' forecasts of long-term earnings growth issued around the equity offerings. If investors rationally anticipate the bias in analysts' forecasts, then the earnings expectation embedded in stock prices will result in $\alpha_0^* = \alpha_0$. Alternatively, if investors naively rely on analysts' forecasts of long-term growth in earnings, then the earnings expectations embedded in stock prices will result in α_0^* being equal to zero. The system of equations is estimated jointly using nonlinear weighted-least-squares and the cross-equation restrictions implied by the competing hypotheses are tested using a likelihood ratio test.[28]

The nonlinear weighted least squares parameter estimates for the system of equations in (5) are reported in panel A of table 6. The estimate of the mean forecast error for all deals, α_0, is -12.1 percent. This figure differs slightly from the corresponding figure in table 2 of -10.6 percent. The reason for the difference is that observations are weighted at the individual analyst level in table 2 and at the individual issuer level (using consensus analyst forecasts) in table 6. The implied estimate of the mean forecast error in the stock price equation, α_0^*, is -2.9 percent.[29] This estimate is significantly different from the

---

[28] This joint estimation procedure has several advantages over a two-step procedure. First, it produces more efficient estimates of parameters because each equation uses information in the other in the estimation process. Second, the joint estimation procedure generates valid test statistics because it accounts for the uncertainty in estimates of the error terms. That is, the joint estimation procedure uses the ε that is *expected* to minimize the mean-square errors in the first equation to form expectations in the second equation, whereas the two-step procedure uses the actual ε that minimized the mean-square errors. Thus, in finite samples, the two-step procedure makes an overly strong assumption about expectation formation. Conceptually, the two-step procedure forms expectations with information from the future as well as from the past, which clearly goes beyond the rational expectations principle. The joint estimation procedure does not suffer from this problem. Details of the joint estimation procedure are described in Mishkin (1983).

[29]Verifying that the asymptotic equivalence of α_0^* and $(-\beta_0/\beta_1)$ holds for the three finite samples examined, we note that all estimates of α_0^* reported in table 6 are equivalent to the ratios of the OLS estimated coefficients $(-\beta_0/\beta_1)$ reported in table 7.

rational value of -12.1 percent (p-value = 0.001), rejecting market efficiency, but is not significantly different from the naive value of zero (p-value = 0.511). Thus, we are unable to reject the hypothesis that investors rely on analysts' forecasts of long-term growth in earnings as if the forecasts are unbiased. Another way of stating this result is that we are unable to reject the hypothesis that investors' naive reliance on analysts' forecasts potentially explains the under performance of stock prices following equity offerings.

The mean forecast errors for the affiliated and unaffiliated deals also differ slightly at the issuer level versus the analyst level. However, we continue to find the affiliated analysts are more overly optimistic than the unaffiliated analysts, with forecast errors of -14.3 percent and -11.8 percent, respectively. Despite the differences in the actual forecast errors, the implied estimates of the forecast errors in stock prices are quite similar. The forecast error implicit in stock prices for the affiliated deals is -2.7 percent, while for the unaffiliated deals it is -2.2 percent. In both cases, market efficiency is rejected, but in neither case is the hypothesis that investors naively rely on analysts' forecasts rejected. Thus, despite the difference in the magnitude of the forecast errors for affiliated and unaffiliated analysts' forecasts, we are unable to reject the hypothesis that they are both incorporated into stock prices.

We present the results of corresponding OLS regressions in order to illustrate the intuition behind the non-linear regression results. Recall from the discussion in section three that the corresponding OLS regressions for the system in panel A are:

$$FE_{t+1} = \alpha_0 + \varepsilon_{t+1}$$

$$AR_{t+1} = \beta_0 + \beta_1 FE_{t+1} + \upsilon_{t+1} \qquad \text{(5-OLS)}$$

Panel B of table 6 presents the results for the stock return regressions only. The forecasting regressions involves no non-linearities, and so the OLS results are identical to those reported using non-linear least squares in panel A. The first thing to note from the OLS stock return regressions in panel B is that β_1, the earnings response coefficient in the stock return regressions, is significantly positive and takes on the same value as its counterpart in the non-linear regressions. This equality again arises because this term involves no non-linearities. The positive coefficient reflects the well-documented positive response of stock returns to earnings surprises.

Non-linearities enter the equation with the coefficients that relate to our competing hypotheses. Recall that market efficiency predicts that stock returns will only respond to the unpredictable component of the forecast error (FE_{t+1} - α_0). Thus, the intercept in the OLS regression, β_0, will load up to remove the predictable component of the forecast error, (α_0), so that β_0 = $-\alpha_0\beta_1$. Concentrating on the *All Deals* column, market efficiency generates the prediction that $\beta_0 = -\alpha_0\beta_1 = -(-0.121) \times (1.171) = 0.142$. On the other hand, the naïve reliance hypothesis implies that the market responds to the entire forecast error (FE_{t+1}), so that $\beta_0 = 0$. The estimated value of β_0 is an insignificant 0.034, which is much closer to the naïve reliance hypothesis prediction of zero, than to the market efficiency hypothesis prediction of 0.142. At an intuitive level, the statistical tests in our non-linear regressions indicate that we can reject the market efficiency hypothesis that $\beta_0 = 0.142$, but we cannot reject the naïve reliance hypothesis that $\beta_0 = 0$. Similarly, we see that β_0 is also insignificantly different from zero in the remaining two columns in panel B of table 6, indicating that the naïve reliance hypothesis is not rejected for either the affiliated or unaffiliated sub-groups.

Table 7 investigates the pricing of analysts' forecast errors after conditioning on the level of forecast growth by estimating the following system of equations developed above in section three:

$$FE_{t+1} = \alpha_0 + \alpha_1 \, Growth_{t+1} + \varepsilon_{t+1}$$

$$AR_{t+1} = \beta_1 \, [FE_{t+1} - \alpha_0^* - \alpha_1^* Growth_{t+1}] + \upsilon_{t+1} \tag{7}$$

The nonlinear weighted least squares parameter estimates for this system of equations are reported in panel A of table 7. Recall from the analyst level results in table 4 that the forecast errors tend to be more optimistic for 'glamour stocks' with high levels of forecast growth. While the magnitude of the coefficients is somewhat different for the issuer-level results, we find the same general relations in table 7. For all deals, the coefficient on forecast growth in earnings, α_1, is -0.329. This indicates that realized growth is only about two-thirds as large as forecast growth. As with the results, in table 4, α_1 is more negative for the affiliated analysts, indicating that affiliated analysts tend to be even more overly optimistic than their unaffiliated counterparts for high growth stocks. If investors rationally anticipate this bias in analysts' forecasts, then the earnings expectation embedded in stock prices will result in $\alpha_1^* = \alpha_1$. Alternatively, if investors naively rely on analysts' forecasts of long-term growth in earnings, then the earnings expectations embedded in stock prices will result in $\alpha_1^* = 0$. The system of equations is again estimated jointly using nonlinear weighted-least-squares and the cross-equation restrictions implied by the competing hypotheses are tested using a likelihood ratio test.

The results indicate that the earnings expectations embedded in stock prices correspond much more closely with the naive reliance hypothesis than with the market efficiency hypothesis. The estimated values of α_1^* are 0.138, -0.032 and 0.380 for *all deals*, *affiliated deals* and *unaffiliated deals*, respectively. These compare to the predicted values of –0.329, -0.621 and –0.325, respectively, under the market efficiency hypothesis and predictions of zero in all three cases for the naïve reliance hypothesis. While the market efficiency hypothesis is rejected in all three cases, the naïve reliance hypothesis cannot be rejected in any of the three cases. It is of interest that the point estimates of α_1^*

are fairly large and positive for the *all deals* and *unaffiliated deals* categories. Under market efficiency, these coefficients are predicted to be negative, while under naive reliance, they are predicted to be zero. One explanation for why these coefficients are positive is that investors are pricing affiliated analysts' forecasts, and this induces a positive bias in the coefficients on the unaffiliated analysts' forecasts when both types of analyst forecasts are present. In other words, investors are pricing the affiliated analysts more overly optimistic forecasts, and because stock prices are responding to the greater optimism in affiliated analysts' forecasts, stock prices reflect more optimistic expectations than those implied by the unaffiliated analysts' forecasts.[30] Overall, the results in table 7 suggest that the poor stock price performance in the years following equity offerings arises because investors naively price affiliated analysts' extreme over-optimism for 'glamour stocks' with strong growth prospects.

The OLS results in panel B of table 7 serve to illustrate the intuition behind the non-linear regression results presented in panel A. Again, panel B presents the OLS regression results for the stock return regressions only, since OLS results for the forecasting regressions are identical to those presented in panel A. Recall that market efficiency predicts that stock returns will only respond to the unpredictable component of the forecast error, ($FE_{t+1} - \alpha_0 - \alpha_1 Growth_{t+1}$). Thus, β_2, the coefficient on forecast growth in the OLS regression will load up to remove the portion of the forecast error that can be predicted by forecast growth ($\alpha_1 Growth_{t+1}$), so that $\beta_2 = -\alpha_1 \beta_1$. Concentrating on the *All Deals* column, market efficiency generates the prediction that $\beta_2 = -\alpha_1 \beta_1 = -(-0.329) \times (1.163) = 0.383$. On the other hand, the naïve reliance hypothesis implies that the market responds to the entire forecast error, (FE_{t+1}), so that $\beta_2 = 0$. The estimated value of β_2 is –0.161, which is insignificantly different from the naïve reliance hypothesis

[30] Replicating the pricing tests for the deals with only unaffiliated analysts provides evidence consistent with this conjecture. The value of α_1^* falls from 0.380 reported in table 8 for the full sample to 0.134.

prediction of zero, but significantly different from the market efficiency hypothesis prediction of 0.383.

VI. CONCLUSION

In this study, we provide evidence consistent with the hypothesis that sell-side analysts make overly optimistic long-term earnings growth forecasts for firms issuing equity. We also show that the overly optimistic forecasts are reflected in stock prices. Together these results suggest that investors' reliance on analysts' overly optimistic forecasts provides one potential explanation for the 'equity issue puzzle.'

Our evidence has potential policy implications. Our evidence suggests that the co-existence of brokerage services and underwriting services in the same institution leads sell-side analysts to compromise their responsibility to brokerage clients in order to attract underwriting business. Investment banks claim to have 'Chinese walls' to prevent such conflicts of interests. Our evidence raises questions about the reliability of these 'Chinese walls.' We document that analysts affiliated with the lead underwriter of an offering tend to issue more overly optimistic growth forecasts than unaffiliated analysts. Further, the magnitude of the affiliated analysts' growth forecasts is positively related to the fee basis paid to the lead underwriters. Finally, equity offerings covered only by affiliated analysts experience the greatest post-offering under performance, suggesting that these offerings are the most overpriced.

Our results also suggest a characterization of over-priced 'glamour stocks' (Lakonishok, Shleifer and Vishny, 1994). 'Glamour stocks' tend to have high growth opportunities and therefore these firms actively seek new equity capital. Consequently, their management has incentives to maximize stock price to lower the cost of raising external capital. Sell-side analysts affiliated with investment banks also have incentives to assist with maximizing the issuing firm's stock price, because doing so generates higher underwriting fees. Under this scenario, the systematic over-optimism in the forecasts of unaffiliated analysts may result from their efforts to attract underwriting clients. An interesting topic for future research would be to examine whether firms' choices of investment banks for underwriting services are influenced by the optimism in the earnings forecasts issued by analysts affiliated with the investment banks.

REFERENCES

Abarbanell, J. S. 1991. Do analysts' earnings forecasts incorporate information in prior stock price changes? *Journal of Accounting and Economics* 14 (June): 147-65.

Ali, A. 1996. Bias in analysts' earnings forecasts as an explanation for the long-run underperformance of stocks following equity offerings. Working Paper. University of Arizona.

Barber, B. M., and J. D. Lyon, 1996. How can long-run abnormal stock returns be both positively and negatively biased? Working Paper. UC-Davis.

Barber, B. M., and J. D. Lyon, 1997. Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics* (March): 341-372

Bhushan, R. 1989. Firm characteristics and analysts following. *Journal of Accounting and Economics* 11(July): 255-274.

Bradshaw, M. T. 1999. Do analysts use their own forecasts? Working paper. University of Michigan.

Brav, A., and P. Gompers. 1997. Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies. *Journal of Finance* 52 (December): 1791-1821.

Brown, P., G. Foster and E. Noreen. 1985. Security analysts multi-year earnings forecasts and the capital market. *Studies in Accounting Research Number 21.* American Accounting Association.

Campbell, J. 1991. A variance decomposition of stock returns, *The Economic Journal* 101: 157-179.

Collins, D., S.P. Kothari, J. Shanken and R. Sloan. 1994. Lack of timeliness and noise as explanations for the low contemporaneous return-earnings association, *Journal of Accounting and Economics* 18: 289-324.

Dechow, P. M., and R. G. Sloan. 1997. Returns to contrarian investment: Tests of the naive expectations hypotheses, *Journal of Financial Economics* 43 (January): 3-28.

Dugar, A., and S. Nathan. 1995. The effect of investment banking relationships on financial analysts' earnings forecasts and investment recommendations. *Contemporary Accounting Research* 12 (Fall 1995): 131-60.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns, *Journal of Finance* 47 (June): 427-65.

Fama, E. F., and K. R. French. 1993. Common risk factors in the returns to stocks and bonds, *Journal of Financial Economics* 33 (February): 3-56.

Hansen, R. S., and A. Sarin. 1996. Is honesty the best policy? An examination of security analyst behavior around seasoned equity offerings. Working Paper. Santa Clara University.

Kothari, S. P., and J. B. Warner. 1997. Measuring long-horizon security price performance. *Journal of Financial Economics* (March): 301-339.

Lakonishok J., A. Shliefer; and R. W. Vishny. 1994. Contrarian investment, extrapolation, and risk. *Journal of Finance* 49 (December): 1541-1578.

La Porta, R. 1996. Expectations and the cross-section of stock returns. *Journal of Finance* 51 (December): 1715-1742.

Lin, H., and M. McNichols. 1998a. Underwriting relationships and analysts' forecasts and investment recommendations. *Journal of Accounting and Economics*. 25 (February): 101-127.

Lin, H., and M. McNichols. 1998b. Analyst coverage of initial public offering firms. Working Paper. Stanford University.

Loughran, T., and J. R. Ritter. 1995. The new issue puzzle. *Journal of Finance* 50 (March): 23-51.

Loughran, T., and J. R. Ritter. 1997. The operating performance of firms conducting seasoned equity offerings. *Journal of Finance* 52 (December): 1823-1850.

Mikhail, M. B., B. R. Walther, and R. H. Willis. 1999, Does forecast accuracy matter to security analysts?, *The Accounting Review* 74 (April): 185-200.

McNichols, M., and P. O'Brien. 1997. Self-selection and analyst coverage. *Journal of Accounting Research* 35 (Supplement): 167-199.

Michaely, R., and K. L. Womack. 1996. Conflict of interest and the credibility of underwriter analyst recommendations. Working Paper. Cornell University.

Mishkin, F. S. 1983. A rational expectations approach to macroeconomics, A National Bureau of Economics Research monograph. The University of Chicago Press. Chicago, Il.

Rajan, R., and H. Servaes. 1997. Analyst following of initial public offerings. *Journal of Finance* (June): 507-529.

Rangan, S. 1998. Earnings management and the performance of seasoned equity offerings. *Journal of Financial Economics* 50 (October): 101-122

Sougiannis, T., E. Oltheten, and T. Yaekura. 1996. Inferring equity values from analysts' earnings forecasts. Working Paper. University of Illinois at Urbana-Champaign.

Spiess, D. K, and J. Affleck-Graves. 1995. The long-run performance following seasoned equity offerings. *Journal of Financial Economics* 38 (July): 243-67.

Stickel, S. E. 1992. Reputation and performance among security analysts. *Journal of Finance*. 47 (December): 1811-1836.

Stickel, S. E., 1998. Analyst incentives and the financial characteristics of Wall Street darlings and dogs. Working paper. Lasalle University.

Teoh, S. H. , I Welch, and T. J. Wong. 1998. Earnings management and the underperformance of seasoned equity offerings. *Journal of Financial Economics* 50 (October): 63-99.

TABLE 1

Sample Formation

7,636	Common stock offerings within the time period 1981-1990 (underwritten offerings only, excludes rights offerings)
-1,723	Firm-offerings not covered on CRSP or Compustat
-218	Firm offerings not covered on I/B/E/S
-3,165	Firm-offerings with incomplete data on CRSP or Compustat
-1,351	Firm-offerings with no long-term earnings growth forecasts within the 12 months (-9 to +3) surrounding the offer date
1,179	Final sample of firm-offerings

Data on common stock offerings is obtained from the Securities Data Company, Inc.

TABLE 2

Profile of Analyst Forecasts and Post-Offering Performance

Panel A: Entire sample

	Mean Abnormal returns	*Mean Realized earnings growth*	*Mean Forecast earnings growth*	*Mean Forecast error*	*Number of observations*
All analysts	-0.127	0.057	0.162	-0.106	7,169

Panel B: Sample stratified by analyst affiliation

	Mean Abnormal returns	*Mean Realized earnings growth*	*Mean Forecast earnings growth*	*Mean Forecast error*	*Number of observations*
Affiliated analysts - Pure deals	-0.323	0.097	0.233	-0.148	131
Affiliated analysts -Mixed deals	-0.213	0.045	0.186	-0.143	491
Unaffiliated analysts -Mixed deals	-0.123	0.048	0.150	-0.105	2,938
Unaffiliated analysts -Pure deals	-0.113	0.064	0.165	-0.100	3,609

Abnormal returns = cumulative five-year buy and hold market adjusted stock returns beginning three months after the issue date. A value-weighted market index is used to adjust for market performance.

Realized earnings growth = five-year annualized growth rates calculated by fitting a least squares growth line to the logarithms of the six annual observations, beginning with the offer year and ending in the fifth year after the offer year.

Forecast earnings growth = analysts' long-term forecast of earnings growth obtained from I/B/E/S.

Forecast error = Realized earnings growth - Forecast earnings growth.

TABLE 3

Profile of Analyst Forecast Errors and Post-Offering Performance Stratified by Analyst Affiliation and Forecast Earnings Growth

Range of forecast earnings growth for each portfolio	*Firms where analysts predict LOW Growth* [-100% - 10%]	*Firms where analysts predict MEDIUM Growth* [11% - 17%]	*Firms where analysts predict HIGH growth* [18% - 100%]
Panel A: Number (percent) of analysts with forecasts			
Affiliated analysts -Pure deals Total number of analysts = 131	19 (15%)	33 (25%)	79 (60%)
Affiliated analysts - Mixed deals Total number of analysts = 491	119 (24%)	140 (29%)	232 (47%)
Unaffiliated analysts - Mixed deals Total number of analysts = 2,938	1,016 (35%)	996 (34%)	926 (31%)
Unaffiliated analysts - Pure deals Total number of analysts = 3,609	1,133 (31%)	1,298 (36%)	1,178 (33%)
All analysts (Total=7,169)	*2,287 (32%)*	*2,467 (34%)*	*2,415 (34%)*
Panel B: Mean Forecast error			
Affiliated analysts -Pure deals	-0.156	-0.108	-0.238
Affiliated analysts - Mixed deals	-0.061	-0.055	-0.239
Unaffiliated analysts - Mixed deals	-0.052	-0.058	-0.215
Unaffiliated analysts - Pure deals	-0.033	-0.091	-0.174
All analysts	*-0.042*	*-0.075*	*-0.198*
Panel C: Abnormal returns			
Affiliated analysts -Pure deals	-0.110	-0.224	-0.431
Affiliated analysts - Mixed deals	0.031	0.003	-0.501
Unaffiliated analysts - Mixed deals	0.002	-0.022	-0.398
Unaffiliated analysts - Pure deals	-0.098	-0.079	-0.166
All analysts	*-0.048*	*-0.054*	*-0.288*

Abnormal returns = cumulative five-year buy and hold market adjusted stock returns beginning three months after the issue date. A value-weighted market index is used to adjust for market performance.

Realized earnings growth = five-year annualized growth rates calculated by fitting a least squares growth line to the logarithms of the six annual observations, beginning with the offer year and ending in the fifth year after the offer year.

Forecast earnings growth = analysts' long-term forecast of earnings growth obtained from I/B/E/S.

Forecast error = Realized earnings growth - Forecast earnings growth.

TABLE 4

Tests of the Bias in Analysts' Forecasts of Long-Term Earnings Growth

Panel A: Distribution of variables for affiliated and unaffiliated analysts (p-values for tests of equal means)

		Mean	*Standard Deviation*	*Lower Quartile*	*Median*	*Upper Quartile*	*Number of Obs.*
Forecast Error	*Affiliated*	-0.144	0.340	-0.241	-0.079	0.012	622
	Unaffiliated	-0.103	0.287	-0.191	-0.059	0.024	6547
	p-value	0.003					
Forecast Earnings Growth	*Affiliated*	0.201	0.133	0.120	0.178	0.250	622
	Unaffiliated	0.159	0.114	0.090	0.130	0.200	6547
	p-value	0.000					
Realized Earnings Growth	*Affiliated*	0.056	0.322	-0.036	0.077	0.203	622
	Unaffiliated	0.057	0.280	-0.026	0.071	0.183	6547
	p-value	0.956					

Panel B: Sensitivity of forecast errors to forecast of long-term earnings growth

$$FE_{t+1} = \alpha_0 + \alpha_1 \text{ Growth}_t + \varepsilon_{t+1}$$

	α_0	α_1	*p-value for equal* α_1	*Adjusted* R^2 *(%)*	*Number of observations*
Entire sample - All analysts	0.004	-0.678**		7.33	7,169
Total affiliated analysts	0.023	-0.832**		10.52	622
			0.048		
Total unaffiliated analysts	0.002	-0.654**		6.79	6,547

** significant at the one percent level.

Realized earnings growth = five-year annualized growth rates calculated by fitting a least squares growth line to the logarithms of the six annual observations, beginning with the offer year and ending in the fifth year after the offer year.

Forecast earnings growth (Growth $_{t+1}$) = analysts' long-term forecast of earnings growth obtained from I/B/E/S.

Forecast error (FE $_{t+1}$) = Realized earnings growth - Forecast earnings growth.

TABLE 5

Tests of the Relation Between Affiliated Analysts' Forecasts of Long-Term Earnings Growth and the Underwriting Fee Paid to Their Employers

Panel A: Sensitivity of affiliated analysts' forecast of long-term earnings growth to the fee paid to the lead manager and realized earnings growth

$$Growth_t = \gamma_0 + \gamma_1 \, Fee_t + \gamma_2 \, R\_Growth_{t+1} + \mu_t$$

	γ_0	γ_1	γ_2	*Adjusted R^2 (%)*	*Number of observations*
Regression 1	0.134**	0.065**		6.80	622
Regression 2	0.134**	0.064**	0.026	7.01	622

Panel B: Sensitivity of affiliated analysts' forecast errors to the fee paid to the lead manager and the forecast of long-term earnings growth

$$FE_{t+1} = \alpha_0 + \alpha_1 \, Growth_t + \alpha_2 \, Fee_t + \varepsilon_{t+1}$$

	α_0	α_1	α_2	*Adjusted R^2 (%)*	*Number of observations*
Regression 1	-0.096**		-0.047*	0.40	622
Regression 2	0.016	-0.840**	0.008	10.38	622

** significant at the one percent level.
* significant at the six percent level.
Realized earnings growth (R_Growth_{t+1}) = five-year annualized growth rates calculated by fitting a least squares growth line to the logarithms of the six annual observations, beginning with the offer year and ending in the fifth year after the offer year.
Forecast earnings growth ($Growth_{t+1}$) = analysts' long-term forecast of earnings growth obtained from I/B/E/S.
Forecast error (FE_{t+1}) = Realized earnings growth - Forecast earnings growth.
Fee basis (Fee_t) = Fee paid to the lead manager divided by the total dollars raised in the offering.

TABLE 6

Results of Nonlinear Weighted and Ordinary Least Squares Regressions Examining the Pricing of the Systematic Bias in Analysts' Forecasts of Long-Term Earnings Growth. Forecast Errors are Conditioned on Proximity to Equity Offerings.

Panel A: Non-linear weighted least squares

$$FE_{t+1} = \alpha_0 + \varepsilon_{t+1}$$
$$AR_{t+1} = \beta_1 [Fe_{t+1} - \alpha_0^*] + \upsilon_{t+1}$$

	All deals	*All deals with affiliated analysts*	*All deals with unaffiliated analysts*
α_0	-0.121**	-0.143**	-0.118**
α_0^*	-0.029	-0.027	-0.022
β_1	1.171**	1.254**	1.102**
Tests of the cross-equation restrictions (p-value for likelihood ratio tests)			
Market efficiency ($\alpha_0^* = \alpha_0$)	0.001	0.001	0.001
Naïve expectations ($\alpha_0^* = 0$)	0.511	0.674	0.602
Number of observations	1179	440	1070

Panel B: Ordinary least squares

$$AR_{t+1} = \beta_0 + \beta_1 FE_{t+1} + \upsilon_{t+1}$$

	All deals	*All deals with affiliated analysts*	*All deals with unaffiliated analysts*
β_0	0.034	0.034	0.024
β_1	1.171**	1.254**	1.102**
Adjusted R^2	0.070	0.089	0.066
Number of observations	1179	440	1070

** significant at the one percent level.

Abnormal returns (AR_{t+1}) = cumulative five-year buy and hold market adjusted stock returns beginning three months after the issue date. A value-weighted market index is used to adjust for market performance.

Realized earnings growth = five-year annualized growth rates calculated by fitting a least squares growth line to the logarithms of the six annual observations, beginning with the offer year and ending in the fifth year after the offer year.

Forecast earnings growth (Growth $_{t+1}$) = analysts' long-term forecast of earnings growth obtained from I/B/E/S.

Forecast error (FE_{t+1}) = Realized earnings growth - Forecast earnings growth.

TABLE 7

Results of Nonlinear Weighted and Ordinary Least Squares Regressions Examining the Pricing of the Systematic Bias in Analysts' Forecasts of Long-Term Earnings Growth. Forecast Errors are Conditioned on Proximity to Equity Offerings and the Level of Forecast Earnings Growth.

Panel A: Non-linear weighted least squares

$$FE_{t+1} = \alpha_0 + \alpha_1 Growth_{t+1} + \varepsilon_{t+1}$$
$$AR_{t+1} = \beta_1 [FE_{t+1} - \alpha_0^* - \alpha_1^* \; Growth_{t+1}] + \upsilon_{t+1}$$

	All deals	All deals with affiliated analysts	All deals with unaffiliated analysts
α_0	-0.058**	-0.018	-0.058**
α_1	-0.329**	-0.621**	-0.325**
α_0^*	-0.054	-0.021	-0.091
α_1^*	0.138	-0.032	0.380
β_1	1.163**	1.257**	1.084**
Tests of the cross-equation restrictions (p-value for likelihood ratio tests)			
Market efficiency ($\alpha_0^* = \alpha_0$ *and* $\alpha_1^* = \alpha_1$)	0.001	0.001	0.001
Naïve expectations ($\alpha_0^* = 0$ *and* $\alpha_1^* = 0$)	0.516	0.894	0.489
Number of observations	1179	440	1070

Panel B: Ordinary least squares

$$AR_{t+1} = \beta_0 + \beta_1 FE_{t+1} + \beta_2 Growth_{t+1} + \upsilon_{t+1}$$

	All deals	*All deals with affiliated analysts*	*All deals with unaffiliated analysts*
β_0	0.063	0.027	0.098
β_1	1.163**	1.257**	1.084**
β_2	-0.161	0.040	-0.412
Adjusted R^2	0.069	0.086	0.066
Number of observations	1179	440	1070

** significant at the one percent level.

Abnormal returns (AR_{t+1}) = cumulative five-year buy and hold market adjusted stock returns beginning three months after the issue date. A value-weighted market index is used to adjust for market performance.

Realized earnings growth = five-year annualized growth rates calculated by fitting a least squares growth line to the logarithms of the six annual observations, beginning with the offer year and ending in the fifth year after the offer year.

Forecast earnings growth ($Growth_{t+1}$) = analysts' long-term forecast of earnings growth obtained from I/B/E/S.

Forecast error (FE_{t+1}) = Realized earnings growth - Forecast earnings growth.

The U.S. Equity Return Premium: Past, Present and Future

J. Bradford DeLong and Konstantin Magin

February 2008

J. Bradford DeLong is Professor of Economics, University of California at Berkeley, Berkeley, California; Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts; and Visiting Scholar, Federal Reserve Bank of San Francisco, San Francisco, California.

Konstantin Magin is visiting professor at the University of California at Berkeley, Berkeley, California.

Their e-mail addresses are <delong@econ.berkeley.edu> and <magin@berkeley.edu>, respectively.

We would like to thank Bob Anderson, Nick Barberis, Roger Craine, Christopher DeLong, Milton Friedman, Jim Hines, Chad Jones, Ulrike Malmendier, Rajnish Mehra, Andrei Shleifer, Timothy Taylor, Robert Waldmann, Marty Weitzman, and especially Jeremy Stein for helpful comments and discussions. Part of this research was supported by the National Science Foundation. The opinions expressed are those of the authors alone, and not of the Federal Reserve Bank of San Francisco or any other institution.

ABSTRACT

For more than a century, diversified long-horizon investors in America's stock market have invariably received much higher returns than investors in bonds: a return gap averaging some six percent per year that Rajnish Mehra and Edward Prescott (1985) labeled the "equity premium puzzle." The existence of this equity return premium has been known for generations: more than eighty years ago financial analyst Edgar L. Smith (1924) publicized the fact that long-horizon investors in diversified equities got a very good deal relative to investors in debt: consistently higher long-run average returns with less risk. As of this writing— October 16, 2007, 11.44 PDT—the annual earnings yield on the value-weighted S&P composite index is 5.53%. This is a wedge of 3.22 % per year when compared to the annual yield on 10-year Treasury inflation-protected bonds of 2.31%. The existence of the equity return premium in the past offered long-horizon investors a chance to make very large returns in return for bearing little risk. It appears likely that the current configuration of market prices offers a similar opportunity to long-horizon investors today.

I. Introduction

For more than a century, diversified long-horizon investors in America's stock market have invariably received much higher returns than investors in bonds: a return gap averaging some six percent per year that Rajnish Mehra and Edward Prescott (1985) labeled the "equity premium puzzle." The existence of this equity return premium has been known for generations: more than eighty years ago financial analyst Edgar L. Smith (1924) publicized the fact that long-horizon investors in diversified equities got a very good deal relative to investors in debt: consistently higher long-run average returns with less risk. It was true, Smith wrote three generations ago, that each individual company's stock was very risky: "subject to the temporary hazard of hard times, and [to the hazard of] a radical change in the arts or of poor corporate management." But these risks could be managed via diversification across stocks: "effectively eliminated through the application of the same principles which make the writing of fire and life insurance policies profitable."

Edgar L. Smith was right.

Common stocks have consistently been extremely attractive as long-term investments.

Over the half century before Smith wrote, the Cowles Commission index of American stock prices deflated by consumer prices shows an average real return on equities of 6.5 percent per year— compared to an average real long-term government bond return of 3.6 percent and an average real bill return of 4.5 percent.[1] Since the start of the twentieth century, the Cowles Commission index linked to the Standard and Poor's Composite shows an average real equity return of 6.0 percent per year, compared to a real bill return of 1.6 percent per year and a real long-term government bond return of 1.8 percent per year. Since World War II equity returns have averaged 6.9 percent per year, bill returns 1.4 percent per year, and bond returns 1.1 percent per year. Similar gaps between stock and bond and bill returns have typically existed in other economies. Mehra (2003)[2] reports an annual equity return premium of 4.6 percent in post-World War II Britain, 3.3 percent in Japan since 1970, and 6.6 percent and 6.3 percent respectively in Germany and Britain since the mid-1970s.

Edgar Smith was right about both his past and our past. It appears likely[3] that Smith is right about our future as well. The arguments that the equity return premium should not be a puzzle in the future appear to imply that the equity return premium should not have existed in the past, yet it did.

The equity return premium has existed in the American stock market since it consisted of

---

[1]In the data set of Robert Shiller (2006): http://www.econ.yale.edu/~shiller/data.htm.

[2]Citing Jeremy Siegel (1998) and John Campbell (2001).

[3]Along with Rajnish Mehra (2006).

a few canal and railroad companies and John Jacob Astor's fur-trading empire. Its existence has been broadly known for 80 years. It is one of the most durable macroeconomic facts in the economy. Thus it appears overwhelmingly likely that the equity return premium has a future as well as a past, and there is little or no apparent reason for us economists to believe that in this case we economists know better than the market.

II. The Arithmetic of the Equity Premium

To pose the equity premium return puzzle, consider a marginal investor with a 20-year horizon—somebody in elementary school receiving a bequest from grandparents, somebody in their 30s with children putting money away to spend on college, somebody age 50 contemplating medical bills or wanting to leave a bequest, a life-insurance company collecting premiums from the middle-aged, or a company offering its workers a defined-benefit pension.

One margin such an investor must consider is the choice between:

(1) investing in a diversified portfolio of equities, reinvesting payouts and rebalancing periodically to maintain diversification;

(2) investing in short-term safe bills, rolling the portfolio over into similar short-term debt instruments as pieces of it mature.

The marginal investor must expect that their marginal dollars would be equally attractively employed in each of these strategies.

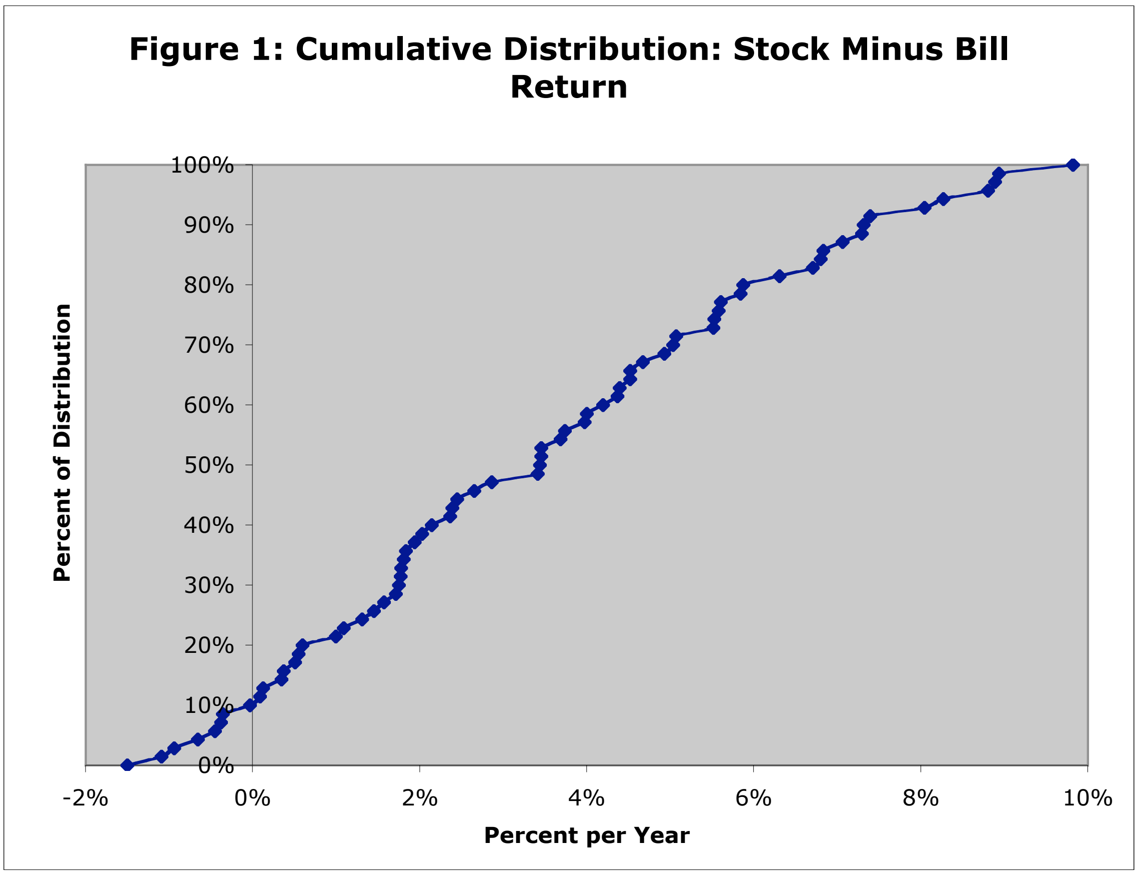
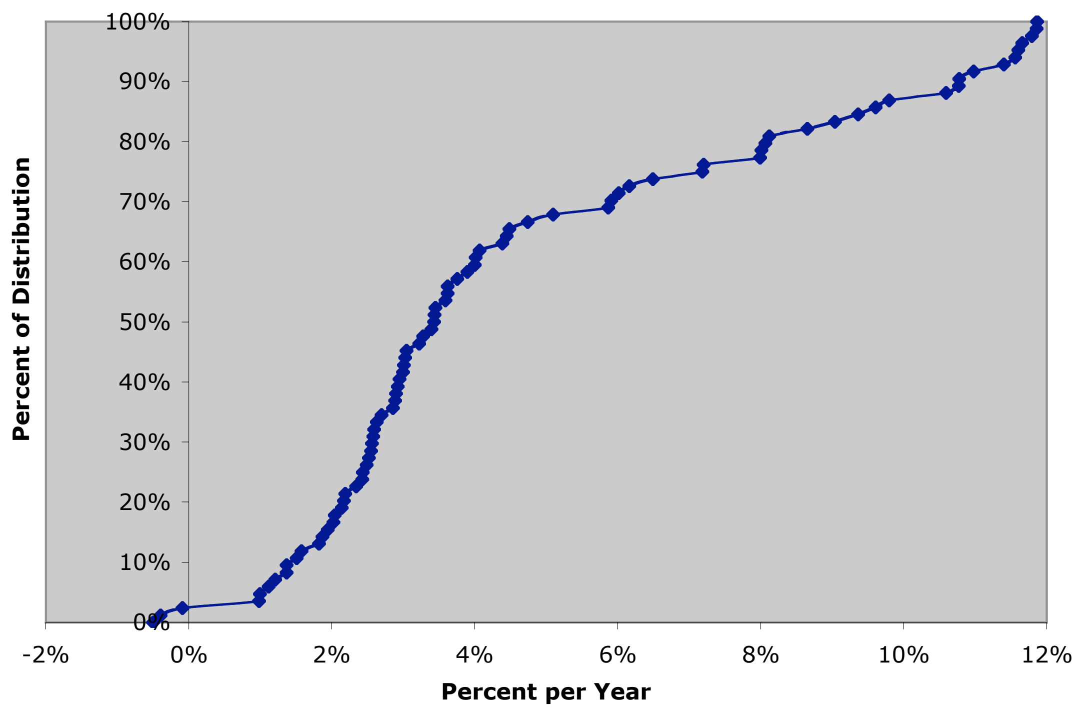


Figure 1 plots the cumulative return distribution for the relative returns for these two twenty-year portfolio strategies starting in each year since the start of the twentieth century. The average geometric return differential since 1901 is some 4.9 percent per year. When the portfolios are cashed in after twenty years, investments in diversified stock portfolios are on average 2.67 times as large as an investment in short-term Treasury bills after twenty years. Stock investors more than double their relative wealth 60 percent of the time, more than quadruple their relative wealth 30 percent of the time, and have a 17 percent chance of a more than seven-fold multiplication of relative wealth. The downside is small: the empirical CDF finds that stocks do worse than bills less than 9 percent of the time. The very worst case observed is the 20 years starting in 1965, when investing in stocks yields a relative cumulative wealth loss of 17 percent compared to investing in bills.

Figure 2: Cumulative Distribution: Stock Minus Bond Return



This equity return premium is not a liquidity effect driven by the special ease with which short-term bills can be turned into cash even in emergencies. Figure 2 shows the CDF of relative returns from the twenty-year strategies of investing in a diversified stock portfolio and investing in a long-term Treasury bond portfolio. This time lower tail is even smaller: in only 2 percent of the cases in the twentieth century would investing in bonds for 20 years outperformed investing in stocks. In the worst relative case—1929—the returns to bonds would have been only 8 percent more than stocks when the portfolios were cashed in 1949.

If the actual twentieth-century CDF is a good proxy for the true underlying *ex ante* return distribution, these return patterns have powerful implications for investors' expectations about their relative marginal utility of wealth. If the marginal investor's marginal dollar is no more advantageously employed in stocks than bonds, it must be the case that:

$$\frac{\text{(chance of loss)}\left[\textit{Average}\left[\text{(amount of loss)}\times\text{(marginal utility of wealth if loss)}\right]\right]}{\text{(chance of gain)}\left[\textit{Average}\left[\text{(amount of gain)}\times\text{(marginal utility of wealth if gain)}\right]\right]}=1$$

Over the twentieth century, the chance of relative gain is ten times the chance of loss. The average amount of gain—167%—is seventeen times the average amount of loss. If the marginal utility in gain states is perfectly correlated with the amount of gain and the marginal utility in loss states uncorrelated with the amount of loss, then the average marginal utility of wealth in "stocks lose" states must be 50 times as great as in "stocks gain" states. This is the equity return premium puzzle at its sharpest: how is one to account for this extraordinary divergence?

The equity premium puzzle appears softer if attention is focused on short-horizon investors who invest for one year only. Stocks are very risky in the short run. 1931 sees a return differential of –60%. And bonds have outperformed stocks in some 35% of the past century's years. Twenty-year investors appear to have turned their backs on nearly riskless opportunities for profit. One-year investors did not. For investors with a time horizon of one year, stocks *are* much more risky than bills[4].

Yet even on a year-to-year scale the equity premium return remains. And there are no visible[5] large year-to-year fluctuations in the consumption of investors correlated with

[4]One reason that the puzzle is softer at short horizons is that a substantial share of year-to-year variability in the stock market appears to be transitory. Stock prices look as though they are somewhat mean reverting: at the level of the stock market as a whole, past performance is not only not a guarantee of future results, past performance is negatively correlated with future results. The variance of 20-year stock returns is only 45% of what it would be if returns were serially uncorrelated (see, for example, Cochrane, 1994; Cochrane, 2006; Campbell and Shiller, 1989). Thus Samuelson (1969)'s proof that horizon is irrelevant for asset allocation fails to go through. Mean reversion can make long-term equity investments more attractive than short-term investments because investments made at one moment insure against investments made at another.

[5] Barro (2005) and others believe that there is here a small numbers problem: with a long enough sample

stock returns that would create a high marginal utility of wealth in “stocks lose” states and so account for the premium. At the one-year horizon an investor would be indifferent at the margin between stocks and bills only if he or she had a marginal utility of wealth in the gain state 83% of the way up the return distribution that was half that of marginal utility in the loss state 17% of the way up. Such a difference in marginal utilities is very difficult to square with the low variability in aggregate consumption: Rajnish Mehra and Edward Prescott (2003) report an annual standard deviation of consumption growth of only 3.6%, which they believe could support an equity return premium for a representative investor of at most two-tenths of a percentage point per year—not six.

The basic point is Richard Thaler and Matthew Rabin (2001): expected utility theory pushes us economists toward the view that agents should be nearly risk-neutral on all bets that do not involve a substantial fraction of lifetime wealth, for only substantial variations in lifetime wealth and thus in current consumption produce enough variation in marginal utility to justify substantial risk aversion. And annual stock market returns do not covary enough with current consumption and lifetime wealth.

Thus order to solve the equity premium puzzle, an economist must propose an explanation that does at least one of:

---

we would see occasional collapses in consumption and stock values that would account for what we have observed.

- providing a reason for a very large gap in the marginal utility of wealth between states of the world in which stocks do well and states of the world in which stocks do poorly.
- demonstrating that the *ex-post* return distribution seen over the twentieth century is very different from the true *ex*-ante distribution in important ways that make stocks no real bargain.
- explaining why it is that, even though stocks have been an extremely attractive investment relative to bonds and bill, money has not flowed out of bonds and bills and into stocks—pushing equity prices up and equity returns down.

A very large number of economists have done excellent work investigating and assessing different potential explanations. Among the most promising lines of work have been investigations of the implications of risk aversion, non-standard preferences; transactions costs; lower-tail risk; persistent mistakes; investor confusion; and cognitive biases.[6] A full and satisfactory explanation of the equity premium return puzzle continues to elude economists. However, none of what appear to be the live possibilities would lead one to anticipate the disappearance of the premium in the future.

III. A Preferences Explanation?

A first potential explanation is simply that rational investors prefer the portfolios they

[6]Of course, space prevents us from even noting the existence of more than a very small fraction of even the most important contributions to the literature. We can only glance at those we regard as most promising.

hold: investors truly are risk averse enough that the observed configuration of returns does not leave unexploited profit opportunities. The difficulties are twofold: first, the low average return debt securities used as a yardstick in measuring the equity return premium are not really low in risk; second, even taking debt to be risk free the degree of risk aversion needed to keep long-term investors from seeing large gains from further investments in equities must be extremely high.

As the late Fisher Black once put it in conversation, in terms of the coefficient of relative risk aversion—the standard way of measuring tolerance for risk—explaining the configuration of asset returns requires a coefficient of about 50. Consider of an agent offered a choice between (a) their current lifetime wealth and (b) a gamble where with probability p they obtain twice and with probability 1-p half their lifetime wealth. An agent with a coefficient of 2 would reject (b) if p were less than 80%; for a coefficient of 10 the critical value is 99.8%; and for a coefficient of 50 the critical value is 99.99999999995%. Many economists argue that both observed purchases of insurance and our intuitions suggest a coefficient of relative risk aversion parameter not of 50 but more in the range of 1 to 3,[7] which corresponds to Mehra-Prescott's estimate of a warranted equity premium of about 0.2 percentage points per year.

Moreover, as we economists learned from Philippe Weil (1989), a standard time-separable utility function with a high degree of risk aversion also generates both a high

[7]See, for example, Partha Dasgupta (2007).

risk-free rate of return (in economies with the roughly two percent per year consumption growth of our own economy) and smooth consumption paths that do not respond to changes in rates of return. Neither of these is observed

The most promising preference-based line of research—exemplified by papers like Lawrence Epstein and Stanley Zin (1991), George Constantinides (1990), Andrew Abel (1990), and John Campbell and John Cochrane (1995)—considers non-standard preferences, making utility dependent not just on consumption but on consumption relative to the consumption of others or to one's own past consumption and separating preferences for risk from preferences for income growth over time. These approaches account for the coexistence of a high degree of effective risk aversion and a low risk-free interest rate: the features of the utility function that make investors extremely averse to stock-market losses have no bearing on the connection between economic growth and the safe real interest rate. But these approaches still require something to generate very high effective risk aversion.

Narayana Kocherlakota (1996) summed up the results from this line of research:

> The risk-free rate puzzle can be resolved as long as the link between individual attitudes toward risk and growth contained in the standard preferences is broken…. [T]he equity premium puzzle is much more robust: individuals must either be highly averse to their own consumption

risk or to per capita consumption risk…

The modern finance literature on the equity premium puzzle is now more than two decades old. The historical investment literature looking back into observers' pasts and noting the existence of a very large equity return premium is now more than eight decades old. Yet to date no critical mass of long-term investors has taken large-enough long-enough-run positions to try to profit from the equity return premium to substantially arbitrage it away.

It is premature to say that these lines of research will never be able to satisfactorily account for the equity premium that has been observed in the past. But they do not to date appear to have done so. It is not clear how they might do so. If, however, they turn out to be correct, they do imply a future equity return premium likely to be about the six percent or so a year observed in the past.

An alternative is offered by behavioral finance economists, for example Benartzi and Thaler (1995), see investors—even professional and highly-compensated investors in it for the long run—as institutionally and psychologically incapable of framing their portfolio-choice problem in a way that allows them to appropriately discount and thus ignore the high short-term risks of equities. If investors could focus instead on the long-term returns of stocks they would realize that there is very little long-term risk in stocks relative to bonds. But they cannot. Rabin and Thaler (2001) argue that expected utility maximization cannot account for most behavior economists label “risk averse,” and should be replaced by “loss aversion” as a model of investor behavior—individuals simply feel the pain of a loss more acutely than the pleasure of an equal-sized gain. Hong and Stein point to “disagreement models” that motivate high trading volumes as a potential explanation for other asset pricing anomalies like the equity premium. Glamor stocks exhibit greater than average turnover rates, high trading volumes, tend to be overpriced and exhibit low rates of return; value stocks exhibit lower than average turnover rates, low trading volumes, tend to be underpriced, and exhibit high rates of return: perhaps this could be built into an explanation of the equity return premium.

It is not clear whether these are explanations of the puzzle or reframings of it. Humans know that they have psychological biases, and build social and economic institutions to compensate for them and to guide them into framing problems in a way that is in their long-term interest. Humans have built mechanisms like automatic payroll deductions, like inducing caution by valuing assets at the lower of cost and market, like entails and trusts. A bias-based psychological explanation must account not just for the bias but for the failure of investors to figure out ex ante how to bind themselves to the mast like Ulysses did with the Sirens.

IV. Transaction Costs and Investor Heterogeneity

Another line of research has attempted to explain the equity premium as due to transaction costs and investor heterogeneity.[8] Gregory Mankiw and Stephen Zeldes (1991) were among the very first to point out that two-thirds of Americans have next to no stock market investments—presumably because of some form of transaction cost that keeps them from being able to recognize and act on the fact that equity investments have a substantial place in every optimal portfolio. Transactions costs keeping a substantial share of the population at a zero position lock up what representative-agent models see as society's risk-bearing capacity, which then cannot be tapped and mobilized to bear equity risk.

Mankiw and Zeldes found that stockholders' consumption does not vary nearly enough to account for the equity premium. If standard representative agent models suggest that the warranted equity return premium should be on the order of 0.2 percentage points per year, a transactions-cost model in which only one-third of agents hold stocks suggests a warranted equity premium on the order of three times as large. This line of research could diminish the magnitude of the equity premium puzzle,[9] but appears to still leave an order of magnitude gap to be accounted for.

---

[8]These go together: if investors are effectively identical they do not trade and transactions costs are irrelevant; if there are no transactions costs than investor heterogeneity does not reduce the net risk-bearing capacity of the economy.

[9] See Vissing-Jorgenson (2002).

This line of research also leaves unanswered the question of just what these transaction costs are. Even back in the nineteenth century "bucket shops"—most of them honest—allowed people with very small amounts of money to "invest:" as little as one dollar could "buy" or "sell" a fractional share at the last ticker price. A bucket shop was not a brokerage. It did not invest its clients' money in the market: it paid today's withdrawals out of yesterday's deposits and relied on commissions and the law of large numbers to make it profitable.[10] And even if there were large transaction costs to buying and selling stocks, could this account for the equity premium puzzle? High costs of buying and selling are amortized over decades when investors follow multi-decade buy-and-hold strategies, and the most vivid advantages of stock investments produced by the equity return premium accrue to those who follow such strategies.

More recently, Constantinides, Donaldson, and Mehra (2002) suggest that the equity premium may be due to transaction costs in the form of borrowing constraints. The relatively young with the option of declaring bankruptcy have difficulty borrowing on a large scale. Because of such borrowing constraints, investors find it optimal to build up stocks of liquid wealth (see, for example, Mark Huggett, 1993; John Heaton and Deborah Lucas, 1995). This argument takes us economists far toward explaining why the risk-free rate of return might be low: people's unwillingness to have even temporarily negative net

[10]Nineteenth-century speculator Daniel Drew found when young that he did better at bucket shops than on Wall Street. His actual purchases and sales generated price pressure against himself, while his notional bucket shop transactions did not.

worth increases saving, increases the capital stock, and so pushes down the rate of interest and profit. But could such borrowing constraints bear much of the weight of accounting for the equity premium? Built-up stocks of wealth could be invested in either stocks or bonds, and stocks offer higher returns with little extra long-horizon risk.

The transaction costs approach that in our view comes closest to accounting for the equity premium puzzle is that of George Constantinides and Darryl Duffie (1996). They propose that investors are subject to uninsurable idiosyncratic income shocks correlated with returns on equities. Thus investors bear a large amount of equity risk embedded in their human capital, and are uninterested in further leveraging their total implicit portfolios. Advancing this explanation would require identifying groups of people whose labor income is subject to shocks correlated with equity returns and demonstrating that those investors' portfolios drive the lack of investment in equities. This has not yet been accomplished.

V. Lower-Tail Risk?

The equity premium return puzzle might be resolved by breaking the assumption that the ex post return distribution over the twentieth century is an adequate proxy for the ex ante return distribution. A high equity premium might be observed in the sample that is our past if that sample does not contain low-probability but large-magnitude economic catastrophe. A small chance of winding up truly far out in the lower tail of a return distribution can have a significant effect on ex ante and—if unobserved in sample—an even more significant effect on ex post return premia. Proposed solutions along these lines have been put forward by authors like Thomas Rietz (1988); Stephen Brown, William Goetzmann, and Stephen Ross (1995); and Robert Barro (2005). If correct, this family of solutions would imply that we economists will continue to observe a large equity premium in-sample for a while—until The Day when the long run arrives while some of us at least are still alive, the economic catastrophe occurs, and investors find their stocks nearly worthless.

This explanation must pass a camel through the eye of a needle. The unobserved-in-sample low-probability catastrophe must occur with a probability small enough that it is plausible that it has not observed. Yet the chance and magnitude of the catastrophe must be large enough to have substantial effects on prices and returns. And the catastrophe must diminish the value of stocks but not of bonds or bills—for a catastrophe that hits stocks and bonds equally has no effect on the equity premium return.[11]

This theory has considerable attractiveness. But it has one principal difficulty: it is not

[11]There is a fourth requirement, for too great a risk of a collapse in the stock market and in consumption will not only produce a high equity premium but a negative real interest rate. The size of collapse must be on a knife-edge in these models: large but not too large—large enough to create the observed equity premium, but small enough to leave a positive safe real interest rate.

obvious what the low-probability economic catastrophes with powerful negative impacts on real equity returns and little effect on bond returns are. Investors and economists can envision a great many potential political and economic catastrophes: defeat in a major war; a populist unraveling of government finances generating hyperinflation; an exhaustion of technological possibilities for innovation; or a banking-sector collapse or other financial crisis that generates a steep but transitory collapse in profits. However these catastrophes are likely to affect both stock and bond values. A permanent decline in the rate of total factor productivity and consumption growth ought to affect stock and bond returns proportionately. War defeat or populist-crisis crashes of government finance are highly likely to produce rapid inflation, which is poison to real debt returns. A transitory collapse in corporate profitability has little effect on far-sighted valuations of equities—unless it is accompanied by a collapse in consumption as well, in which case the reduced tax base is likely to lead to substantial money printing and inflation.

A large deflationary episode like the Great Depression itself could serve as a source of risk to stocks but not bonds. Few, however, believe that any future central bank would allow such a steep and persistent deflation as the Federal Reserve allowed in the 1930s. And the Great Depression is already in our sample. It is hard to argue that its absence from our sample is the cause of the observed equity return premium puzzle.

This difficulty applies also to the "survivorship" argument that looking across countries the U.S. is a large positive outlier in stock returns. It is a large positive outlier in bond returns as well.

There is one possible source that can be envisioned of a collapse in real equity values that would not much affect the real values of government bonds. If the U.S. government were to decide to put extraordinarily heavy taxes on corporate profits or to impose extraordinarily heavy regulatory burdens on corporations, those policies could redirect a substantial amount of cash flow away from shareholders without affecting bond values. Yet is the rational fear of future tax increases or regulatory burdens narrowly targeted on corporate profits large enough to support anything like the observed equity premium? But perhaps we overestimate the competence of our government, and underestimate the strength of a populism that really does believe that when the government taxes corporations no individual pays. Moreover, as public finance economists like James Hines (2005) point out, in a world of mobile capital tax competition restrains governments from pursuing tax policies very different from those of other nations. A radical failure of such tax competition would have to be required as well.

An analogous argument to Rietz (1988) and Barro (2005) is made by Martin Weitzman (2006). Weitzman argues not that lower tail risk is large, but that investors do not and cannot know what the lower tail risk truly is: Knightian uncertainty rather than von Neumann-Morgenstern risk. Once again, the principal difficulty is to identify the potential the events that investors believe might generate a long fat lower tail of equity returns and yet leave real government debt returns unaffected.

A final unresolved difficulty with the unobserved lower-tail hypothesis is that, as Barro (2005) points out, this explanation carries the implication that the greater the chance of a collapse the higher are equity prices. In this theory, 2000 is a year in which investors expected a high, and 1982 a year in which investors expected a low, probability of macroeconomic disaster.[12]

If the arguments for heretofore unobserved lower-tail risk hold true, then the appearance of an equity premium puzzle will not persist forever. At some point the risks that underpin the asset price configuration would manifest themselves, at which point it will become very clear that the equity premium puzzle never really existed at all.

VI. Learning About the Return Distribution

Yet another path assumes that economic agents are not extraordinarily risk averse, that economic agents are not limited in their risk-bearing capacity by transactions costs and heterogeneity, that the in-sample return distribution is a good proxy to the ex-ante return distribution, but that investors early in the twentieth century mistook the parameters of

[12] This is a somewhat disturbing artifact of the Lucas (1978) model that underpins papers like Rietz (1988), Barro (2005), Weitzmann (2006), and Mehra and Prescott (1985).

the fundamental return distribution, and that it has taken them a very long time indeed to learn what the true parameters of the fundamental return distribution are. Thus misperceptions created the equity premium. And the process of correcting these misperceptions has given a boost to stock prices that has further driven up the in-sample equity premium. This argument carries a corollary: the equity premium has a solid past, but it will not have as much of a future: investors have learned and will continue learn from experience over time, and if there is an equity return premium still in existence today it is likely to shrink relatively rapidly.

McGrattan and Prescott (2003) develop this argument by pointing to changing institutions as a source of the equity premium in the past that is not present today. Regulatory restrictions imposed by legislatures and courts that had too great a fear of the riskiness of equities used to encourage over-investment in debt by pension funds. Until the passage of ERISA in the mid-1970s it was unclear what a pension fund trustee could and could not do without risking legal liability. But it was clear that a trustee who invested in investment-grade bonds was in a safe harbor with respect to any possible legal liability for maladministration. And it was clear that a trustee who invested in stocks was not in a safe harbor. As time passed and as even government officials learned that the riskiness of stocks had been overstated, these regulatory restrictions fell. Thus changing expectations working through the channel of the creation of better financial institutions greatly contributed to this fall in the market risk premium on stocks.

Yet another exploration of this alternative is Olivier Blanchard (1993), who sees two major macroeconomic events driving the movements of the equity premium from 1927 until the early 1990s. He sees high equity premiums as a reaction to the shock of the Great Crash of 1929-1933, and a subsequent decline as the memory and thus the perceived likelihood of a repetition of that extraordinary event has dimmed. He also sees, as do others like Modigliani and Cohn (1979), Campbell and Vuolteenaho (20040), and Randolph Cohen, Chris Polk, and Tuomo Vuolteenaho (2005), a strong correlation of the equity premium and inflation in the 1970s and the 1980s. John Campbell and Tuomo Vuolteenaho (2004) call this effect of inflation on the equity premium a "mispricing" attributed to expectations implicit in market prices "deviating from the rational forecast." They point to Wall Street traders' use of the 'Fed model' to value stocks—believing that the *nominal* coupon yield on debt ought to be in some equilibrium relationship with the *real* earnings yield on equity—as a conceptual error that generates inflation illusion.[13]

These factors led Blanchard back in 1993 to predict that the future equity premium would "remain small," because inflation was likely to remain low and because the memory of the Great Depression was dim and would continue to erode. But Blanchard's regressions were reduced forms, and changing economic institutions and structures would lead one to fear that reduced forms might not track their future very well, and indeed this did not.

[13]It is not clear whether Campbell and Vuoleenaho view this as a misperception to be corrected by learning or as the result of psychological biases that cause confusion between real and nominal magnitudes that will persist.

Over the fourteen years from 1993 to 2007 the real return on Treasury bills has been 2.1 percent while the real return on stocks has been 7.6 percent, for an equity premium of 5.5 percent per year. Perhaps post-1993 estimates of the equity premium are high because of the stock market boom of the late 1990s, but the data since the early 1990s provides little evidence that the equity premium faded away with the vanishing of the memory of the Great Depression and the inflation of the 1970s. An 18 year-old runner from the floor of the New York Stock Exchange in 1929 would have turned 96 in 2007.

What appears as the most powerful attempt to flesh out this alternative is Fama and French (2002). Over the medium run, they argue, the risk premium on stocks has fallen as a result of the correction of misapprehensions about riskiness. Such a fall in the risk premium shows up as a jump in stock prices. Thus learning that the ex-ante equity premium should be lower than in the past produces an in-sample past equity premium even higher than its misperceived ex-ante value.

Fama and French thus argue that one should not estimate the post-World War II ex ante equity premium by looking at ex-post returns—that is, adding dividend yields to the rate of growth of stock prices. That procedure is biased because it includes this unanticipated windfall from learning about the world. One should, instead, estimate expected stock returns via the Gordon Equation:

$$r = D/P + g$$

where D/P is the dividend yield and g is the expected rate of capital gain. The dividend yield is directly observable. The expected capital gain is not, and must be estimated.

VII. The Future of the Equity Premium

What are the implications of taking Fama and French's advice, and estimating the future equity premium via this Gordon-equation approach? A natural way to estimate expected capital gains is to look at their value over the past. But estimating the expected capital gain by averaging past capital gains will be biased upward when—as Fama and French argue—the past contains learning about reduced risks that lowered required rates of return. On the other hand, estimating the expected capital gain by averaging past rates of dividend growth will be biased downward when—as has happened over the past two generations—firms have substituted stock buybacks for dividends as a wa y of pushing money out of the firm. Estimating the expected capital gain in the Gordon mode from the average of past rates of earnings growth avoids much but not all of this last bias: today's higher rate of retained earnings should fuel somewhat faster earnings growth than was generated by lower rates of retained earnings in the past.

Estimating future stock returns via the Gordon model from today's dividend yield and using the post-WWII average rate of earnings growth to forecast expected capital gains produces an expected equity premium of 4.3% per year.

But, as Fama and French further observe, we economists have had good macroeconomic news over the past century: earnings growth since 1950 has probably exceeded what would have been rational expectations formed in the shadow of the Great Depression. Thus Fama and French assess the likely equity premium going forward as likely to be less than this 4.3% per year.

The Gordon equation approach, however, faces a Modigliani-Miller problem. Optimizing firms have chosen their dividend yields for a reason. If dividend yields are currently low it might be because opportunities to invest retained earnings are especially high—in which case properly anticipated likely capital gains in the future will be higher than past historical averages. If dividend yields are currently high it might be because opportunities to invest retained earnings are especially poor—in which case properly anticipated likely capital gains in the future will be lower than past historical averages. An alternative favored by Siegel (2007) is to attempt to estimate equity returns by looking at earnings yields.

The wedge between accounting earnings yields and bond rates is not necessarily the expected equity premium. Do accounting earnings overstate or understate the true Haig-Simons earnings of the corporation, and by how much? By how much do stock options granted but not yet exercised dilute ownership, and so reduce earnings per share? What proportion of the current earnings yield is a cyclical phenomenon? To what extent do retained earnings reinvested inside of firms earn higher rates of return than outside investments subject to information and incentive problems? To what extent to retained earnings reinvested inside of firms earn lower rates of return than outside investments because of corporate control issues? Are there expectations of changes in expected rates of return which thus induce expected capital gains and losses that drive a further wedge between accounting profitability and expected real returns?

Cutting through this Gordian knot of issues, if expected rates of return are constant, accounting earnings equal Haig-Simons earnings, stock options do not much dilute ownership, earnings are not much boosted or depressed by the business cycle, and retained earnings yield the same return as outside investments, then the accounting earnings yield is the expected rate of return. As of this writing— October 16, 2007, 11.44 PDT—the annual earnings yield on the value-weighted S&P composite index is 5.530%. This is a wedge of 3.220 percent per year when compared to the annual yield on 10-year Treasury inflation-protected bonds of 2.310%.

Thus both Gordon and earnings-based approaches confirm the research-surveying judgment in Rajnish Mehra (2003) that the equity premium is likely to persist into the future, but at a level somewhat but not enormously smaller than the original estimated Mehra and Prescott (1985) 6 percent per year. As Mehra (2003) wrote—based not on his commitment to a particular model of the equity return premium but rather on agnostic uncertainty about the sources of the equity return:

> The data used to document the equity premium over the past 100 years are as good an economic data set as analysts have, and 100 years is long series when it comes to economic data. Before the equity premium is dismissed, not only do researchers need to understand the observed phenomena, but they also need a plausible explanation as to why the future is likely to be any different from the past. In the absence of this explanation, and on the basis of what is currently known, I make the following claim: Over the long term, the equity premium is likely to be similar to what it has been in the past and returns to investment in equity will continue to substantially dominate returns to investment in T-bills for investors with a long planning horizon.

Many Wall Street observers appear to agree that there remains a substantial equity premium. Ivo Welch (2000) surveyed 226 financial economists, asking them to provide their estimates of the future equity premium. Their consensus was that stocks will outperform bills by 6-7% per year for the next ten to thirty years. Gram and Harvey (2007) surveyed nonfinancial corporations' Chief Financial Officers (CFOs). Their 7,316 responses produce an expected annual equity premium of 3.2% per year. There appears to be no compelling reason why CFOs' expectations should be biased in one direction or another.

The modern finance literature on the equity premium puzzle is now more than two decades old. The historical investment literature looking back into observers' pasts and noting the existence of a very large equity return premium is now more than eight decades old. Yet to date no critical mass of long-term investors has taken large-enough long-enough-run positions to try to profit from the equity return premium to substantially arbitrage it away.

Keynes (1936) proposed an explanation. He believed that the finance practitioner professon selects for financial practitioners who are especially vulnerable to these behavioral-finance biases. He wrote that the craft of managing investments is "intolerably boring and over-exacting to any one who is entirely exempt from the gambling instinct." Thus those who would be able to ignore the short-run risks of equities do not stay in the profession. And for those who do have "the gambling instinct"? "He who has it must pay to this propensity the appropriate toll."

From Keynes's proto-behavioral-finance perspective, our collective failure to date to build institutions that will curb psychological propensities for long-run investors to overweight the short-run risks of equity investments is not a thing of the past that the finance practitioners can learn was a mistake and adjust for, but rather a sign that the equity premium return is here for a long run to come.

It would, however, be surprising if the equity premium were as large today as it has been over the past century. The memory of the Great Depression *has* faded. Institutional changes like ERISA *have* removed constraints on investing in equities. Private equity *does* lock investors' money away and so rescues it from the propensity to churn. Individual investors who control their own retirement planning through defined-contribution pension plans *do* find it easier to invest in equities, and the rise in mutual funds *has* in theory made it easier to achieve the benefits of diversification—even if a look at the spread of mutual fund returns shows that the typical mutual fund carries an astonishing amount of idiosyncratic risk.

It would be astonishing if these institutional developments had no effect on the equity return premium.

Yet if the market can be trusted, the equity premium persists today at a level difficult to account for as compensation for the long-term risks of equity investment. There are powerful expected utility-theoretic arguments that the economy has the risk-bearing capacity to make an appropriate equity return premium for visible long-run risks equal to no more than tenths of a percent per year. The existence of the equity return premium in the past offered long-horizon investors a chance to make very large returns in return for bearing little risk. It appears likely that the current configuration of market prices offers a similar opportunity to long-horizon investors today.

How damaging to the economy is this market failure to mobilize its risk-bearing capacity and drive the equity premium down by orders of magnitude? If the failure makes the cost of capital higher because capital ownership involves risk, then the throwing-away of the economy's risk-bearing capacity implies that the economy's capital-output ratio is likely to be much too low. Institutional changes that mobilized some of this absent risk-bearing capacity would then promise enormous dividends. But there is another possibility: perhaps we economists have not an equity return premium but instead a debt return discount puzzle. Firms must then overpay for equity only to the extent that investors overpay for debt. In this case the distortions created are more subtle ones of organizational form—a disfavoring of equity and a favoring of debt-heavy modes—and are presumably smaller in magnitude.

A great many agents and institutions in the economy should have a strong interest in profiting from the extremely high value of the equity return premium. There are lots of long-horizon investors who know that they will not need the money they are investing now until twenty or thirty years in the future. Think of parents of newborns looking forward to their children's college, the middle-aged looking at rapidly-escalating health-care costs, the elderly looking forward to bequeathing some of their wealth, workers with defined-contribution pensions, businesses with defined-benefit pensions, life insurance companies, governments facing an aging population, the rapidly-growing exchange reserve accounts of the world's central banks. On the other side of the market, there are companies that appear underleveraged: replacing high-priced equity capital with low-priced debt capital would seem to be as profitable a strategy for a long-lived company as investing in high-return equity rather than low-return debt is for a long-term investor.

It is understandable that some of these groups chose the aggregate debt-heavy portfolios that they must have done in order to generate the equity return premiums observed over the past century. We economists can build models about principal-agent problems in financial institutions that make portfolio managers seek trades that have high payoffs in a small fraction of a career rather than a large fraction of a lifetime. We economists can speculate about how imperatives of organizational survival lead managers to be strongly averse to putting themselves in a position where they could be bankrupted by unlikely risks that are unknown to them. And we economists can point to institutions and portfolio managers that do borrow long-term to invest in equities: many leveraged buyouts, junk bonds, private equity partnerships, Warren Buffett's career at Berkshire-Hathaway spent buying up insurance companies and putting their reserves to work buying equities. But does this add up to an explanation?

These considerations suggest a strong case for revisiting issues of financial institution design, in order to give the market a push toward being more willing to invest in equities. Economists need to think about institutions that would make long-run buy-and-hold bets on equities easier and more widespread. Mandatory personal retirement or savings accounts with default investments in equity index funds? Automatical investment of tax refunds into diversified equity funds via personal savings accounts? Investing the Social Security trust fund balance in equities as well?

References

Abel, Andrew. 1990. "Asset Prices under Habit Formation and Catching Up with the Joneses." *American Economic Review* 80:2, pp. 38-42.

Alvagari, S. Rao and Gertler, Mark. 1991. "Asset Returns with Transactions Costs and Uninsured Individual Risk." *Journal of Monetary Economics* 27:3, pp. 311-31.

Barro, Robert. 2005. "Rare Events and the Equity Premium." Cambridge: NBER Working Paper 11310.

Barsky, Robert and DeLong J. Bradford. 1993. "Why Does the Stock Market Fluctuate?" *Quarterly Journal of Economics*.

Benartzi, Shlomo and Thaler Richard H. 1995. "Mypic Loss Aversion and the Equity Premium Puzzle."

Blanchard, Olivier J. 1993. "Movements in the Equity Premium." *Brookings Papers on Economic Activity* 1993:2, pp. 75-138

Brown, Stephen, William Goetzmann, and Stephen Ross. 1995. "Survival." *Journal of Finance* 50:3, pp.853-73

Bush, Georgia. 2007. “Volatility Makes a Comeback.” New York: Citi Global Wealth Management Investment Research.

Campbell, John. 2003. “Consumption Based Asset Pricing.” In Constantinides, George, Mark Harris, and Rene Stulz, eds. *Handbook of the Economics of Finance*. Amsterdam: North Holland.

Campbell, John and Cochrane, John. 1995. “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior.” Cambridge: NBER Working Paper 4995.

Campbell, John and Shiller, Robert. 1989. “The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors.” *Journal of Financial Studies.* 1, pp. 195-228.

Campbell, John and Tuomo Vuolteenaho. 2004. “Inflation Illusion and Stock Prices.” Cambridge: NBER working paper 10263.

Cochrane, John. 1994. “Permanent and Transitory Components of GDP and Stock Prices.” *Quarterly Journal of Economics* 109, pp. 241-65.

Cochrane, John. 2006. “The Dog that Did Not Bark: A Defense of Return Predictability.”

Cambridge: NBER working paper 12026.

Cohen, Randolph, Polk Christopher, and Tuomo Vuolteenaho, 2005. "Money Illusion in the Stock Market: The Modigliani-Cohn Hypothesis." Cambridge: NBER working paper 11018.

Constantinides, George. 1990. "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy* 98:3, pp. 519-543.

Constantinides, George and Duffie, Darrel. 1996. "Asset Pricing with Heterogeneous Consumers." Bonn: University of Bonn Discussion Paper Series A 1992.

Constantinides, George, Donaldson, John, and Mehra, Rajnish. 2002. "Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle." *Quarterly Journal of Economics,* 117:1, pp. 269-96.

Dasgupta, Partha. 2007. "Discounting Climate Change." Cambridge: University of Cambridge working paper.

Dimson, Elroy, Marsh, Paul and Staunton, Mike (2006) "The Worldwide Equity Premium: A Smaller Puzzle." Working Paper, London School of Economics

Epstein, Larry and Stanley Zin. 1990. "'First-Order' Risk Aversion and the Equity Premium Puzzle." *Journal of Monetary Economics* 26:3, pp. 387-407.

Fama, Eugene F. and French, Kenneth R. (1988), Permanent and Temporary Components of Stock Prices, Journal of Political Economy, April 1988; 96(2): 246-73

Friedman, Milton and Anna Jacobson Schwartz. 1963. *A Monetary History of the United States, 1857-1960*. Princeton: Princeton University Press.

Graham, John R. and Campbell Harvey. 2007. "The Equity Risk Premium in January 2007: Evidence from the Global CFO Outlook Survey." Durham: Duke University <http://ssrn.com/abstract=959703>

Heaton, John and Deborah Lucas. 1995. "The Importance of Investor Heterogeneity and Financial Market Imperfections for the Behavior of Asset Prices." *Carnegie-Rochester Conference Series on Public Policy*. 42, pp. 1-32.

Hines, James R. 2005. "Corporate Taxation and International Competition." Ann Arbor: University of Michigan.

Huggett, Mark. 1993. "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies." *Journal of Economic Dynamics and Control.* 17:5, pp. 953-69.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan.

Kocherlakota, Narayana. 1996, “The Equity Premium: It's Still a Puzzle.” *Journal of Economic Literature* 34:1, pp. 42-71.

Lucas, Robert. 1978. “Asset Prices in an Exchange Economy.” *Econometrica* 46:6.

Macurdy, Thomas and John Shoven. 1992. “Stocks, Bonds, and Pension Wealth.” In Wise, David, ed. *Topics in the Economics of Aging.* Chicago: University of Chicago Press, pp. 61-75.

Mankiw, N. Gregory and Stephen Zeldes. 1991. “The Consumption of Stockholders and Nonstockholders.” *Journal of Financial Economics* 29:1, pp. 97-112.

McGrattan, Ellen R. and Edward C. Prescott. 2003. “Average Debt and Equity Returns: Puzzling?” Minneapolis: Federal Reserve Bank of Minneapolis, Research Department Staff Report 313.

Mehra, Rajinish. 2003. “The Equity Premium: Why Is It a Puzzle?” Cambridge: NBER Working Paper 9512.

Mehra, Rajinish and Edward C. Prescott. 1985. “The Equity Premium: a Puzzle.” *Journal of Monetary Economics* 15:2, pp. 145-162.

Mehra, Rajinish and Edward C. Prescott. 2003. “The Equity Premium in Retrospect.” Cambridge: NBER Working Paper 9525.

Modigliani, Franco and Richard A. Cohn. 1979. “Money Illusion in the Stock Market.” *Financial Analysts Journal* pp. 3-23.

Polk, Christopher, Samuel Thompson, and Tuomo Vuolteenaho, 2004. “New Forecasts of the Equity Premium.” Cambridge: NBER working paper 10406.

Rietz, Thomas. 1998. “The Equity Risk Premium: A Solution.” *Journal of Monetary Economics.*

Samuelson, Paul. 1969. “Lifetime Portfolio Selection by Dynamic Stochastic Programming,” *Review of Economics and Statistics* 51(3), pp. 239-46.

Shiller, Robert. 2006. Updated Data Set from Chapter 26 of Irrational Exuberance <http://www.econ.yale.edu/~shiller/data.htm>.

Siegel, Jeremy. 2007. “The Future for Investors” Yahoo Finance.

Siegel, Jeremy. 2002. Stocks for the Long Run, Second ed. New York: McGraw Hill.

Smith, Edgar L. 1924. Common Stocks as Long-Term Investments. New York: Macmillan.

Summers, Lawrence H. 1983. “The Nonadjustment of Nominal Interest Rates: A Study of the Fisher Effect.” In James Tobin, ed., *A Symposium in Honor of Arthur Okun*. Washington: Brookings Institution, pp. 201-241.

Vissing-Jorgenson, Annette. 2002. "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution." *Journal of Political Economy* 110:41, pp. 825-837.

Weil, Philippe. 1989. “The Equity Premium Puzzle and the Risk-Free Rate Puzzle.” *Journal of Monetary Economics* 24:3, pp.401-21.

Weitzman, Martin L. 2005. “A Unified Bayesian Theory of Equity ‘Puzzles’,”

Welch, Ivo. 2000. “Views of Financial Economists on the Equity Premium and on Professional Controversies.” *Journal of Business* 73:4, pp. 501-537.

Title: Equity Risk Premium: Expectations Great and Small

Authors: Richard A. Derrig and Elisha D. Orr

Contact: Richard A. Derrig
Senior Vice President
Automobile Insurers Bureau of Massachusetts
101 Arch Street
Boston, MA 02110
617-439-4542 x718
617-439-6789 (fax)
richard@aib.org

Version 3.0
August 28, 2003

Disclaimer: The views expressed in this paper are solely the responsibility of the authors, and do not necessarily represent the views of the AIB or of any member of AIB.

Equity Risk Premium: Expectations Great and Small

What I actually think is that our prey, called the equity risk premium, is extremely elusive.

Stephen A. Ross 2001

Abstract:
The Equity Risk Premium (ERP) is an essential building block of the market value of risk. In theory, the collective action of all investors results in an equilibrium expectation for the return on the market portfolio excess of the risk-free return, the equity risk premium. The ability of the valuation actuary to choose a sensible value for the ERP, whether as a required input to CAPM valuation, or any of its descendants, is as important as choosing risk-free rates and risk relatives (betas) to the ERP for the asset at hand. The historical realized ERP for the stock market appears to be at odds with pricing theory parameters for risk aversion. Since 1985, there has been a constant stream of research, each of which reviews theories of estimating market returns, examines historical data periods, or both. Those ERP value estimates vary widely from about minus one percent to about nine percent, based on a geometric or arithmetic averaging, short or long horizons, short or long-run expectations, unconditional or conditional distributions, domestic or international data, data periods, and real or nominal returns. This paper will examine the principal strains of the recent research on the ERP and catalogue the empirical values of the ERP implied by that research. In addition, the paper will supply several time series analyses of the standard Ibbotson Associates 1926-2002 ERP data using short Treasuries for the risk-free rate. Recommendations for ERP values to use in common actuarial valuation problems will be offered.

Acknowledgments
The authors acknowledge the helpful comments from participants in the 2003 Bowles Symposium, Louise Francis, Francis Analytics & Actuarial Data Mining Inc., and four anonymous referees. The authors would also like to thank Jack Wilson for supplying his data series from Wilson & Jones (2002).

Keywords: Equity Risk Premium, Risk Premium Puzzle, Market Return Models, CAPM, Dividend Growth Models, Actuarial Valuations.

Introduction

The Equity Risk Premium (ERP) is an essential building block of the market value of risk. In theory, the collective action of all investors results in an equilibrium expectation for the return on the market portfolio excess of the risk-free return, the equity risk premium. The ability of the valuation actuary to choose a sensible value for the ERP, whether as a required input to CAPM valuation, or any of its descendants[1], is as important as choosing risk-free rates and risk relatives (betas) to the ERP for the asset at hand. Risky discount rates, asset allocation models, and project costs of capital are common actuarial uses of ERP as a benchmark rate.

The equity risk premium should be of particular interest to actuaries. For pensions and annuities backed by bonds and stocks, the actuary needs to have an understanding of the ERP and its variability compared to fixed horizon bonds. Variable products, including Guaranteed Minimum Death Benefits, require accurate projections of returns to ensure adequate future assets. With the latest research producing a relatively low equity risk premium, the rationale for including equities in insurers' asset holdings is being tested. In describing individual investment account guarantees, LaChance and Mitchell (2003) point out an underlying assumption of pension asset investing that, based only on the historical record, future equity returns will continue to outperform bonds; they clarify that those higher expected equity returns come with the additional higher risk of equity returns. Ralfe et al. (2003) support the risky equity view and discuss their pension experience with an all bond portfolio. Recent projections in some literature of a zero or negative equity risk premium challenge the assumptions underlying these views. By reviewing some of the most recent and relevant work on the issue of the equity risk premium, actuaries will have a better understanding of how these values were estimated, critical assumptions that allowed for such a low EPR, and the time period for the projection. Actuaries can then make informed decisions for expected investment results going forward.[2]

In 1985, Mehra and Prescott published their work on the so-called Equity Risk Premium Puzzle: The fact that the historical realized ERP for the stock market 1889-1978 appeared to be at odds with and, relative to Treasury bills, far in excess of asset pricing theory values based on investors with reasonable risk aversion parameters. Since then, there has been a constant stream of research, each of which reviews theories of estimating market returns, examines historical data periods, or both.[3] Those ERP value estimates vary widely from about minus one percent to about nine percent, based on geometric or arithmetic averaging, short or long horizons, short or long-run means, unconditional or conditional expectations, using domestic or international data, differing data periods, and real or nominal returns. Brealey and Myers, in the sixth edition of their standard corporate finance textbook, believe a range of 6% to 8.5% for the US ERP is reasonable for practical project valuation. Is that a fair estimate?

---

[1] The multifactor arbitrage pricing theory (APT) of Ross (1976), the three-factor model of Fama and French (1992) and the recent Mamaysky (2002) five-factor model for stocks and bonds are all examples of enhanced CAPM models.

[2] See Appendix D

[3] For example, see Cochrane (1997), Cornell (1999), or Leibowitz (2001).

Current research on the equity risk premium is plentiful (Leibowitz, 2001). This paper covers a selection of mainstream articles and books that describe different approaches to estimating the ex ante equity risk premium. We select examples of the research that cover the most important approaches to the ERP. We begin by describing the methodology of using historical returns to predict future estimates. We identify the many varieties of ERPs in order to alert the reader to the fact that numerical estimates of the ERP that appear different may instead be about the same under a common definition. We examine the well-known Ibbotson Associates 1926-2002 data series for stationarity, i.e. time invariance of the mean ERP. We show by several statistical tests that stationarity cannot be rejected and the best estimate going forward, ceteris paribus, is the realized mean. This paper will examine the principal strains of the recent research on the ERP and catalogue the empirical values of the ERP implied by that research.[4]

We first discuss how the Social Security Administration derives estimates of the equity risk premium. Then, we survey the puzzle research, that is, the literature written in response to the Equity Premium Puzzle suggested by Mehra and Prescott (1985). We cover five major approaches from the literature. Next, we report from two surveys of ”experts” on the equity risk premium. Finally, after we describe the main strains of research, we explore some of the implications for practicing actuaries.

We do not discuss the important companion problem of estimating the risk relationship of an individual company, line of insurance, or project with the overall market. Within a CAPM or Fama-French framework, the problem is estimating a market beta.[5] Actuaries should be aware, however, that simple 60-month regression betas are biased low where size or non-synchronous trading is a substantial factor (Kaplan and Peterson (1998), Pratt (1998), p86). Adjustments are made to historical betas in order to remove the bias and derive more accurate estimates. Elton and Gruber (1995) explain that by testing the relationship of beta estimates over time, empirical studies have shown that an adjustment toward the mean should be made to project future betas.[6]

The Equity Risk Premium

Based on the definition in Brealey and Myers, *Principles of Corporate Finance* textbook, the equity risk premium (ERP) is the “expected additional return for making a risky investment rather than a safe one”. In other words, the ERP is the difference between the market return and a risk-free return. Market returns include both dividends and capital gains. Because both the historical ERP and the prospective ERP have been referred to simply as the equity risk premium, the terms *ex post* and *ex ante* are used to differentiate between them but are often omitted. Table 1 shows the historical annual

---

[4] The research catalogued appears as Appendix B.

[5] According to CAPM, investors are compensated only for non-diversifiable, or market, risk. The market beta becomes the measurement of the extent to which returns on an individual security covary with the market. The market beta times the ERP represents the non-diversifiable expected return from an individual security.

[6] Elton and Gruber (1995), p148.

average returns from 1926-2002 for large company equities (S&P 500), Treasury Bills and Bonds, and their arithmetic differences using the Ibbotson data (Ibbotson Associates, 2003).[7]

US Equity Risk Premia 1926-2002 Annual Equity Returns and Premia versus Treasury Bills, Intermediate, and Long Term Bonds			
Horizon	**Equity Returns**	**Risk-Free Return**	**ERP**
Short	12.20%	3.83%	8.37%
Intermediate	12.20%	4.81%	7.40%
Long	12.20%	5.23%	6.97%
Source: Ibbotson Yearbook (2003)			

Table 1

In 1985, Mehra and Prescott introduced the idea of the equity risk premium puzzle. The puzzling result is that the historical realized ERP for the stock market using 1889-1978 data appeared to be at odds with and, relative to Treasury bills, far in excess of asset pricing theory values based on normal parametrizations of risk aversion. When using standard frictionless return models and historical growth rates in consumption, the real risk-free rate, and the equity risk premium, the resulting relative risk aversion parameter appears too high. By choosing a maximum relative risk aversion parameter to be 10 and using the growth in consumption, Mehra and Prescott's model produces an ERP much lower than the historical.[8] Their result inspired a stream of finance literature that attempts to solve the puzzle. Two different research threads have emerged. One thread, including behavioral finance, attempts to explain the historical returns with new models and different assumptions about investors.[9] A second thread is from a group that provides estimates of the ERP that are derived from historical data and/or standard economic models. Some in this latter group argue that historical returns may have been higher than those that should be required in the future. In a curiously asymmetric way, there are no serious studies yet concluding that the historical results are too low to serve as ex ante estimates. Although both groups have made substantial and provocative contributions, the behavioral models do not give any ex ante ERP estimates other than explaining and supporting the historical returns. We presume, until results show otherwise, the behavioralists support the historical average as the ex ante unconditional long-run expectation. Therefore, we focus on the latter to catalogue equity risk premium estimates other than the historical approach, but we will discuss both as important strains for puzzle research.

Equity Risk Premium Types

Many different types of equity risk premium estimates can be given even though they are labeled by the same general term. These estimates vary widely; currently the estimates range from about nine percent to a small negative. When ERP estimates are

[7] Ibbotson's 1926-2002 series from the 2003 Yearbook, Valuation Edition. The entire series is shown in Appendix A.

[8] Campbell, Lo, and MacKinlay (1997) perform a similar analysis as Mehra and Prescott and find a risk-aversion coefficient of 19, larger than the reasonable level suggested in Mehra and Prescott's paper, pp307-308.

[9] See, for example, Benartzi and Thaler (1995) and Mehra (2002).

given, one should determine the type before comparing to other estimates. We point out seven important types to look for when given an ERP estimate. They include:

- Geometric vs. arithmetic averaging
- Short vs. long investment horizon
- Short vs. long-run expectation
- Unconditional vs. conditional on some related variable
- Domestic US vs. international market data
- Data sources and periods
- Real vs. nominal returns

The average market return and ERP can be stated as a geometric or arithmetic mean return. An arithmetic mean return is a simple average of a series of returns. The geometric mean return is the compound rate of return; it is a measure of the actual average performance of a portfolio over a given time period. Arithmetic returns are the same or higher than geometric returns, so it is not appropriate to make a direct comparison between an arithmetic estimate and a geometric estimate. However, those two returns can be transformed one to the other. For example, arithmetic returns can be approximated from geometric returns by the formula.[10]

$$AR = GR + \frac{s^2}{2}, s^2 \text{ the variance of the (arithmetic) return process}$$

Arithmetic averages of periodic returns are to be preferred when estimating next period returns since they, not geometric averages, reproduce the proper probabilities and means of expected returns.[11] ERPs can be generated by arithmetic differences (Equity – Risk Free) or by geometric differences ([(1 + Equity)/(1 + Risk Free)]-1). Usually, the arithmetic and geometric differences produce similar estimates.[12]

A second important difference in ERP estimate types is the horizon. The horizon indicates the total investment or planning period under consideration. For estimation purposes, the horizon relates to the term or maturity of the risk-free instrument that is used to determine the ERP.[13] The Ibbotson Yearbook (2003) provides definitions for three different horizons.[14] The short-horizon expected ERP[15] is defined as "the large company stock total returns minus U.S. Treasury bill *total* returns". Note, the income return and total return are the same for U.S. Treasury bills. The intermediate-horizon expected ERP is "the large company stock total returns minus intermediate-term government bond *income* returns". Finally, the long-horizon expected ERP is "the large company stock total returns minus long-term government bond *income* returns". For the Ibbotson data, Treasury bills have a maturity of approximately one month; intermediate-term government bonds have a maturity around five years; long-term government bonds

[10] See Welch (2000), Dimson et al. (2002), Ibbotson and Chen (2003).
[11] For example, see Ibbotson Yearbook, Valuation Edition (2003), pp71-73 for a complete discussion of the arithmetic/geometric choice. See also Dimson et al. (2000), p35 and Brennan and Schwartz (1985).
[12] The arithmetic difference is the geometric difference multiplied by 1 + Risk Free.
[13] See Table 1.
[14] See Ibbotson 2003 Yearbook, p177.
[15] Table 1 displays the short horizon ERP calculation for the 1926-2002 Ibbotson Data.

have a maturity of about 20 years. Although the Ibbotson definitions may not apply to other research, we will classify equity risk premium estimates based on these guidelines to establish some consistency among the current research. The reader should note that Ibbotson Associates recommends the income return (or the yield) when using a bond as the risk free rate rather than the total return.[16]

A third type is the length of time of the equity risk premium forecast. We distinguish between short-run and long-run expectations. Short-run expectations refer to the current equity risk premium, or for this paper, a prediction of up to ten years. In contrast, the long-run expectation is a forecast over ten years to as much as seventy-five years for social security purposes. Ten years appears an appropriate breaking point based on the current literature surveyed.

The next difference is whether the equity risk premium estimate is unconditional or conditioned on one or more related variables. In defining this type, we refer to an admonition by Constantinides (2002, p1568) of the differences in these estimates:

> "First, I draw a sharp distinction between *conditional, short-term forecasts* of the mean equity return and premium and *estimates of the unconditional mean*. I argue that the currently low conditional short-term forecasts of the return and premium do not lessen the burden on economic theory to explain the large unconditional mean equity return and premium, as measured by their sample average over the past one hundred and thirty years."

Many of the estimates we catalogue below will be conditional ones, conditional on dividend yield, expected earnings, capital gains, or other assumptions about the future.

ERP estimates can also exhibit a US versus international market type depending upon the data used for estimation purposes and the ERP being estimated. Dimson, et al. (2002) notes that at the start of 2000, the US equity market, while dominant, was slightly less than one-half (46.1%) of the total international market for equities, capitalized at 52.7 trillion dollars. Data from the non-US equity markets are clearly different from US markets and, hence, will produce different estimates for returns and ERP.[17] Results for the entire world equity market will, of course, be a weighted average of the US and non-US estimates.

[16] The reason for this is two-fold. First, when issued, the yield is the expected market return for the entire horizon of the bond. No net capital gains are expected for the market return for the entire horizon of the bond. No capital gains are expected at the default-free maturity. Second, historical annual capital gains on long-term Government Bonds average near zero (0.4%) over the 1926-2002 period (Ibbotson Yearbook, 2003, Table 6-7).

[17] One qualitative difference can arise from the collapse of equity markets during war time.

Worldwide Equity Risk Premia, 1900-2000 Annual Equity Risk Premium Relative to Treasury Bills		
Country	**Geometric Mean**	**Arithmetic Mean**
United States	5.8%	7.7%
World	4.9%	6.2%
Source: Dimson, et al. (2002), pages 166-167		

Table 2

The next type is the data source and period used for the market and ERP estimates. Whether given an historical average of the equity risk premium or an estimate from a model using various historical data, the ERP estimate will be influenced by the length, timing, and source of the underlying data used. The time series compilations are primarily annual or monthly returns. Occasionally, daily returns are analyzed, but not for the purpose of estimating an ERP. Some researchers use as much as 200 years of history; the Ibbotson data currently uses S&P 500 returns from 1926 to the present.[18] As an example, Siegel (2002) examines a series of real US returns beginning in 1802.[19] Siegel uses three sources to obtain the data. For the first period, 1802 to 1870, characterized by stocks of financial organizations involved in banking and insurance, he cites Schwert (1990). The second period, 1871-1925, incorporates Cowles stock indexes compiled in Shiller (1989). The last period, beginning in 1926, uses CRSP data; these are the same data underlying Ibbotson Associates calculations.

Goetzmann et al. (2001) construct a NYSE data series for 1815 to 1925 to add to the 1926-1999 Ibbotson series. They conclude that the pre-1926 and post-1926 data periods show differences in both risk and reward characteristics. They highlight the fact that inclusion of pre-1926 data will generally produce lower estimates of ERPs than relying exclusively on the Ibbotson post-1926 data, similar to that shown in Appendix A. Several studies that rely on pre-1926 data, catalogued in Appendix B, show the magnitudes of these lower estimates.[20] Table 3 displays Siegel's ERPs for three subperiods. He notes that subperiod III, 1926-2001, shows a larger ERP (4.7%), or a smaller real risk-free mean (2.2%), than the prior subperiods[21].

[18] For the Ibbotson analysis of the small stock premium, the NYSE/AMEX/NASDAQ combined data are used with the S&P 500 data falling within deciles 1 and 3 (Ibbotson 2002 Yearbook, pp122-136.)

[19] A more recent alternative is Wilson and Jones (2002) as cited by Dimson et al. (2002), p39.

[20] Using Wilson and Jones' 1871-2002 data series, time series analyses show no significant ERP difference between the 1871-1925 period and the 1926-2002 period; one cannot distinguish the old from the new. The overall average is lower with the additional 1871-1925 data, but on a statistical basis, they are not significantly different. Assuming the equivalency of the two data series for 1871 to 1925 (series of Goetzmann et al. and Wilson & Jones), the risk difference found by Goetzmann et al. must be determined by a significantly different ERP in the pre-1871 data. The 1871-1913 return is prior to personal income tax and appears to be about 35% lower than the 1926-2002 period average of 11.8%, might reflect a zero valuation for income taxes in the pre-1914 returns. Adjusting the pre-1914 data for taxes would most likely make the ERP for the entire period (1871-2002) approximately equal to 7.5%, the 1926-2002 average.

[21] The low risk-free return is indicative of the "risk-free rate puzzle", the twin of the ERP puzzle. For details see Weil (1989).

Short-Horizon Equity Risk Premium by Subperiods			
	Subperiod I	**Subperiod II**	**Subperiod III**
	1802-1870	**1871-1925**	**1926-2001**
Real Geometric Stock Returns	7.0%	6.6%	6.9%
Real Geometric Long Term Governments	4.8%	3.7%	2.2%
Equity Risk Premium	2.2%	2.9%	4.7%
Source: Siegel (2002), pages 13 and 15.			

Table 3

Smaller subperiods will show much larger variations in equity, bill and ERP returns. Table 4 displays the Ibbotson returns and short horizon risk premia for subperiods as small as 5 years. The scatter of results is indicative of the underlying large variation (20% sd) in annual data.

Average Short-Horizon Risk Premium over Various Time Period				
		Common Stocks	U. S. Treasury Bills	Short-Horizon
Year		Total Annual Returns	Total Annual Returns	Risk Premium
All Data	1926-2002	12.20%	3.83%	8.37%
50 Year	1953-2002	12.50%	5.33%	7.17%
40 Year	1963-2002	11.80%	6.11%	5.68%
30 Year	1943-1972	14.55%	2.54%	12.02%
	1973-2002	12.21%	6.61%	5.60%
15 Year	1928-1942	5.84%	0.95%	4.89%
	1943-1957	17.14%	1.20%	15.94%
	1958-1972	11.96%	3.87%	8.09%
	1973-1987	11.42%	8.20%	3.22%
	1988-2002	13.00%	5.03%	7.97%
10 Year	1933-1942	12.88%	0.15%	12.73%
	1943-1952	17.81%	0.81%	17.00%
	1953-1962	15.29%	2.19%	13.11%
	1963-1972	10.55%	4.61%	5.94%
	1973-1982	8.67%	8.50%	0.17%
	1983-1992	16.80%	6.96%	9.84%
	1993-2002	11.17%	4.38%	6.79%
5 Year	1928-1932	- 8.25%	2.55%	-10.80%
	1933-1937	19.82%	0.22%	19.60%
	1938-1942	5.94%	0.07%	5.87%
	1943-1947	15.95%	0.37%	15.57%
	1948-1952	19.68%	1.25%	18.43%
	1953-1957	15.79%	1.97%	13.82%
	1958-1962	14.79%	2.40%	12.39%
	1963-1967	13.13%	3.91%	9.22%
	1968-1972	7.97%	5.31%	2.66%
	1973-1977	2.55%	6.19%	- 3.64%
	1978-1982	14.78%	10.81%	3.97%
	1983-1987	16.93%	7.60%	9.33%
	1988-1992	16.67%	6.33%	10.34%
	1993-1997	21.03%	4.57%	16.46%
	1998-2002	1.31%	4.18%	- 2.88%

Table 4

In calculating an expected market risk premium by averaging historical data, projecting historical data using growth models, or even conducting a survey, one must determine a proxy for the "market". Common proxies for the US market include the S&P 500, the NYSE index, and the NYSE, AMEX, and NASDAQ index.[22] For the purpose of this paper, we use the S&P 500 and its antecedents as the market. However, in the various research surveyed, many different market proxies are assumed. We have already discussed using international versus domestic data when describing different MRP types. With international data, different proxies for other country, region, or world markets are used.[23] For domestic data, different proxies have been used over time as stock market exchanges have expanded.[24] Fortunately, as shown in the Ibbotson Valuation yearbook, the issue of a US market proxy does not have a large effect on the MRP estimate because the various indices are highly correlated. For example, the S&P 500 and the NYSE have a correlation of 0.95, the S&P 500 and NYSE/AMEX/NASDAQ 0.97, and the NYSE and NYSE/AMEX/NASDAQ 0.90.[25] Therefore, the market proxy selected is one reason for slight differences in the estimates of the market risk premium.

As a final note, stock returns and risk-free rates can be stated in nominal or real terms. Nominal includes inflation; real removes inflation. The equity risk premium should not be affected by inflation because either the stock return and risk-free rate both include the effects of inflation (both stated in nominal terms) or neither have inflation (both stated in real terms). If both returns are nominal, the difference in the returns is generally assumed to remove inflation. Otherwise, both terms are real, so inflation is removed prior to finding the equity risk premium. While numerical differences in the real and nominal approaches may exist, their magnitudes are expected to be small.

Equity Risk Premia 1926-2002

As an example of the importance of knowing the types of equity risk premium estimates under consideration, Table 5 displays ERP returns that each use the same historical data, but are based on arithmetic or geometric returns and the type of horizon. The ERP estimates are quite different.[26]

[22] 2003 Ibbotson Valuation Yearbook, p92.
[23] For example, Dimson (2002) and Claus and Thomas (2001) use international market data.
[24] For a data series that is a mixture of the NYSE exchange, NYSE, AMEX, and NASDAQ stock exchange, and the Wilshire 5000, see Dimson (2002), p306.
[25] 2003 Ibbotson Valuation Yearbook, p93; using data from October 1997 to September 2002.
[26] The nominal and real ERPs are identical in Table 5 because the ERPs are calculated as arithmetic differences, and the same value of inflation will reduce the market return and the risk-free return equally. Geometric differences would produce minimally different estimates for the same types.

ERP using same historical data (1926-2002)		
RFR Description	**ERP Description**	**ERP Historical Return**
Short nominal	Arithmetic Short-horizon	8.4%
Short nominal	Geometric Short-horizon	6.4%
Short real	Arithmetic Short-horizon	8.4%
Short real	Geometric Short-horizon	6.4%
Intermediate nominal	Arithmetic Inter-horizon	7.4%
Intermediate nominal	Geometric Inter-horizon	5.4%
Intermediate real	Arithmetic Inter-horizon	7.4%
Intermediate real	Geometric Inter-horizon	5.4%
Long nominal	Arithmetic Long-horizon	7.0%
Long nominal	Geometric Long-horizon	5.0%
Long real	Arithmetic Long-horizon	7.0%
Long real	Geometric Long-horizon	5.0%

Table 5

Historical Methods

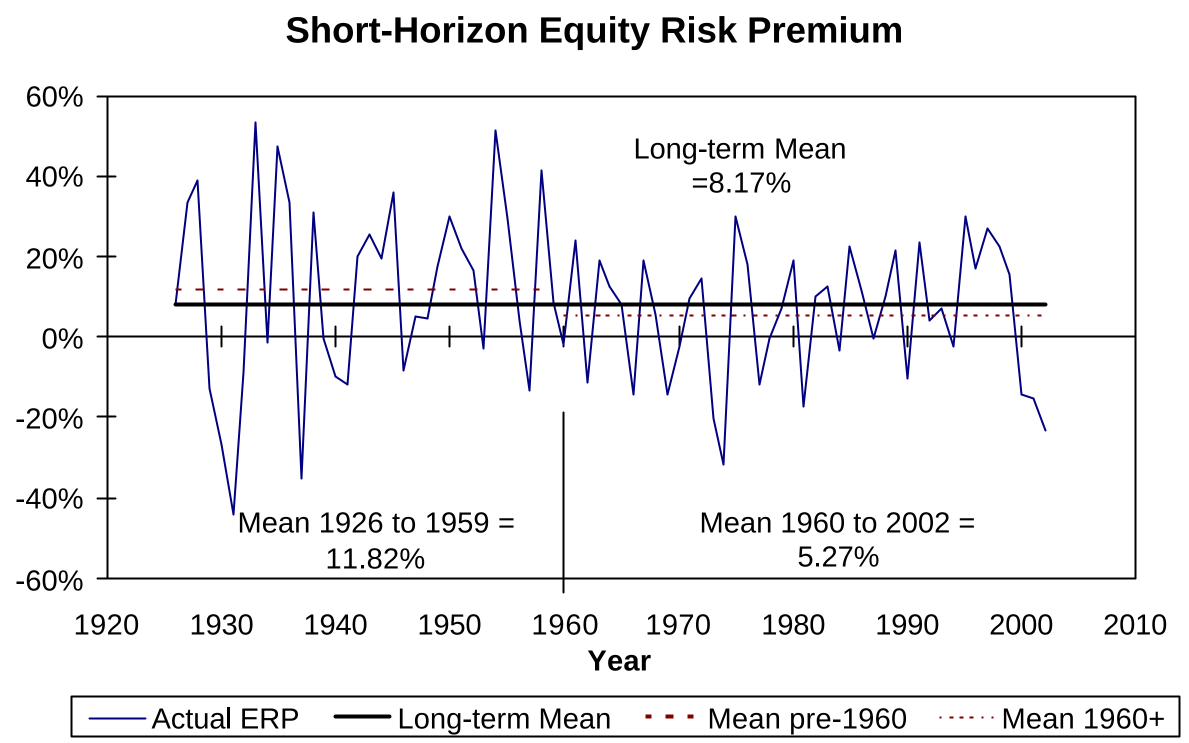
The historical methodology uses averages of past returns to forecast future returns. Different time periods may be selected, but the two most common periods arise from data provided by either Ibbotson or Siegel. The Ibbotson series begins in 1926 and is updated each year. The Siegel series begins in 1802 with the most recent compilation using returns through 2001. Appendix A provides equity risk premium estimates using Ibbotson data for the 1926-2002 period that we use in this paper for most illustrations. We begin with a look at the ERP history through a time series analysis of the Ibbotson data.

Time Series Analysis

Much of the analysis addressing the equity risk premium puzzle relies on the annual time series of market, risk-free and risk premium returns. Two opposite views can be taken of these data. One view would have the 1926-2002 Ibbotson data, or the 1802-2001 Siegel data, represent one data point; i.e., we have observed one path for the ERP through time from the many possible 77 or 200 year paths. This view rests upon the existence or assumption of a stochastic process with (possibly) inter-temporal correlations. While mathematically sophisticated, this model is particularly unhelpful without some testable hint at the details of the generating stochastic process. The practical view is that the observed returns are random samples from annual distributions that are iid, independent and identically distributed about the mean. The obvious advantage is that we have at hand 77 or 200 observations on the iid process to analyze. We adopt the latter view.

Some analyses adopt the assumption of stationarity of ERP, i.e., the true mean does not change with time. Figure 1 displays the Ibbotson ERP data and highlights two subperiods, 1926-1959 and 1960-2002.[27] While the mean ERP for the two subperiods appear quite different (11.82% vs. 5.27%), the large variance of the process (std dev 20.24%) should make them indistinguishable statistically speaking.

[27] The ERP shown here are the geometric differences (calculated) rather than the simple arithmetic differences in Table 1; i.e. ERP = $[(1+r_m)/(1+r_f)] - 1$. The test results are qualitatively the same for the arithmetic differences.



Source: Ibbotson Yearbook (2003), Geometric Differences

Figure 1

T-Tests

The standard T-test can be used for the null hypothesis H_o : mean 1960-2002 = 8.17%, the 77 year mean.[28] The outcome of the test is shown in Table 6; the null hypothesis cannot be rejected.

T-Test Under the Null Hypothesis that ERP (1960-2002) = ERP (1926-2002) = 8.17%	
Sample mean 1960-2002	5.27%
Sample s.d. 1960-2002	15.83%
T value (DF=42)	-1.20
PR > |T|	0.2374
Confidence Interval 95%	(0.0040, 0.1014)
Confidence Interval 90%	(0.0121, 0.0933)

Table 6

Another T-Test can be used to test whether the subperiod means are different in the presence of unequal variances.[29] The result is similar to Table 6 and the difference of subperiod means equal to zero cannot be rejected.[30]

[28] Standard statistical procedures in SAS 8.1 have been used for all tests.
[29] Equality of variances is rejected at the one percent level by an F test (F=2.39, DF=33,42)
[30] t-value 1.35, PR> |T| = 0.1850 with the Cochran method.

Time Trends

The supposition of stationarity of the ERP series can be supported by ANOVA regressions. The results of regressing the ERP series on time is shown in Table 7.

ERP ANOVA Regressions on Time		
Period	**Time Coefficient**	**P-Value**
1926-1959	0.004	0.355
1960-2002	0.001	0.749
1926-2002	-0.001	0.443

Table 7

There are no significant time trends in the Ibbotson ERP data.[31]

ARIMA Model

Time series analysis using the well established Box-Jenkins approach can be used to predict future series values through the lag correlation structure.[32] The SAS ARIMA procedure applied to the full 77 time series data shows:

(1) No significant autocorrelation lags.
(2) An identification of the series as white noise.
(3) ARIMA projection of year 78+ ERP is 8.17%, the 77 year average.

All of the above single time series tests point to the reasonability of the stationarity assumption for (at least) the Ibbotson ERP 77 year series.[33]

Social Security Administration

In the current debate on whether to allow private accounts that may invest in equities, the Office of the Chief Actuary of the Social Security Administration has selected certain assumptions to assess various proposals (Goss, 2001). The relevant selection is to use 7 percent as the real (geometric) annual rate of return for equities.[34] This assumption is based on the historical return of the 20th century. SSA received further support that showed the historical return for the last 200 years is consistent with this estimate, along with the Ibbotson series beginning in 1926. For SSA, the calculation of the equity risk premium uses a long-run real yield on Treasury bonds as the risk-free rate. From the assumptions in the 1995 Trustees Report, the long-run real yield on Treasury bonds that the Advisory Council proposals use is 2.3%. Using a future Treasury securities real yield of 2.3% produces a geometric equity risk premium of 4.7% over long-term Treasury securities. More recently, the Treasury securities assumption has increased to 3%[35], yielding a 4% geometric ERP over long-term Treasury securities.

[31] The result is confirmed by a separate Chow test on the two subperiods.
[32] See Harvey (1990), p30.
[33] The same tests applied to the Wilson and Jones 1871-2002 data series show similar results: Neither the 1871-1925 period nor the 1926-2002 period is different from the overall 1871-2002 period. The overall period and subperiods also show no trends over time.
[34] Compare Table 3, subperiod III.
[35] 1999 Social Security Trustees Report.

At the request of the Office of the Chief Actuary of the Social Security Administration (OCACT), John Campbell, Peter Diamond, and John Shoven were engaged to give their expert opinions on the assumptions Social Security mode. Each economist begins with the Social Security assumptions and then explains any difference he feels would be more appropriate.

In John Campbell's response, he considers valuation ratios as a comparison to the returns from the historical approach (Campbell 2001). The current valuation ratios are at unusual levels, with a low dividend-price ratio and high price-earnings ratio. He reasons that the prices are what have dramatically changed these ratios. Campbell presents two views as to the effect of valuation ratios in their current state. One view is that valuations will remain at the current level, suggesting much lower expected returns. The second view is a correction to the ratios, resulting in less favorable returns until the ratios readjust. He decides to give some weight to both possibilities, so he lowers the geometric equity return estimate to 5-5.5% from 7%. For the risk-free rate, he uses the yield on the long-term inflation-indexed bonds[36] of 3.5% or the OCACT assumption of 3%. Therefore, his geometric equity premium estimate is around 1.5 to 2.5%.

Peter Diamond uses the Gordon growth formula to calculate an estimate of the equity return (Diamond 2001). The classic Gordon Dividend Growth model is[37]:

$$K = (D_1 / P_0) + g$$

K = Expected Return or Discount Rate $\quad$ P_0 = Price this period
D_1 = Expected Dividend next period $\quad$ g = Expected growth in dividends in perpetuity

Based on his analysis, he feels that the equity return assumption of 7% for the next 75 years is not consistent with a reasonable level of stock value compared to GDP. Even when increasing the GDP growth assumption, he still does not feel that the equity return is plausible. By reasoning that the next decade of returns will be lower than normal, only then is the equity return beyond that time frame consistent with the historical return. By considering the next 75 years together, he would lower the overall projected equity return to 6-6.5%. He argues that the stock market is overvalued, and a correction is required before the long-run historical return is a reasonable projection for the future. By using the OCACT assumption of 3.0% for the long-term real yield on Treasury bonds, Diamond estimates a geometric equity risk premium of about 3-3.5%.

John Shoven begins by explaining why the traditional Gordon growth model is not appropriate, and he suggests a modernized Gordon model that allows share repurchases to be included instead of only using the dividend yield and growth rate (Shoven 2001). By assuming a long-term price-earnings ratio between its current and historical value, he comes up with an estimate for the long-term real equity return of 6.125%. Using his general estimate of 6-6.5% for the equity return and the OCACT assumptions for the long-term bond yield, he projects a long-term equity risk premium of approximately 3-3.5%. All the SSA experts begin by accepting the long-run historical

[36] See discussion of current yields on TIPS below.
[37] Brealey and Myers (2000), p67.

ERP analyses and then modifying that by changes in the risk-free rate or by decreases in the long-term ERP based on their own personal assessments. We now turn to the major strains in ERP puzzle research.

ERP Puzzle Research

Campbell and Shiller (2001) begin with the assumption of mean reversion of dividend/price and price/earnings ratios. Next, they explain the result of prior research which finds that the dividend-price ratio predicts future prices, and historically, the price corrects the ratio when it diverts from the mean.[38] Based on this result, they then use regressions of the dividend-price ratio and the price-smoothed-earnings[39] ratio to predict future stock prices out ten years. Both regressions predict large losses in stock prices for the ten year horizon. Although Campbell and Shiller do not rerun the regression on the dividend-price ratio to incorporate share repurchases, they point out that the dividend-price ratio should be upwardly adjusted, but the adjustment only moves the ratio to the lower range of the historical fluctuations (as opposed to the mean). They conclude that the valuation ratios indicate a bear market in the near future[40]. They predict for the next ten year period negative real stock returns. They caution that because valuation ratios have changed so much from their normal level, they may not completely revert to the historical mean, but this does not change their pessimism about the next decade of stock market returns.

Arnott and Ryan (2001) take the perspective of fiduciaries, such as pension fund managers, with an investment portfolio. They begin by breaking down the historical stock returns (past 74 years since December 1925) by analyzing dividend yields and real dividend growth. They point out that the historical dividend yield is much higher than the current dividend yield of about 1.2%. They argue that the changes from stock repurchases, reinvestment, and mergers and acquisitions, which affect the lower dividend yield, can be represented by a higher dividend growth rate. However, they cap real dividend or earnings growth at the level of real economic growth. They add the dividend yield and the growth in real dividends to come up with an estimate for the future equity return; the current dividend yield of 1.2% and the economic growth rate of 2.0% add to the 3.2% estimated real stock return. This method corresponds to the dividend growth model or earnings growth model and does not take into account changing valuation levels. They cite a TIPS yield of 4.1% for the real risk-free rate return.[41] These two estimates yield a negative geometric long-horizon conditional equity risk premium.

Arnott and Bernstein (2002) begin by arguing that in 1926 investors were not expecting the realized, historical compensation that they later received from stocks. They cite bonds' reaction to inflation, increasing valuations, survivorship bias[42], and changes in

[38] Campbell and Shiller (1989).
[39] Earnings are "smoothed" by using ten year averages.
[40] The stock market correction from year-end 1999 to year-end 2002 is a decrease of 37.6% or 14.6% per year. Presumably, the "next ten years" refers to 2000 to 2010.
[41] See the current TIPS yield discussion near end of paper.
[42] See Brown et al. (1992, 1995) for details on potential survivorship bias.

regulation as positive events that helped investors during this period. They only use the dividend growth model to predict a future expected return for investors. They do not agree that the earnings growth model is better than the dividend growth model both because earnings are reported using accounting methods and earnings data before 1870 are inaccurate. Even if the earnings growth model is chosen instead, they find that the earnings growth rate from 1870 only grows 0.3% faster than dividends, so their results would not change much. Because of the Modigliani-Miller theorem[43], a change in dividend policy should not change the value of the firm. They conclude that managers benefited in the "era of 'robber baron' capitalism" instead of the conclusion reached by others that the dividend growth model under-represents the value of the firm.

By holding valuations constant and using the dividend yield and real growth of dividends, Arnott and Bernstein calculate the equity return that an investor might have expected during the historical time period starting in 1802. They use an expected dividend yield of 5.0%, close to the historical average of 1810 to 2001. For the real growth of dividends, they choose the real per capita GDP growth less a reduction for entrepreneurial activity in the economy plus stock repurchases. They conclude that the net adjustment is negative, so the real GDP growth is reduced from 2.5-3% to only 1%. A fair expectation of the stock return for the historical period is close to 6.1% by adding 5.0% for the dividend yield and a net real GDP per capita growth of 1.1%. They use a TIPS yield of 3.7% for the real risk-free rate, which yields a geometric intermediate-horizon equity risk premium of 2.4% as a fair expectation for investors in the past. They consider this a "normal" equity risk premium estimate. They also opine that the current ERP is zero; i.e. they expect stocks and (risk-free) bonds to return the same amounts.

Fama and French (2002) use both the dividend growth model and the earnings growth model to investigate three periods of historical returns: 1872 to 2000, 1872 to 1950, and 1951 to 2000. Their ultimate aim is to find an unconditional equity risk premium. They cite that by assuming the dividend-price ratio and the earnings-price ratio follow a mean reversion process, the result follows that the dividend growth model or earnings growth model produce approximations of the unconditional equity return. Fama and French's analysis of the earlier period of 1872 to 1950 shows that the historical average equity return and the estimate from the dividend growth model are about the same. In contrast, they find that the 1951 to 2000 period has different estimates for returns when comparing the historical average and the growth models' estimates. The difference in the historical average and the model estimates for 1951 to 2000 is interpreted to be "unexpected capital gains" over this period. They find that the unadjusted growth model estimates of the ERP, 2.55% from the dividend model and 4.32% from the earnings model, fall short of the realized average excess return for 1951-2000. Fama and French prefer estimates from growth models instead of the historical method because of the lower standard error using the dividend growth model. Fama and French provide 3.83% as the unconditional expected equity risk premium return (referred to as the annual bias-adjusted ERP estimate) using the dividend growth model with underlying data from 1951 to 2000. They give 4.78% as the unconditional expected equity risk

[43] Brealey and Myers (2000), p447. See also discussion in Ibbotson and Chen (2003).

premium return using the earnings growth model with data from 1951 to 2000. Note that using a one-month Treasury bill instead of commercial paper for the risk-free rate would increase the ERP by about 1% to nearly 6% for the 1951-2000 period.

Ibbotson and Chen (2003) examine the historical real geometric long-run market and long risk-free returns using their "building block" methodology.[44] They use the full 1926-2000 Ibbotson Associates data and consider as building blocks all of the fundamental variables of the prior researchers. Those blocks include (not all simultaneously):

- Inflation
- Real risk-free rates (long)
- Real capital gains
- Growth of real earnings per share
- Growth of real dividends
- Growth in payout ratio (dividend/earnings)
- Growth in book value
- Growth in ROE
- Growth in price/earnings ratio
- Growth in real GDP/population
- Growth in equities excess of GDP/POP
- Reinvestment

Their calculations show that a forecast real geometric long run return of 9.4% is a reasonable extrapolation of the historical data underlying a realized 1926-2000 return of 10.7%, yielding a long horizon arithmetic ERP of 6%, or a short horizon arithmetic ERP of about 7.5%.

The authors construct six building block methods; i.e., they use combinations of historic estimates to produce an expected geometric equity return. They highlight the importance of using both dividends and capital gains by invoking the Modigliani-Miller theorem. The methods, and their component building blocks are:

- Method 1: Inflation, real risk free rate, realized ERP
- Method 2: Inflation, income, capital gains and reinvestment
- Method 3: Inflation, income, growth in price/earnings, growth in real earnings per share and reinvestment.
- Method 4: Inflation, growth rate of price/earnings, growth rate of real dividends, growth rate of payout ratio dividend yield and reinvestment
- Method 5: Inflation, income growth rate of price/earnings, growth of real book value, ROE growth and reinvestment
- Method 6: Inflation, income, growth in real GDP/POP, growth in equities excess GDP/POP and reinvestment.

[44] See Appendix D for a summary of their building block estimates. See also Pratt (1998) for a discussion of the Building Block, or Build-Up Model, cost of capital estimation method.

All six methods reproduce the historical long horizon geometric mean of 10.70% as shown in Appendix D. Since the source of most other researchers' lower ERP is the dividend yield, the authors recast the historical results in terms of ex ante forecasts for the next 75 years. Their estimate of 9.37% using supply side methods 3 and 4 is approximately 130 basis points lower than the historical result. Within their methods, they also show how the substantially lower expectation of 5.44% for the long mean geometric return is calculated by omitting one or more relevant variables. Underlying these ex ante methods are the assumptions of stationarity of the mean ERP return and market efficiency, the absence of the assumption that the market has mispriced equities. All of their methods are aimed at producing an unconditioned estimate of the ex ante ERP.

As opposed to short-run, conditional estimates from Campbell and Shiller and others, Constantinides (2002) seeks to estimate the unconditional equity risk premium, more in line with the goal of Fama and French (2002) and Ibbotson and Chen (2003). He begins with the premise that the unconditional ERP can be estimated from the historical average using the assumption that the ERP follows a stationary path. He suggests most of the other research produces conditional estimates, conditioned upon beliefs about the future paths of fundamentals such as dividend growth, price-earnings ratio and the like. While interesting in themselves, they add little to the estimation of the unconditional mean ERP.

Constantinides uses the historical return and adjusts downward by the growth in the price-earnings ratio to calculate the unconditional equity risk premium. He removes the growth in the price-earnings ratio because he is assuming no change in valuations in the unconditional state. He gives estimates using three periods. For 1872-2000, he uses the historical equity risk premium which is 6.9%, and after amortizing the growth in the price-dividend ratio or price-earnings ratio over a period as long as 129 years, the effect of the potential reduction is no change. Therefore, he finds an unconditional arithmetic, short-horizon equity risk premium of 6.9% using the 1872-2000 underlying data. For 1951-2000, he again starts with the historical equity risk premium which is 8.7% and lowers this estimate by the growth in the price-earnings ratio of 2.7% to find an unconditional arithmetic, short-horizon equity risk premium of 6.0%. For 1926-2000, he uses the historical equity risk premium which is 9.3% and reduces this estimate by the growth in the price-earnings ratio of 1.3% to find an unconditional arithmetic, short-horizon equity risk premium of 8.0%. He appeals to behavioral finance to offer explanations for such high unconditional equity risk premium estimates.

From the perspective of giving practical investor advice, Malkiel (1999) discusses "the age of the millennium" to give some indication of what investors might expect for the future. He specifically estimates a reasonable expectation for the first few decades of the twenty-first century. He estimates the future bond returns by giving estimates if bonds are held to maturity with corporate bonds of 6.5-7%, long-term zero-coupon Treasury bonds of about 5.25%, and TIPS with a 3.75% return. Depending on the desired level of risk, Malkiel indicates bondholders should be more favorably compensated in the future compared to the historical returns from 1926 to 1998. Malkiel uses the earnings growth model to predict future equity returns. He uses the current dividend yield of 1.5% and an earnings growth estimate of 6.5%, yielding an 8% equity return estimate compared with an 11% historical return. Malkiel's estimated range of the equity risk premium is from 1% to 4.25%, depending on the risk-free instrument selected. Although his equity risk premium is lower than the historical return, his selection of a relatively high earnings growth rate is similar to Ibbotson and Chen's forecasted models. In contrast with Ibbotson and Chen, Malkiel allows for a changing equity risk premium and advises investors to not rely solely on the past "age of exuberance" as a guide for the future. Malkiel points out the impact of changes in valuation ratios, but he does not attempt to predict future valuation levels.

Finally, Mehra (2002) summarizes the results of the research since the ERP puzzle was posed. The essence of the puzzle is the inconsistency of the ERPs produced by descriptive and prescriptive economic models of asset pricing on the one hand and the historical ERPs realized in the US market on the other. Mehra and Prescott (1985) speculated that the inconsistency could arise from the inadequacy of standard models to incorporate market imperfections and transaction costs. Failure of the models to reflect reality rather than failure of the market to follow the theory seems to be Mehra's conclusion as of 2002. Mehra points to two promising threads of model-modifying research. Campbell and Cochrane (1999) incorporate economic cycles and changing risk aversion while Constantinides et al. (2002) propose a life cycle investing modification, replacing the representative agent by segmenting investors into young, middle aged, and older cohorts. Mehra sums up by offering:

> "Before we dismiss the premium, we not only need to have an understanding of the observed phenomena but also why the future is likely to be different. In the absence of this, we can make the following claim based on what we know. Over the long horizon the equity premium is likely to be similar to what it has been in the past and the returns to investment in equity will continue to substantially dominate those in bonds for investors with a long planning horizon."

<u>Financial Analyst Estimates</u>

Claus and Thomas (2001) and Harris and Marston (2001) both provide equity premium estimates using financial analysts' forecasts. However, their results are rather different. Claus and Thomas use an abnormal earnings model with data from 1985 to 1998 to calculate an equity risk premium as opposed to using the more common dividend growth model. Financial analysts project five year estimates of future earnings growth rates. When using this five year growth rate for the dividend growth rate in perpetuity in the Gordon growth model, Claus and Thomas explain that there is a potential upward bias in estimates for the equity risk premium. Therefore, they choose to use the abnormal earnings model instead and only let earnings grow at the level of inflation after five years. The abnormal earnings model replaces dividends with "abnormal earnings" and discounts each flow separately instead of using a perpetuity. The average estimate that they find is 3.39% for the equity risk premium. Although it is generally recognized that financial analysts' estimates have an upward bias, Claus and Thomas propose that in the current literature, financial analysts' forecasts have underestimated short-term earnings in order for management to achieve earnings estimates in the slower economy. Claus and Thomas conclude that their findings of the ERP using data from the past fifteen years are not in line with historical values.

Harris and Marston use the dividend growth model with data from 1982 to 1998. They assume that the dividend growth rate should correspond to investor expectations. By using financial analysts' longest estimates (five years) of earnings growth in the model, they attempt to estimate these expectations. They argue that if investors are in accord with the optimism shown in analysts' estimates, even biased estimates do not pose a drawback because these market sentiments will be reflected in actual returns. Harris and Marston find an equity risk premium estimate of 7.14%. They find fluctuations in the equity risk premium over time. Because their estimates are close to historical returns, they contend that investors continue to require a high equity risk premium.

Survey Methods

One method to estimate the ex ante equity risk premium is to find the consensus view of experts. John Graham and Campbell Harvey perform a survey of Chief Financial Officers to determine the average cost of capital used by firms. Ivo Welch surveys financial economists to determine the equity risk premium that academic experts in this area would estimate.

Graham and Harvey administer surveys from the second quarter of 2000 to the third quarter of 2002 (Graham and Harvey, 2002). For their survey format, they show the current ten year bond yield and then ask CFOs to provide their estimate of the S&P 500 return for the next year and over the next ten years. CFOs are actively involved in setting a company's individual hurdle[45] rate and are therefore considered knowledgeable about investors' expectations.[46] When comparing the survey responses of the one and ten year returns, the one year returns have so much volatility that they conclude that the ten-year equity risk premium is the more important and appropriate return of the two when making financial decisions such as hurdle rates and estimating cost of capital. The average ten-year equity risk premium estimate varies from 3% to 4.7%.

The most current Welch survey compiles the consensus view of about five hundred financial economists (Welch 2001). The average arithmetic estimate for the 30-year equity risk premium relative to Treasury bills is 5.5%; the one-year arithmetic equity risk premium consensus is 3.4%. Welch deduces from the average 30-year geometric

[45] A "hurdle" rate is a benchmark cost of capital used to evaluate projects to accept (expected returns greater than hurdle rate) or reject (expected returns less than hurdle rate).

[46] Graham and Harvey claim three-fourths of the CFOs use CAPM to estimate hurdle rates.

equity return estimate of 9.1% that the arithmetic equity return forecast is approximately 10%.[47]

Welch's survey question allows the participants to self select into different categories based upon their knowledge of ERP. The results indicate that the responses of the less ERP knowledgeable participants showed more pessimism than those of the self reported experts. The experts gave 30-year estimates that are 30 to 150 basis points above the estimates of the non-expert group.

Differences in Forecasts across Expertise Level				
Relative Expertise	**Statistic**	**Stock Market**	**Equity Premium**	
		30-Year Geometric	30-Year Arithmetic	30-Year Geometric
188 Less Involved	Mean	8.5%	4.9%	4.4%
	Median	8%	5%	4%
	IQ Range	6%-10%	3%-6%	2%-5.5%
235 Average	Mean	9.2%	5.8%	4.8%
	Median	9%	5%	4%
	IQ Range	7.5%-10%	3.5%-7%	3%-6%
72 Experts	Mean	10.1%	6.2%	5.4%
	Median	9%	5.4%	5%
	IQ Range	8%-11%	4%-7.5%	3.4%-6%
Data Source: Welch (2001), Table 5				

Table 8

Table 8 shows that there may be a "lemming" effect, especially among economists who are not directly involved in the ERP question. Stated differently, all the academic and popular press, together with the prior Welch survey[48] could condition the non-expert, the "less involved", that the expected ERP was lower than historic levels.

The Behavioral Approach

Benartzi and Thaler (1995) analyze the equity risk premium puzzle from the point of view of prospect theory (Kahneman and Tversky; 1979). Prospect theory[49] has "loss aversion", the fact that individuals are more sensitive to potential loss than gain, as one of its central tenets. Once an asymmetry in risk aversion is introduced into the model of the rational representative investor or agent, the unusual risk aversion problem raised initially by Mehra and Prescott (1985) can be "explained" within this behavioral model of decision-making under uncertainty. Stated differently, given the historical ERP series, there exists a model of investor behavior that can produce those or similar results. Benartzi and Thaler combine loss aversion with "mental accounting", the behavioral process people use to evaluate their status relative to gains and losses compared to expectations, utility and wealth, to get "myopic loss aversion". In particular, mental

[47] For the Ibbotson 1926-2002 data, the arithmetic return is about 190 basis points higher than the geometric return rather than the inferred 90 basis points. This suggests the participant's beliefs may not be internally consistent.
[48] The prior Welch survey in 1998 had a consensus ERP of about 7%.
[49] A current survey of the applications of prospect theory to finance can be found in Benartzi et al. (2001).

accounting for a portfolio needs to take place infrequently because of loss aversion, in order to reduce the chances of observing loss versus gain. The authors concede that there is a puzzle with the standard expected utility-maximizing paradigm but that the myopic loss aversion view may resolve the puzzle. The authors' views are not free of controversy; any progress along those lines is sure to match the advance of behavioral economics in the large.

The adoption of other behavioral aspects of investing may also provide support for the historical patterns of ERPs we see from 1802-2002. For example, as the true nature of risk and rewards has been uncovered by the virtual army of 20th century researchers, and as institutional investors held sway in the latter fifty years of the century, the demand for higher rewards seen in the later historical data may be a natural and rational response to the new and expanded information set. Dimson et al. (2002, Figure 4-6) displays increasing real US equity returns of 6.7, 7.4, 8.2 and 10.2 for periods of 101, 75, 50 and 25 years ending in 2001 consistent with this "risk-learning" view.

Next Ten Years

The "next ten years" is an issue that experts reviewing Social Security assumptions and Campbell and Shiller address either explicitly or implicitly. Experts evaluating Social Security's proposals predicted that the "next ten years", indicating a period beginning around 2000, of returns were likely to be below the historical return. However, a historical return was recommended as appropriate for the remaining 65 of the 75 years to be projected. For Campbell and Shiller (2001), the period they discuss is approximately 2000-2010. Based on the current state of valuation ratios, they predict lower stock market returns over "the next ten years". These expert predictions, and other pessimistic low estimates, have already come to fruition as market results 2000 through 2002.[50] The US equities market has decreased 37.6% since 1999, or an annual decrease of 14.6%. Although these forecasts have proved to be accurate in the short term, for future long-run projections, the market is not at the same valuation today as it was when these conditional estimates were originally given. Therefore, actuaries should be wary of using the low long-run estimates made prior to the large market correction of 2000-2002.

Treasury Inflation Protection Securities (TIPS)

Several of the ERP researchers refer to TIPS when considering the real risk free rates. Historically, they adjust Treasury yields downward to a real rate by an estimate of inflation, presumably for the term of the Treasury security. As Table 3 shows using the Siegel data, the modern era data show a low real long-term risk-free rate of return (2.2%). This contrasts with the initial[51] TIPS issue yields of 3.375%. Some researchers use those TIPS yields as (market) forecasts of real risk-free returns for intermediate and long-horizon, together with reduced (real) equity returns to produce low estimates of ex ante ERPs. None consider the volatility of TIPS as indicative of the accuracy of their ERP estimate.

---

[50] The Social Security Advisory Board will revisit the seventy five year rate of return assumption during 2003, Social Security Advisory Board (2002).

[51] TIPS were introduced by the Treasury in 1996 with the first issue in January, 1997.

Table 9 shows a recent market valuation of ten and thirty year TIPS issued in 1998-2002.

Inflation-Indexed Treasury Securities		
Maturity	**Coupon Issue Rate**	**Yield to Maturity**
1/11	3.500	1.763
1/12	3.375	1.831
7/12	3.000	1.878
4/28	3.625	2.498
4/29	3.875	2.490
4/32	3.375	2.408
Source: WSJ 1 2/24/2003		

Table 9

Note the large 90-180 basis point decrease in the current "real" yields from the issue yields as recent as ten months ago. While there can be several explanations for the change (revaluation of the inflation option, flight to Treasury quality, paucity of 30 year Treasuries), the use of these current "real" risk free yields, with fixed expected returns, would raise ERPs by at least one percent.

Conclusion

This paper has sought to bring the essence of recent research on the equity risk premium to practicing actuaries. The researchers covered here face the same ubiquitous problems that actuaries face daily: Do I rely on past data to forecast the future (costs, premiums, investments) or do I analyze the past and apply informed judgment as to future differences, if any, to arrive at actuarially fair forecasts? Most of the ERP estimates lower than the unconditional historical estimate have an undue reliance on recent lower dividend yields (without a recognition of capital gains[52]) and/or on data prior to 1926.

Despite a spate of research suggesting ex ante ERPs lower than recent realized ERPs, actuaries should be aware of the range of estimates covered here (Appendix B); be aware of the underlying assumptions, data and terminology; and be aware that their independent analysis is required before adopting an estimate other than the historical average. We believe that the Ibbotson-Chen (2003) layout, reproduced here as Appendix D, offers the actuary both an understanding of the fundamental components of the historical ERP and the opportunity to change the estimates based upon good judgment and supportable beliefs. We believe that reliance solely on "expert" survey averages, whether of financial analysts, academic economists, or CFOs, is fraught with risks of statistical bias to fair estimates of the forward ERP.

---

[52] Under the current US tax code, capital gains are tax-advantaged relative to dividend income for the vast majority of equity holders (households and mutual funds are 55% of the total equity holders, Federal Flow of Funds, 2002 Q3, Table L-213). Curiously, the reverse is true for property-liability insurers because of the 70% stock dividend exclusion afforded insurers.

It is dangerous for actuaries to engage in simplistic analyses of historical ERPs to generate ex ante forecasts that differ from the realized mean.[53] The research we have catalogued in Appendix B, the common level ERPs estimated in Appendix C, and the building block (historical) approach of Ibboston and Chen in Appendix D all discuss important concepts related to both ex post and ex ante ERPs and cannot be ignored in reaching an informed estimate. For example, Richard Wendt, writing in a 2002 issue of Risks and Rewards, a newsletter of the Society of Actuaries, concludes that a linear relationship is a better predictor of future returns than a "constant" ERP based on the average historical return. He arrives at this conclusion by estimating a regression equation[54] relating long bond yields with 15-year geometric mean market returns starting monthly in 1960. First, there is no significant relationship between short, intermediate or long-term income returns over 1926-2002 (or 1960-2002) and ERPs, as evidenced by simple regressions using Ibbotson data.[55] Second, if the linear structural equation indeed held, there would be no need for an ERP since the (15-year) return could be predicted within small error bars. Third, there is always a negative bias introduced when geometric averages are used as dependent variables (Brennan and Schwartz, 1985). Finally, the results are likely to be spurious due to the high autocorrelations of the target and independent variables; an autocorrelation correction would eliminate any significant relationship of long-yields to the ERP.

Actuaries should also be aware of the variability of both the ERP and risk-free rate estimates discussed in this paper (see Tables 4 and 9). All too often, return estimates are made without noting the error bars and that can lead to unexpected "surprises". As one example, recent research by Francis Longstaff (2002), proposes that a 1991-2001 "flight to quality" has created a valuation premium (and lowered yields) in the entire yield curve of Treasuries. He finds a 10 to 16 basis point liquidity premium throughout the zero coupon Treasury yield curve. He translates that into a 10% to 15% pricing difference at the long end. This would imply a simple CAPM market estimate for the long horizon might be biased low.

Finally, actuaries should know that the research catalogued in Appendix B is not definitive. No simple model of ERP estimation has been universally accepted. Undoubtedly, there will be still more empirical and theoretical research into this data rich financial topic. We await the potential advances in understanding the return process that the behavioral view may uncover.

Post Script: Appendices A-D

We provide four appendices that catalogue the ERP approaches and estimates discussed in the paper. Actuaries, in particular, should find the numerical values, and descriptions of assumptions underlying those values, helpful for valuation work that

[53] ERPs are derived from historical or expected after corporate tax returns. Pre-tax returns depend uniquely on the tax schedule for the differing sources of income.
[54] 15-year mean returns = 2.032 (Long Government Bond Yield) – 0.0242, R^2 = 0.882.
[55] The p-values on the yield-variables in an ERP/Yield regression using 1926-2002 annual data are 0.1324, 0.2246, and 0.3604 for short, intermediate and long term yields respectively with adjusted r square virtually zero.

adjusts for risk. Appendix A provides the annual Ibbotson data from 1926 through 2002 from Ibbotson Associates referred to throughout this paper. The equity risk-premium shown is a simple difference of the arithmetic stock returns and the arithmetic U.S. Treasury Bills total returns. Appendix B is a compilation of articles and books related to the equity risk premium. The puzzle research section contains the articles and books that were most related to addressing the equity risk premium puzzle. Page 1 of Appendix B gives each source, along with risk-free rate and equity risk premium estimates. Then, each source's estimate is classified by type (indicated with an X for the appropriate type). Page 2 of Appendix B shows further details collected from each source. This page adds the data period used, if applicable, and the projection period. We also list the general methodology used in the reference. The final three pages of Appendix B provide the footnotes which give additional details on the sources' intent.

Appendix C adjusts all the equity risk premium estimates to a short-horizon, arithmetic, unconditional ERP estimate. We begin with the authors' estimates for a stock return (the risk-free rate plus the ERP estimate). Next, we make adjustments if the ERP "type" given by the author(s) is not given in this format. For example, to adjust from a geometric to an arithmetic ERP estimate, we adjust upwards by the 1926-2002 historical difference in the arithmetic large company stocks' total return and the geometric large company stocks' total return of 2%. Next, if the estimate is given in real instead of nominal terms, we adjust the stock return estimate upwards by 3.1%, the 1926-2002 historical return for inflation.

We make an approximate adjustment to move the estimate from a conditional to unconditional estimate based on Fama and French (2002). Using the results for the 1951-2000 period shown in Table 4 of their paper and the standard deviations provided in Table 1, we have four adjustments based on their data. For the 1951-2000 period, Fama and French use an adjustment of 1.28% for the dividend growth model and 0.46% for the earnings growth model. Following a similar calculation, the 1872-2000 period would require a 0.82% adjustment using a dividend growth model; the 1872-1950 period would require a 0.54% adjustment using a dividend growth model. Earnings growth models were used by Fama and French only for the 1951-2000 data period. Therefore, we selected the lowest adjustment (0.46%) as a minimum adjustment from a conditional estimate to an unconditional estimate. Finally, we subtract the 1926-2002 historical U.S. Treasury Bills' total return to arrive at an adjusted equity risk premium.

These adjustments are only approximations because the various sources rely on different underlying data, but the changes in the ERP estimate should reflect the underlying concept that different "types" of ERPs cannot be directly compared and require some attempt to normalize the various estimates.

Page 1 of Appendix D is a table from Ibbotson and Chen which breaks down historical returns using various methods that correspond to their 2003 paper (reprinted with permission of Ibbotson Associates). The bottom portion provides forward-looking estimates. Page 2 of Appendix D is provided to show the formulas that Ibbotson and Chen develop within their paper.

References

Arnott, Robert D., and Peter L. Bernstein, 2002, What Risk Premium Is "Normal"?. *Financial Analysts Journal,* 58, 64-85.

Arnott, Robert D., and Ronald J. Ryan, 2001, The Death of the Risk Premium: Consequences of the 1990s, *Journal of Portfolio Management,* 27, 61-74.

Barberis, Nicholas, Ming Huang, and Tano Santos, 2001, Prospect Theory and Asset Prices, *Quarterly Journal of Economics,* 116, 1-53.

Benartzi, Shlomo, and Richard H. Thaler, 1995, Myopic Loss Aversion and the Equity Premium Puzzle, *Quarterly Journal of Economics,* 110, 73-92.

Brealey, Richard A., and Stewart C. Myers, 2000, *Principles of Corporate Finance*, 6th ed. (Boston: Irwin McGraw-Hill).

Brennan, Michael J., and Eduardo S. Schwartz, 1985, On the Geometric Mean Index: A note, *Journal of Financial and Quantitative Analysis*, 20,119-122.

Brown, Stephen J., William Goetzmann, Roger G. Ibboston and Stephen A. Ross 1992, Survivorship Bias in Performance Studies, *Review of Financial Studies,* 5, 553-580.

Brown, Stephen J., William N. Goetzmann, and Stephen A. Ross, 1995, Survival, *Journal of Finance,* 50, 853-873.

Campbell, John Y., 2001, Forecasting U.S. Equity Returns in the 21st Century, *Estimating the Real Rate of Return on Stocks Over the Long Term*, manuscript presented to the Social Security Advisory Board. http://www.ssab.gov/estimated%20rate%20of%20return.pdf

Campbell, John Y., and John H. Cochrane, 1999, By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior, *Journal of Political Economics* 107: 205-251.

Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton: Princeton University Press).

Campbell, John Y., and Robert J. Shiller, 1989, The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, *Review of Financial Studies* 1, 195-228.

Campbell, John Y., and Robert J. Shiller, 2001, Valuation Ratios and the Long-Run Stock Market Outlook: An Update, Working paper #8221, NBER, forthcoming in

Advances in Behavioral Finance, Volume II, Nicholas Barberis and Richard Thaler eds., Russell Sage Foundation, 2003.

Claus, James, and Jacob Thomas, 2001, Equity Premia as Low as Three Percent?, *Journal of Finance,* 56, 1629-1666.

Cochrane, John H., 1997, Where is the Market Going? Uncertain Facts and Novel Theories, *Economic Perspectives,* 21, 3-37.

Constantinides, George M., John B. Donaldson, and Rajnish Mehra, 2002, Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle, *Quarterly Journal of Economics* 117: 269-296.

Constantinides, George M., 2002, Rational Asset Prices, *Journal of Finance,* 57, 1567-1591.

Cornell, Bradford, 1999, *The Equity Risk Premium: The Long-Run Future of the Stock Market* (New York: John Wiley & Sons).

Diamond, Peter A., 1999, What Stock Market Returns to Expect for the Future?, *Center for Retirement at Boston College number 2.*

Diamond, Peter A., 2001, What Stock Market Returns to Expect for the Future: An Update, *Estimating the Real Rate of Return on Stocks Over the Long Term*, manuscript presented to the Social Security Advisory Board. http://www.ssab.gov/estimated%20rate%20of%20return.pdf

Dimson, Elroy, Paul Marsh, and Mike Staunton, 2002, *Triumph of the Optimists: 101 Years of Global Investment Returns* (Princeton: Princeton University Press).

Elton, Edwin J., and Martin J. Gruber, 1995, *Modern Portfolio Theory and Investment Analysis*, 5th ed. (New York: John Wiley & Sons).

Fama, Eugene F., and Kenneth R. French, 2002, The Equity Premium, *Journal of Finance,* 57, 637-659.

Fama, Eugene F., and Kenneth R French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance,* 47, 427-465.

Goetzmann, William N., Roger G. Ibbotson, and Liang Peng, 2001, A New Historical Database for the NYSE 1815 to 1925: Performance and Predictability, *Journal of Financial Markets,* 4, 1-32.

Goss, Stephen C., 2001, Equity Yield Assumptions Used by the Office of the Chief Actuary, Social Security Administration, to Develop Estimates for Proposals with Trust

Fund and/or Individual Account Investments, *Estimating the Real Rate of Return on Stocks Over the Long Term*, manuscript presented to the Social Security Advisory Board.
http://www.ssab.gov/estimated%20rate%20of%20return.pdf

Graham, John R., and Campbell R. Harvey, 2002, Expectations of Equity Risk Premia, Volatility and Asymmetry, Working paper, Duke University.

Grinold, Richard, and Kenneth Kroner, 2002, The Equity Risk Premium, *Barclays Global Investors' Investment Insights* 5.

Harris, Robert S., and Felicia C. Marston, 2001, The Market Risk Premium: Expectational Estimates Using Analysts' Forecasts, *Journal of Applied Finance,* 11, 6-16.

Harvey, Andrew, 1990, *The Econometric Analysis of Time Series*, 2nd ed. (Cambridge, MA: MIT Press).

Ibbotson Associates, 2003, *Stocks, Bonds, Bills, and Inflation Yearbook* (Ibbotson Associates, Chicago).

Ibbotson Associates, 2003, *Stocks, Bonds, Bills, and Inflation Yearbook*, Valuation Edition (Ibbotson Associates, Chicago).

Ibbotson, Roger G., and Peng Chen, 2003, Long-Run Stock Returns: Participating in the Real Economy, *Financial Analysts Journal,* 59, 88-98.

Kaplan, Paul D., and James D. Peterson, 1998, Full-Information Industry Betas, *Financial Management,* 27, 85-93.

LaChance, Marie-Eve, and Olivia S. Mitchell, 2003, Understanding Individual Account Guarantees, University of Michigan Retirement Research Center working paper WP 2003-035.

Leibowitz, Martin L., 2001, Equity Risk Premium Forum, TIAA-CREF, New York City, November.
http://www.aimrpubs.org/ap/issues/v2002n1/toc.html

Longstaff, Francis A., 2002, The Flight-to-Liquidity Premium in U.S. Treasury Bond Prices, Working paper #9312, NBER, forthcoming in *Journal of Business*.

Mamaysky, Harry, 2002, A Model for Pricing Stocks and Bonds with Default Risk, Yale ICF Working Paper No. 02-13.

Malkiel, Burton G., 1999, *A Random Walk Down Wall Street* (New York: W. W. Norton & Company).

Mehra, Rajnish, 2002, Equity Premium Puzzle, in *Mastering Investments*, Ed. James Pickford (London: Financial Times-Prentice Hall).

Mehra, Rajnish, and Edward C. Prescott, 1985, The Equity Premium: A Puzzle, *Journal of Monetary Economics,* 15, 145-161.

Pratt, Shannon, P., 1998, *Cost of Capital Estimation and Applications*, John Wiley & Sons, New York.

Ralfe, John, Cliff Speed, and Jon Palin, 2003, Pensions and Capital Structure: Why hold equities in the pension fund?, Society of Actuaries Symposium, Vancouver, June.

Ross, Stephen A., 1976, The Arbitrage Theory of Capital Asset Pricing, *Journal of Economic Theory,* 13, 341-360.

Schwert, G. William, 1990, Indexes of United States Stock Prices from 1802 to 1987, *Journal of Business,* 63, 399-426.

Shiller, Robert, 1989, *Market Volatility* (Cambridge, MA: MIT Press).

Shoven, John B., 2001, What Are Reasonable Long-Run Rates of Return to Expect on Equities?, *Estimating the Real Rate of Return on Stocks Over the Long Term*, manuscript presented to the Social Security Advisory Board. http://www.ssab.gov/estimated%20rate%20of%20return.pdf

Siegel, Jeremy J., 1999, The Shrinking Equity Premium, *Journal of Portfolio Management,* 26, 10-17.

Siegel, Jeremy J., 2002, *Stocks for the Long Run: The Definitive Guide to Financial Market Returns and Long-Term Investment Strategies*, 3rd ed. (New York: McGraw-Hill).

Social Security Advisory Board, 2002, Fiscal Year Annual Report http://www.ssab.gov/annualreport2002.pdf

Tversky, Amos, and Daniel Kahneman, 1991, Loss Aversion and Riskless Choice: A Reference Dependent Model, *Quarterly Journal of Economics,* 152, 1039-1061.

Weil, Philippe, 1989, The Equity Premium Puzzle and the Risk-free Rate Puzzle, *Journal of Monetary Economics,* 24, 401-421.

Welch, Ivo, 2000, Views of Financial Economists on the Equity Premium and on Professional Controversies, *Journal of Business,* 73, 501-537.

Welch, Ivo, 2001, The Equity Premium Consensus Forecast Revisited, Cowles Foundation discussion paper, No. 1325.

Wendt, Richard, 2002, Understanding Equity Risk Premium, *Risk and Rewards*, No. 38, Society of Actuaries, February.

Wilson, Jack W., and Charles P. Jones, 2002, An Analysis of the S&P 500 Index and Cowles's Extensions: Price Indexes and Stock Returns, 1870-1999, *Journal of Business*, 75:3, 505-533.

Appendix A Ibbotson Market Data 1926-2002*			
	Common Stocks	**U. S. Treasury Bills**	**Arithmetic Short-Horizon**
Year	**Total Annual Returns**	**Total Annual Returns**	**Equity Risk Premia**
1926	11.62%	3.27%	8.35%
1927	37.49%	3.12%	34.37%
1928	43.61%	3.56%	40.05%
1929	- 8.42%	4.75%	-13.17%
1930	-24.90%	2.41%	-27.31%
1931	-43.34%	1.07%	-44.41%
1932	- 8.19%	0.96%	- 9.15%
1933	53.99%	0.30%	53.69%
1934	- 1.44%	0.16%	- 1.60%
1935	47.67%	0.17%	47.50%
1936	33.92%	0.18%	33.74%
1937	-35.03%	0.31%	-35.34%
1938	31.12%	- 0.02%	31.14%
1939	- 0.41%	0.02%	- 0.43%
1940	- 9.78%	0.00%	- 9.78%
1941	-11.59%	0.06%	-11.65%
1942	20.34%	0.27%	20.07%
1943	25.90%	0.35%	25.55%
1944	19.75%	0.33%	19.42%
1945	36.44%	0.33%	36.11%
1946	- 8.07%	0.35%	- 8.42%
1947	5.71%	0.50%	5.21%
1948	5.50%	0.81%	4.69%
1949	18.79%	1.10%	17.69%
1950	31.71%	1.20%	30.51%
1951	24.02%	1.49%	22.53%
1952	18.37%	1.66%	16.71%
1953	- 0.99%	1.82%	- 2.81%
1954	52.62%	0.86%	51.76%
1955	31.56%	1.57%	29.99%
1956	6.56%	2.46%	4.10%

Appendix A
Ibbotson Market Data 1926-2002*

	Common Stocks	U. S. Treasury Bills	Arithmetic Short-Horizon
Year	Total Annual Returns	Total Annual Returns	Equity Risk Premia
1957	-10.78%	3.14%	-13.92%
1958	43.36%	1.54%	41.82%
1959	11.96%	2.95%	9.01%
1960	0.47%	2.66%	- 2.19%
1961	26.89%	2.13%	24.76%
1962	- 8.73%	2.73%	-11.46%
1963	22.80%	3.12%	19.68%
1964	16.48%	3.54%	12.94%
1965	12.45%	3.93%	8.52%
1966	-10.06%	4.76%	-14.82%
1967	23.98%	4.21%	19.77%
1968	11.06%	5.21%	5.85%
1969	- 8.50%	6.58%	-15.08%
1970	4.01%	6.52%	- 2.51%
1971	14.31%	4.39%	9.92%
1972	18.98%	3.84%	15.14%
1973	-14.66%	6.93%	-21.59%
1974	-26.47%	8.00%	-34.47%
1975	37.20%	5.80%	31.40%
1976	23.84%	5.08%	18.76%
1977	- 7.18%	5.12%	-12.30%
1978	6.56%	7.18%	- 0.62%
1979	18.44%	10.38%	8.06%
1980	32.42%	11.24%	21.18%
1981	- 4.91%	14.71%	-19.62%
1982	21.41%	10.54%	10.87%
1983	22.51%	8.80%	13.71%
1984	6.27%	9.85%	- 3.58%
1985	32.16%	7.72%	24.44%
1986	18.47%	6.16%	12.31%
1987	5.23%	5.47%	- 0.24%
1988	16.81%	6.35%	10.46%
1989	31.49%	8.37%	23.12%

Appendix A Ibbotson Market Data 1926-2002*			
	Common Stocks	**U. S. Treasury Bills**	**Arithmetic Short-Horizon**
Year	**Total Annual Returns**	**Total Annual Returns**	**Equity Risk Premia**
1990	- 3.17%	7.81%	-10.98%
1991	30.55%	5.60%	24.95%
1992	7.67%	3.51%	4.16%
1993	9.99%	2.90%	7.09%
1994	1.31%	3.90%	- 2.59%
1995	37.43%	5.60%	31.83%
1996	23.07%	5.21%	17.86%
1997	33.36%	5.26%	28.10%
1998	28.58%	4.86%	23.72%
1999	21.04%	4.68%	16.36%
2000	- 9.11%	5.89%	-15.00%
2001	-11.88%	3.83%	-15.71%
2002	-22.10%	1.65%	-23.75%
mean=	**12.20%**	**3.83%**	**8.37%**
Standard Dev=	**20.49%**	**3.15%**	**20.78%**

* *2003 SBBI Yearbook pages 38 and 39*

Appendix B

Source	Risk-free-Rate	ERP Estimate	Real risk-free rate	Nominal risk-free rate	Geometric	Arithmetic	Long-horizon	Short-horizon	Short-run expectation	Long-run expectation	Conditional	Unconditional
Historical												
Ibbotson Associates	3.8% [7]	8.4% [31]		X		X		X		X		X
Social Security												
Office of the Chief Actuary [1]	2.3%,3.0% [8]	4.7%,4.0% [32]	X		X		X			X		X
John Campbell [2]	3% to 3.5% [9]	1.5-2.5%, 3-4% [33]	X		X	X	X	X		X	X	
Peter Diamond	2.2% [10]	<4.8% [34]	X		X		X			X	X	
Peter Diamond [3]	3.0% [11]	3.0% to 3.5% [35]	X		X		X			X	X	
John Shoven [4]	3.0%,3.5% [12]	3.0% to 3.5% [36]	X		X		X			X	X	
Puzzle Research												
Robert Arnott and Peter Bernstein	3.7% [13]	2.4% [37]	X		X		X			X	X	
Robert Arnott and Ronald Ryan	4.1% [14]	-0.9% [38]	X		X		X			X	X	
John Campbell and Robert Shiller	N/A	Negative [39]	X		?		?		X		X	
James Claus and Jacob Thomas	7.64% [15]	3.39% or less [40]		X		X	X			X	X	
George Constantinides	2.0% [16]	6.9% [41]	X			X		X		X		X
Bradford Cornell	5.6%, 3.8% [17]	3.5-5.5%, 5-7% [42]		X		X	X	X		X	X	
Dimson, Marsh, & Staunton	1.0% [18]	5.4% [43]	X			X		X		X	X	
Eugene Fama and Kenneth French	3.24% [19]	3.83% & 4.78% [44]	X			X		X		X		X
Robert Harris and Felicia Marston	8.53% [20]	7.14% [45]		X		X	X		X		X	
Roger Ibbotson and Peng Chen	2.05% [21]	4% and 6% [46]	X		X	X	X			X		X
Jeremy Siegel	4.0% [22]	-0.9% to -0.3% [47]	X		X		X			X	X	
Jeremy Siegel	3.5% [23]	2-3% [48]	X		X		X			?	X	
Surveys												
John Graham and Campbell Harvey	? by survey [24]	3-4.7% [49]		X		?	X		X		X	
Ivo Welch	N/A [25]	7% [50]		X		X		X		X	X	
Ivo Welch [5]	5% [26]	5.0% to 5.5% [51]		X		X		X		X	X	
Misc.												
Barclays Global Investors	5% [27]	2.5%, 3.25% [52]		X	X		X		X		X	
Richard Brealey and Stewart Myers	N/A [28]	6 to 8.5% [53]		X		X		X		X		X
Burton Malkiel	5.25% [29]	2.75% [54]		X	X		X			X	X	
Richard Wendt [6]	5.5% [30]	3.3% [55]		X		X	X			X	X	

Long-run expectation considered to be a forecast of more than 10 years.
Short-run expectation considered to be a forecast of 10 years or less.

Source	Risk-free Rate	ERP Estimate	Data Period	Methodology
Historical				
Ibbotson Associates	3.8% [7]	8.4% [31]	1926-2002	Historical
Social Security				
Office of the Chief Actuary [1]	2.3%, 3.0% [8]	4.7%, 4.0% [32]	1900-1995, Projecting out 75 years	Historical
John Campbell [2]	3% to 3.5% [9]	1.5-2.5%, 3-4% [33]	Projecting out 75 years	Historical & Ratios (Div/Price & Earn Gr)
Peter Diamond	2.2% [10]	<4.8% [34]	Last 200 yrs for eq/ 75 for bonds, Proj 75 yrs	Fundamentals: Div Yld, GDP Gr
Peter Diamond [3]	3.0% [11]	3.0% to 3.5% [35]	Projecting out 75 years	Fundamentals: Div/Price
John Shoven [4]	3.0%, 3.5% [12]	3.0% to 3.5% [36]	Projecting out 75 years	Fundamentals: P/E, GDP Gr
Puzzle Research				
Robert Arnott and Peter Bernstein	3.7% [13]	2.4% [37]	1802 to 2001, normal	Fundamentals: Div Yld & Gr
Robert Arnott and Ronald Ryan	4.1% [14]	-0.9% [38]	Past 74 years, 74 year projection [56]	Fundamentals: Div Yld & Gr
John Campbell and Robert Shiller	N/A	Negative [39]	1871 to 2000, ten-year projection	Ratios: P/E and Div/Price
James Claus and Jacob Thomas	7.64% [15]	3.39% or less [40]	1985-1998, long-term	Abnormal Earnings model
George Constantinides	2.0% [16]	6.9% [41]	1872 to 2000, long-term	Hist. and Fund.: Price/Div & P/E
Bradford Cornell	5.6%, 3.8% [17]	3.5-5.5%, 5-7% [42]	1926-1997, long run forward-looking	Weighing theoretical and empirical evid
Dimson, Marsh, & Staunton	1.0% [18]	5.4% [43]	1900-2000, prospective	Adj hist ret, Var of Gordon gr model
Eugene Fama and Kenneth French	3.24% [19]	3.83% & 4.78% [44]	Estimate for 1951-2000, long-term	Fundamentals: Dividends and Earnings
Robert Harris and Felicia Marston	8.53% [20]	7.14% [45]	1982-1998, expectational	Fin analysts' est, div gr model
Roger Ibbotson and Peng Chen	2.05% [21]	4% and 6% [46]	1926-2000, long-term	Historical and supply side approaches
Jeremy Siegel	4.0% [22]	-0.9% to -0.3% [47]	1871 to 1998, forward-looking	Fundamentals: P/E, Div Yld, Div Gr
Jeremy Siegel	3.5% [23]	2-3% [48]	1802-2001, forward-looking	Earnings yield
Surveys				
John Graham and Campbell Harvey	? by survey [24]	3-4.7% [49]	2Q 2000 thru 3Q 2002, 1 & 10 year proj	Survey of CFO's
Ivo Welch	N/A [25]	7% [50]	30-Year forecast, surveys in 97/98 & 99	Survey of financial economists
Ivo Welch [5]	5% [26]	5.0% to 5.5% [51]	30-Year forecast, survey around August 2001	Survey of financial economists
Misc.				
Barclays Global Investors	5% [27]	2.5%, 3.25% [52]	Long-run (10-year) expected return	Fundamentals: Inc, Earn Gr, & Repricing
Richard Brealey and Stewart Myers	N/A [28]	6 to 8.5% [53]	1926-1997	Predominantly Historical
Burton Malkiel	5.25% [29]	2.75% [54]	1926 to 1997, estimate millennium [57]	Fundamentals: Div Yld, Earn Gr
Richard Wendt [6]	5.5% [30]	3.3% [55]	1960-2000, estimate for 2001-2015 period	Linear regression model

Footnotes:

[1] Social Security Administration.

[2] Presented to the Social Security Advisory Board.

[3] Presented to the Social Security Advisory Board. Update of 1999 article.

[4] Presented to the Social Security Advisory Board.

[5] Update to Welch 2000.

[6] Newsletter of the Investment Section of the Society of Actuaries.

[7] Arithmetic mean of U.S. Treasury bills annual total returns from 1926-2002.

[8] 2.3% Long-run real yield on Treasury bonds; used for Advisory Council proposals. 3.0% Long-term real yield on Treasury bonds; used in 1999 Social Security Trustees Report.

[9] Estimate for safe real interest rates in the future based on yield of long-term inflation-indexed Treasury securities of 3.5% and short-term real interest rates recently averaging about 3%.

[10] Real long-term bond yield using 75 year historical average.

[11] Real yield on long-term Treasuries (assumption by OCACT).

[12] 3.0% is the OCACT assumption. 3.5% is the real return on long-run (30-year) inflation-indexed Treasury securities.

[13] Long-term expected real geometric bond return (10 year-horizon).

[14] The yield on US government inflation-indexed bonds (starting bond real yield in Jan 2000).

[15] Average 10-year Government T-bond yield between 1985 and 1998 (yield of 11.43% in 1985 to 5.64% in 1998. The mean 30-year risk-free rate for each year of the U.S. sample period is 31 basis points higher than the mean 10-year risk-free rate.

[16] Rolled-over real arithmetic return of three-month Treasury bills and certificates.

[17] Historical 20-year Treasury bond return of 5.6%. Yield on 20-year Treasury bonds in 1998 was approximately 6%. Historical 1-month Treasury bill return of 3.8%. Yield on 1-month Treasury bills in 1998 was approximately 4%.

[18] United States historical arithmetic real Treasury bill return over 1900-2000 period. 0.9% geometric Treasury bill return.

[19] Average real return on six-month commercial paper (proxy for risk-free interest rate). Substituting the one-month Treasury bill rate for the six-month commercial paper rate causes estimates of the annual equity premium for 1951-2000 to rise by about 1.00%.

[20] Average yield to maturity on long-term U.S. government bonds, 1982-1998.

[21] Real, geometric risk-free rate. Geometric risk-free rate with inflation (nominal) 5.13%. Nominal yield equivalent to historical geometric long-term government bond income return for 1926-2000.

[22] The ten- and thirty-year TIPS bond yielded 4.0% in August 1999.

[23] Return on inflation-indexed securities.

[24] Current 10-year Treasury bond yield. Survey administered from June 6, 2000 to June 4, 2002. The rate on the 10-year Treasury bond changes in each survey. For example, in the Dec. 1, 2000 survey, the current annual yield on the 10-year Treasury bond was 5.5%. For the June 6, 2001 survey, the current annual yield on the 10-year Treasury bond was 5.3%.

[25] Arithmetic per-annum average return on rolled-over 30-day T-bills.

[26] Average forecast of arithmetic risk-free rate of about 5% by deducting ERP from market return.

[27] Current nominal 10-year bond yield.

[28] Return on Treasury bills. Treasury bills yield of about 5 percent in mid-1998. Average historical return on Treasury bills 3.8 percent.

[29] Good quality corporate bonds will earn approximately 6.5% to 7%. Long-term zero-coupon Treasury bonds will earn about 5.25%. Long-term TIPS will earn a real return of 3.75%.

[30] 1/1/01 Long T-Bond yield; uses initial bond yields in predictive model.

[31] Arithmetic short-horizon expected equity risk premium. Arithmetic intermediate-horizon expected equity risk premium 7.4%. Arithmetic long-horizon expected equity risk premium 7.0%. Geometric short-horizon expected equity risk premium 6.4%.

[32] Geometric equity premium over long-term Treasury securities. OCACT assumes a constant geometric real 7.0% stock return.

[33] Long-run average equity premium of 1.5% to 2.5% in geometric terms and 3% to 4% in arithmetic terms.

[34] Lower return over the next decade, followed by a geometric, real 7.0% stock return for remaining 65 years or lower rate of return for entire 75-year period (obscures pattern of returns).

[35] Most likely poor return over the next decade followed by a return to historic yields. Working from OCACT stock return assumption, he gives a single rate of return on equities for projection purposes of 6.0 to 6.5% (geometric, real).

[36] Geometric real stock return over the geometric real return on long-term government bonds.

[37] Expected geometric return over long-term government bonds. Their current risk premium is approximately zero, and their recommended expectation for the future real return for both stocks and bonds is 2-4 percent. The "normal" level of the risk premium is modest (2.4 percent or quite possibly less).

[38] Geometric real returns on stocks are likely to be in the 3%-4% range for the foreseeable future (10-20 years).

[39] Substantial declines in real stock prices, and real stock returns below zero, over the next ten years (2001-2010).

[40] The equity premium for each year between 1985 and 1998 in the United States. Similar results for five other markets.

[41] Unconditional, arithmetic mean aggregate equity premium over the 1872-2000 period. Over the period 1951 to 2000, the adjusted estimate of the unconditional mean premium is 6.0%. The corresponding estimate over the 1926 to 2000 period is 8.0%. Sharp distinction between conditional, short-term forecasts of the mean equity return and premium and estimates of the unconditional mean.

[42] Long run arithmetic future ERP of 3.5% to 5.5% over Treasury bonds and 5% to 7% over Treasury bills. Compares estimates to historical returns of 7.4% for bond premium and 9.2% for bill premium.

[43] 5.4% United States arithmetic expected future ERP relative to bills. 4.0% World (16 countries) arithmetic expected future ERP relative to bills. 4.1% United States geometric expected future ERP relative to bills. 3.0% World (16 countries) geometric expected future ERP relative to bills.

[44] 3.83% unconditional expected annual simple equity premium return (referred to as the annual-bias adjusted estimate of the annual equity premium) using dividend growth model. 4.78% unconditional expected annual simple equity premium return (referred to as the annual-bias adjusted estimate of the annual equity premium) using earnings growth model. Compares these results against historical real equity risk premium of 7.43% for 1951-2000.

[45] Average expectational risk premium. Because of the possible bias of analysts' optimism, the estimates are interpreted as "upper bounds" for the market premium. The average expectational risk premium is approximately equal to the arithmetic (7.5%) long-term differential between returns on stocks and long-term government bonds.

[46] 4% geometric (real) and 6% arithmetic (real). Forward looking long-horizon sustainable equity risk premium.

[47] Using the dividend discount model, the forward-looking real long-term geometric return on equity is 3.3%. Based on the earnings yield, the forward-looking real long-term geometric retum on equity is between 3.1% and 3.7%.

[48] Future geometric equity premium. Future real return on equities of about 6%.

[49] The 10-year premium. The one-year risk premium averages between 0.4 and 5.2% depending on the quarter surveyed.

[50] Arithmetic 30-year forecast relative to short-term bills; 10-year same estimate. Second survey 6.8% for 30 and 10-year estimate. 1-year horizon between 0.5% and 1.5% lower. Geometric 30-year forecast around 5.2% (50% responded to this question).

[51] Arithmetic 30-year equity premium (relative to short-term T-bills). Geometric about 50 basis points below arithmetic. Arithmetic 1-year equity premium 3 to 3.5%.

[52] 2.5% current (conditional) geometric equity risk premium. 3.25% long-run, geometric normal or equilibrium equity risk premium.

[53] Extra arithmetic return versus Treasury bills. "Brealey and Myers have no official position on the exact market risk premium, but we believe a range of 6 to 8.5 percent is reasonable for the United States. We are most comfortable with figures towards the upper end of the range."

[54] The projected geometric (nominal) total return for the S&P 500 is 8 percent per year.

[55] Arithmetic mean 15 year horizon.

[56] 74 years since Dec 1925 and 74 years starting Jan 2000.

[57] Estimate the early decades of the twenty-first century.

Appendix C
Estimating a Short-Horizon Arithmetic Unconditional Equity Risk Premium

Source	Risk-free Rate	ERP Estimate	Stock Return Estimate	Geometric to arithmetic	Real to nominal	Conditional to unconditional [60]	Fixed short-horizon RFR	Short-horizon arithmetic unconditional ERP estimate
	I	II	III	IV	V	VI	VII	VIII
Historical								
Ibbotson Associates	3.8% [7]	8.4% [31]	12.2%	0.0%	0.0%	0.00%	3.8%	8.4%
Social Security								
Office of the Chief Actuary [1]	2.3%,3.0% [8]	4.7%,4.0% [32]	7.0%	2.0%	3.1%	0.00%	3.8%	8.3%
John Campbell [2]	3% to 3.5% [9]	1.5-2.5%, 3-4% [33]	6.0%-7.5%	0.0%	3.1%	0.46%	3.8%	5.8%-7.3%
Peter Diamond	2.2% [10]	<4.8% [34]	<7.0%	2.0%	3.1%	0.46%	3.8%	<8.8%
Peter Diamond [3]	3.0% [11]	3.0% to 3.5% [35]	6.0%-6.5%	2.0%	3.1%	0.46%	3.8%	7.8%-8.3%
John Shoven [4]	3.0%,3.5% [12]	3.0% to 3.5% [36]	6.0%-7.0%	2.0%	3.1%	0.46%	3.8%	7.8%-8.8%
Puzzle Research								
Robert Arnott and Peter Bernstein	3.7% [13]	2.4% [37]	6.1%	2.0%	3.1%	0.46%	3.8%	7.9%
Robert Arnott and Ronald Ryan	4.1% [14]	-0.9% [38]	3.2%	2.0%	3.1%	0.46%	3.8%	5.0%
John Campbell and Robert Shiller	N/A	Negative [39]	Negative	N/A	N/A	N/A	N/A	N/A
James Claus and Jacob Thomas	7.64% [15]	3.39% or less [40]	11.03%	0.0%	0.0%	0.46%	3.8%	7.69%
George Constantinides	2.0% [16]	6.9% [41]	8.9%	0.0%	3.1%	0.00%	3.8%	8.2%
Bradford Cornell	5.6%, 3.8% [17]	3.5-5.5%, 5-7% [42]	8.8%-10.8%	0.0%	0.0%	0.46%	3.8%	5.5%-7.5%
Dimson, Marsh, & Staunton	1.0% [18]	5.4% [43]	6.4% [58]	0.0%	3.1%	0.46%	3.8%	6.2% [61]
Eugene Fama and Kenneth French	3.24% [19]	3.83% & 4.78% [44]	7.07%-8.02%	0.0%	3.1%	0.00%	3.8%	6.37%-7.32%
Robert Harris and Felicia Marston	8.53% [20]	7.14% [45]	12.34% [59]	0.0%	0.0%	0.46%	3.8%	9.00%
Roger Ibbotson and Peng Chen	2.05% [21]	4% and 6% [46]	8.05%	0.0%	3.1%	0.00%	3.8%	7.35%
Jeremy Siegel	4.0% [22]	-0.9% to -0.3% [47]	3.1%-3.7%	2.0%	3.1%	0.46%	3.8%	4.9%-5.5%
Jeremy Siegel	3.5% [23]	2-3% [48]	5.5%-6.5%	2.0%	3.1%	0.46%	3.8%	7.3%-8.3%
Surveys								
John Graham and Campbell Harvey	? by survey [24]	3-4.7% [49]	8.3%-10.2%	N/A	0.0%	0.46%	3.8%	5.0%-6.9%
Ivo Welch	N/A [25]	7% [50]	N/A	0.0%	0.0%	0.46%	0.0%	7.5%
Ivo Welch [5]	5% [26]	5.0% to 5.5% [51]	10.0%-10.5%	0.0%	0.0%	0.46%	3.8%	6.7%-7.2%
Misc.								
Barclays Global Investors	5% [27]	2.5%, 3.25% [52]	7.5%,8.25%	2.0%	0.0%	0.46%	3.8%	6.16%-6.91%
Richard Brealey and Stewart Myers	N/A [28]	6 to 8.5% [53]	N/A	0.0%	0.0%	0.00%	0.0%	6.0%-8.5%
Burton Malkiel	5.25% [29]	2.75% [54]	8.0%	2.0%	0.0%	0.46%	3.8%	6.7%
Richard Wendt [6]	5.5% [30]	3.3% [55]	8.8%	0.0%	0.0%	0.46%	3.8%	5.5%

Column formulas:
III = I + II
VIII = III + IV + V + VI –VII

Source for adjustments:
2003 Ibbotson Yearbook Table 2-1 page 33
Fama French 2002 (see footnote 60)

Footnotes (1-57 from Appendix B):
[58] World estimate of 5.0%.
[59] Long risk-free of 5.2% plus 7.14%.
[60] For the 1951-2000 period, Fama and French (2002) adjust the conditional dividend growth model estimate upwards by 1.28% for an unconditional estimate, and they make a 0.46% upwards adjustment to the earnings growth model. We select the smaller of the two as an approximate minimum adjustment. For the longer period of 1872-2000, a comparable adjustment would be 0.82% for the dividend growth model and 0.54% for the 1872-1950 period using a dividend growth model. Earnings growth rates are shown by Fama and French only for the 1951-2000 period.
[61] World estimate of 4.8%.

Appendix D

Historical and Forecasted Equity Returns- All Ibbotson and Chen Models (Percent).

Method/ Model	Sum	Inflation	Real Risk-Free Rate	Equity Risk Premium	Real Capital Gain	g(Real EPS)	g(Real Div)	- g (Pay out Ratio)	g (BV)	g (ROE)	g P/E)	g(Real GDP/ POP)	g(FS-GDP/ POP)	Income Return	Re-Investment + Interaction	Additional Growth	Forecast Earnings Growth
Column #	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII
Historical																	
Method 1	10.70	3.08	2.05	5.24											0.33		
Method 2	10.70	3.08			3.02									4.28	0.32		
Method 3	10.70	3.08				1.75					1.25			4.28	0.34		
Method 4	10.70	3.08					1.23	0.51			1.25			4.28	0.35		
Method 5	10.70	3.08							1.46	0.31	1.25			4.28	0.31		
Method 6	10.70	3.08										2.04	0.96	4.28	0.32		
Forecast with Historical Dividend Yield																	
Model 3F	9.37	3.08				1.75								4.28	0.26		
Model 3F (ERP)	9.37	3.08	2.05	3.97											0.27		
Forecast with Current Dividend Yield																	
Model 4F	5.44	3.08					1.23							1.10 [a]	0.03		
Model 4F (ERP)	5.44	3.08	2.05	0.24											0.07		
Model $4F_2$	9.37	3.08					1.23	0.51						2.05 [b]	0.21	2.28	
Model $4F_2$ (FG)	9.37	3.08												1.10 [a]	0.21		4.98

Source: The data and format was made available by Ibbotson/Chen and is reprinted with permission by Ibbotson Associates. Corresponds to Ibbotson/Chen Table 2 Exhibit; column numbers have been added.

[a] 2000 dividend yield

[b] Assuming the historical average dividend-payout ratio, the 2000 dividend yield is adjusted up 0.95 pps.

	Formula*	Description of Method
Historical		
Method 1	I=(1+II)*(1+III)*(1+IV)-1	Building Blocks Method: inflation, real risk-free rate, and equity risk premium.
Method 2	I=[(1+II)*(1+V)-1]+XIV+XV	Capital Gain and Income Method: inflation, real capital gain, and income return.
Method 3	I=[(1+II)*(1+VI)*(1+XI)-1]+XIV+XV	Earnings Model: inflation, growth in earnings per share, growth in price to earnings ratio, and income return.
Method 4	I=[(1+II)*(1+XI)*(1+VII)/(1-VIII)-1]+XIV+XV	Dividends Model: inflation, growth rate of price earnings ratio, growth rate of the dollar amount of dividends after inflation, growth rate of payout ratio, and dividend yield (income return).
Method 5	I=[(1+II)*(1+XI)*(1+IX)*(1+X)-1]+XIV+XV	Return on Book Equity Model: inflation, growth rate of price earnings ratio, growth rate of book value, growth rate of ROE, and income return.
Method 6	I=[(1+II)*(1+XII)*(1+XIII)-1]+XIV+XV	GDP Per Capita Model: inflation, real growth rate of the overall economic productivity (GDP per capita), increase of the equity market relative to the overall economic productivity, and income return.
Forecast with Historical Dividend Yield		
Model 3F	I=[(1+II)*(1+VI)-1]+XIV+XV	Forward-looking Earnings Model: inflation, growth in real earnings per share, and income return.
Model 3F (ERP)	IV=(1+I)/[(1+II)*(1+III)]-1	Using Model 3F result to calculate ERP.
Forecast with Current Dividend Yield		
Model 4F	I=[(1+II)*(1+VII)-1]+XIV+XV	Forward-looking Dividends Model: inflation, growth in real dividend, and dividend yield (income return); also referred to as Gordon model.
Model 4F (ERP)	IV=(1+I)/[(1+II)*(1+III)]-1	Using Model 4F result to calculate ERP.
Model $4F_2$	I=[(1+II)*(1+VII)*(1+VIII)-1]+XIV+XV+XVI	Attempt to reconcile Model 4F and Model 3F.
Model $4F_2$ (FG)	XVII=[(1+I)/(1+II)-1]-XIV-XV	Using Method $4F_2$ result to calculate forecasted earnings.

Explanation of Ibbotson/Chen Table 2 Exhibit; using column numbers to represent formula.

High stock prices, together with projected slow economic growth, are not consistent with the 7.0 percent return that the Office of the Chief Actuary has generally used when evaluating proposals with stock investments. Routes out of the inconsistency include assuming higher GDP growth, a lower long-run stock return, or a lower short-run stock return with a 7.0 percent return on a lower base thereafter. In short, either the stock market is overvalued and requires a correction to justify a 7.0 percent return thereafter, or it is correctly valued and the long-run return is substantially lower than 7.0 percent (or some combination of the two). This article argues that the former view is more convincing, since accepting the "correctly valued" hypothesis implies an implausibly small equity premium.

This article originally appeared as an Issue in Brief of the Center for Retirement Research at Boston College (No. 2, September 1999). The research reported herein was performed pursuant to a grant from the Social Security Administration (SSA) funded as part of the Retirement Research Consortium. The opinions and conclusions expressed are solely those of the author and should not be construed as representing the opinions or policy of SSA, any agency of the federal government, or the Center for Retirement Research at Boston College.

*The author is a professor at the Massachusetts Institute of Technology. *(Manuscript received November 1999; submitted for external review December 1999; revise and resubmit recommended April 2000; revision received June 2000; paper accepted July 2000.)*

What Stock Market Returns to Expect for the Future?

*by Peter A. Diamond**

Summary

In evaluating proposals for reforming Social Security that involve stock investments, the Office of the Chief Actuary (OCACT) has generally used a 7.0 percent real return for stocks. The 1994-96 Advisory Council specified that OCACT should use that return in making its 75-year projections of investment-based reform proposals. The assumed ultimate real return on Treasury bonds of 3.0 percent implies a long-run equity premium of 4.0 percent. There are two equity-premium concepts: the *realized* equity premium, which is measured by the actual rates of return; and the *required* equity premium, which investors expect to receive for being willing to hold available stocks and bonds. Over the past two centuries, the realized premium was 3.5 percent on average, but 5.2 percent for 1926 to 1998.

Some critics argue that the 7.0 percent projected stock returns are too high. They base their arguments on recent developments in the capital market, the current high value of the stock market, and the expectation of slower economic growth.

Increased use of mutual funds and the decline in their costs suggest a lower required premium, as does the rising fraction of the American public investing in stocks. The size of the decrease is limited, however, because the largest cost savings do not apply to the very wealthy and to large institutional investors, who hold a much larger share of the stock market's total value than do new investors. These trends suggest a lower equity premium for projections than the 5.2 percent of the past 75 years. Also, a declining required premium is likely to imply a temporary increase in the realized premium because a rising willingness to hold stocks tends to increase their price. Therefore, it would be a mistake during a transition period to extrapolate what may be a temporarily high realized return. In the standard (Solow) economic growth model, an assumption of slower long-run growth lowers the marginal product of capital if the savings rate is constant. But lower savings as growth slows should partially or fully offset that effect.

The present high stock prices, together with projected slow economic growth, are not consistent with a 7.0 percent return. With a plausible level of adjusted dividends (dividends plus net share repurchases), the ratio of stock value to gross domestic product (GDP) would rise more than 20-fold over 75 years. Similarly, the steady-state Gordon formula—that stock returns equal the adjusted dividend yield plus the growth rate of stock prices (equal to that of GDP)—suggests a return of roughly 4.0 percent to 4.5 percent. Moreover, when relative stock values have been high, returns over the following decade have tended to be low.

To eliminate the inconsistency posed by the assumed 7.0 percent return, one could assume higher GDP growth, a lower long-run stock return, or a lower short-run stock return with a 7.0 percent return on a lower base thereafter. For example, with an adjusted dividend yield of 2.5 percent to 3.0 percent,

the market would have to decline about 35 percent to 45 percent in real terms over the next decade to reach steady state.

In short, either the stock market is overvalued and requires a correction to justify a 7.0 percent return thereafter, or it is correctly valued and the long-run return is substantially lower than 7.0 percent (or some combination). This article argues that the "overvalued" view is more convincing, since the "correctly valued" hypothesis implies an implausibly small equity premium. Although OCACT could adopt a lower rate for the entire 75-year period, a better approach would be to assume lower returns over the next decade and a 7.0 percent return thereafter.

Introduction

All three proposals of the 1994-96 Advisory Council on Social Security (1997) included investment in equities. For assessing the financial effects of those proposals, the Council members agreed to specify a 7.0 percent long-run real (inflation-adjusted) yield from stocks.[1] They devoted little attention to different short-run returns from stocks.[2] The Social Security Administration's Office of the Chief Actuary (OCACT) used this 7.0 percent return, along with a 2.3 percent long-run real yield on Treasury bonds, to project the impact of the Advisory Council's proposals.

Since then, OCACT has generally used 7.0 percent when assessing other proposals that include equities.[3] In the 1999 Social Security Trustees Report, OCACT used a higher long-term real rate on Treasury bonds of 3.0 percent.[4] In the first 10 years of its projection period, OCACT makes separate assumptions about bond rates for each year and assumes slightly lower real rates in the short run.[5] Since the assumed bond rate has risen, the assumed equity premium, defined as the difference between yields on equities and Treasuries, has declined to 4.0 percent in the long run.[6] Some critics have argued that the assumed return on stocks and the resulting equity premium are still too high.[7]

This article examines the critics' arguments and, rather than settling on a single recommendation, considers a range of assumptions that seem reasonable.[8] The article:

- Reviews the historical record on rates of return,
- Assesses the critics' reasons why future returns may be different from those in the historical record and examines the theory about how those rates are determined, and
- Considers two additional issues: the difference between gross and net returns, and investment risk.

Readers should note that in this discussion, a decline in the equity premium need not be associated with a decline in the return on stocks, since the return on bonds could increase. Similarly, a decline in the return on stocks need not be associated with a decline in the equity premium, since the return on bonds could also decline. Both rates of return and the equity premium are relevant to choices about Social Security reform.

Historical Record

Realized rates of return on various financial instruments have been much studied and are presented in Table 1.[9] Over the past 200 years, stocks have produced a real return of 7.0 percent per year. Even though annual returns fluctuate enormously, and rates vary significantly over periods of a decade or two, the return on stocks over very long periods has been quite stable (Siegel 1999).[10] Despite that long-run stability, great uncertainty surrounds both a projection for any particular period and the relevance of returns in any short period of time for projecting returns over the long run.

The equity premium is the difference between the rate of return on stocks and on an alternative asset—Treasury bonds, for the purpose of this article. There are two concepts of equity premiums. One is the *realized* equity premium, which is measured by the actual rates of return. The other is the *required* equity premium, which equals the premium that investors expect to get in exchange for holding available quantities of assets. The two concepts are closely related but different—significantly different in some circumstances.

The realized equity premium for stocks relative to bonds has been 3.5 percent for the two centuries of available data, but it has increased over time (Table 2).[11, 12] That increase has resulted from a significant decline in bond returns over the past

Table 1.
Compound annual real returns, by type of investment, 1802-1998 (in percent)

Period	Stocks	Bonds	Bills	Gold	Inflation
1802-1998	7.0	3.5	2.9	-0.1	1.3
1802-1870	7.0	4.8	5.1	0.2	0.1
1871-1925	6.6	3.7	3.2	-0.8	0.6
1926-1998	7.4	2.2	0.7	0.2	3.1
1946-1998	7.8	1.3	0.6	-0.7	4.2

Source: Siegel (1999).

Table 2.
Equity premiums: Differences in annual rates of return between stocks and fixed-income assets, 1802-1998

	Equity premium (percent)	
Period	With bonds	With bills
1802-1998	3.5	5.1
1802-1870	2.2	1.9
1871-1925	2.9	3.4
1926-1998	5.2	6.7
1946-1998	6.5	7.2

Source: Siegel (1999).

200 years. The decline is not surprising considering investors' changing perceptions of default risk as the United States went from being a less-developed country (and one with a major civil war) to its current economic and political position, where default risk is seen to be virtually zero.[13]

These historical trends can provide a starting point for thinking about what assumptions to use for the future. Given the relative stability of stock returns over time, one might initially choose a 7.0 percent assumption for the return on stocks—the average over the entire 200-year period. In contrast, since bond returns have tended to decline over time, the 200-year number does not seem to be an equally good basis for selecting a long-term bond yield. Instead, one might choose an assumption that approximates the experience of the past 75 years—2.2 percent, which suggests an equity premium of around 5.0 percent. However, other evidence, discussed below, argues for a somewhat lower value.[14]

Why Future Returns May Differ from Past Returns

Equilibrium and Long-Run Projected Rates of Return

The historical data provide one way to think about rates of return. However, thinking about how the future may be different from the past requires an underlying theory about how those returns are determined. This section lists some of the actions by investors, firms, and government that combine to determine equilibrium; it can be skipped without loss of continuity.

In asset markets, the demand by individual and institutional investors reflects a choice among purchasing stocks, purchasing Treasury bonds, and making other investments.[15] On the supply side, corporations determine the supplies of stocks and corporate bonds through decisions on dividends, new issues, share repurchases, and borrowing. Firms also choose investment levels. The supplies of Treasury bills and bonds depend on the government's budget and debt management policies as well as monetary policy. Whatever the supplies of stocks and bonds, their prices will be determined so that the available amounts are purchased and held by investors in the aggregate.

The story becomes more complicated, however, when one recognizes that investors base decisions about portfolios on their projections of both future prices of assets and future dividends.[16] In addition, market participants need to pay transactions costs to invest in assets, including administrative charges, brokerage commissions, and the bid-ask spread. The risk premium relevant for investors' decisions should be calculated net of transactions costs. Thus, the greater cost of investing in equities than in Treasuries must be factored into any discussion of the equity premium.[17] Differences in tax treatments of different types of income are also relevant (Gordon 1985; Kaplow 1994).

In addition to determining the supplies of corporate stocks and bonds, corporations also choose a debt/equity mix that affects the risk characteristics of both bonds and stocks. Financing a given level of investment more by debt and less by equity leaves a larger interest cost to be paid from the income of corporations before determining dividends. That makes both the debt and the equity more risky. Thus, changes in the debt/equity mix (possibly in response to prevailing stock market prices) should affect risk and, therefore, the equilibrium equity premium.[18]

Since individuals and institutions are generally risk averse when investing, greater expected variation in possible future yields tends to make an asset less valuable. Thus, a sensible expectation about long-run equilibrium is that the expected yield on equities will exceed that on Treasury bonds. The question at hand is how much more stocks should be expected to yield.[19] That is, assuming that volatility in the future will be roughly similar to volatility in the past, how much more of a return from stocks would investors need to expect in order to be willing to hold the available supply of stocks. Unless one thought that stock market volatility would collapse, it seems plausible that the premium should be significant. For example, equilibrium with a premium of 70 basis points (as suggested by Baker 1999a) seems improbable, especially since transactions costs are higher for stock than for bond investments. In considering this issue, one needs to recognize that a greater willingness to bear the risk associated with stocks is likely to be accompanied by greater volatility of stock prices if bond rates are unchanged. That is, fluctuations in expected growth in corporate profits will have bigger impacts on expected discounted returns (which approximate prices) when the equity premium, and so the discount rate, is lower.[20]

Although stocks should earn a significant premium, economists do not have a fully satisfactory explanation of why stocks have yielded so much more than bonds historically, a fact that has been called the equity-premium puzzle (Mehra and Prescott 1985; Cochrane 1997). Ongoing research is trying to develop more satisfactory explanations, but the theory still has inadequacies.[21] Nevertheless, to explain why the future may be different from the past, one needs to rely on some theoretical explanation of the past in order to have a basis for projecting a different future.

Commentators have put forth three reasons as to why future returns may be different from those in the historical record. First, past and future long-run trends in the capital market may imply a decline in the equity premium. Second, the current valuation of stocks, which is historically high relative to various benchmarks, may signal a lower future rate of return on equities. Third, the projection of slower economic growth may suggest a lower long-run marginal product of capital, which is the source of returns to financial assets. The first two issues are discussed in the context of financial markets; the third, in the context of physical assets. One should distinguish between arguments that suggest a lower equity premium and those that suggest lower returns to financial assets generally.

Equity Premium and Developments in the Capital Market

The capital market has experienced two related trends—the decrease in the cost of acquiring a diversified portfolio of stocks and the spread of stock ownership more widely in the economy. The relevant equity premium for investors is the equity premium net of the costs of investing. Thus, if the cost of investing in some asset decreases, that asset should have a higher price and a lower expected return gross of investment costs. The availability of mutual funds and the decrease in the cost of purchasing them should lower the equity premium in the future relative to long-term historical values. Arguments have also been raised about investors' time horizons and their understanding of financial markets, but the implications of those arguments are less clear.

Mutual Funds. In the absence of mutual funds, small investors would need to make many small purchases in different companies in order to acquire a widely diversified portfolio. Mutual funds provide an opportunity to acquire a diversified portfolio at a lower cost by taking advantage of the economies of scale in investing. At the same time, these funds add another layer of intermediation, with its costs, including the costs of marketing the funds.

Nevertheless, as the large growth of mutual funds indicates, many investors find them a valuable way to invest. That suggests that the equity premium should be lower in the future than in the past, since greater diversification means less risk for investors. However, the significance of the growth of mutual funds depends on the importance in total equity demand of "small" investors who purchase them, since this argument is much less important for large investors, particularly large institutional investors. According to recent data, mutual funds own less than 20 percent of U.S. equity outstanding (Investment Company Institute 1999).

A second development is that the average cost of investing in mutual funds has decreased. Rea and Reid (1998) report a drop of 76 basis points (from 225 to 149) in the average annual charge of equity mutual funds from 1980 to 1997. They attribute the bulk of the decline to a decrease in the importance of front-loaded funds (funds that charge an initial fee when making a deposit in addition to annual charges). The development and growth of index funds should also reduce costs, since index funds charge investors considerably less on average than do managed funds while doing roughly as well in gross rates of return. In a separate analysis, Rea and Reid (1999) also report a decline of 38 basis points (from 154 to 116) in the cost of bond mutual funds over the same period, a smaller drop than with equity mutual funds. Thus, since the cost of stock funds has fallen more than the cost of bond funds, it is plausible to expect a decrease in the equity premium relative to historical values. The importance of that decline is limited, however, by the fact that the largest cost savings do not apply to large institutional investors, who have always faced considerably lower charges.

A period with a declining required equity premium is likely to have a temporary increase in the realized equity premium. Assuming no anticipation of an ongoing trend, the divergence occurs because a greater willingness to hold stocks, relative to bonds, tends to increase the price of stocks. Such a price rise may yield a realized return that is higher than the required return.[22] The high realized equity premium since World War II may be partially caused by a decline in the required equity premium over that period. During such a transition period, therefore, it would be a mistake to extrapolate what may be a temporarily high realized return.

Spread of Stock Ownership. Another trend that would tend to decrease the equity premium is the rising fraction of the American public investing in stocks either directly or indirectly through mutual funds and retirement accounts (such as 401(k) plans). Developments in tax law, pension provision, and the capital markets have expanded the base of the population who are sharing in the risks associated with the return to corporate stock. The share of households investing in stocks in any form increased from 32 percent in 1989 to 41 percent in 1995 (Kennickell, Starr-McCluer, and Sundén 1997). Numerous studies have concluded that widening the pool of investors sharing in stock market risk should lower the equilibrium risk premium (Mankiw and Zeldes 1991; Brav and Geczy 1996; Vissing-Jorgensen 1997; Diamond and Geanakoplos 1999; Heaton and Lucas 2000). The importance of that trend must be weighted by the low size of investment by such new investors.[23]

Investors' Time Horizons. A further issue relevant to the future of the equity premium is whether the time horizons of investors, on average, have changed or will change.[24] Although the question of how time horizons should affect demands for assets raises subtle theoretical issues (Samuelson 1989), longer horizons and sufficient risk aversion should lead to greater willingness to hold stocks given the tendency for stock prices to revert toward their long-term trend (Campbell and Viceira 1999).[25]

The evidence on trends in investors' time horizons is mixed. For example, the growth of explicit individual retirement savings vehicles, such as individual retirement accounts (IRAs) and 401(k)s, suggests that the average time horizons of individual investors may have lengthened. However, some of that growth is at the expense of defined benefit plans, which may have longer horizons. Another factor that might suggest a longer investment horizon is the increase in equities held by institutional investors, particularly through defined benefit pension plans. However, the relevant time horizon for such holdings may not be the open-ended life of the plan but rather the horizon of the plans' asset managers, who may have career concerns that shorten the relevant horizon.

Other developments may tend to lower the average horizon. Although the retirement savings of baby boomers may currently add to the horizon, their aging and the aging of the population generally will tend to shorten horizons. Finally, individual stock ownership has become less concentrated (Poterba and Samwick 1995), which suggests a shorter time horizon because less wealthy investors might be less concerned about passing assets on to younger generations. Overall, without detailed calculations that would go beyond the scope of this article, it is not clear how changing time horizons should affect projections.

Investors' Understanding. Another factor that may affect the equity premium is investors' understanding of the properties of stock and bond investments. The demand for stocks might be affected by the popular presentation of material, such as Siegel (1998), explaining to the general public the difference between short- and long-run risks. In particular, Siegel highlights the risks, in real terms, of holding nominal bonds. While the creation of inflation-indexed Treasury bonds might affect behavior, the lack of wide interest in those bonds (in both the United States and the United Kingdom) and the failure to fully adjust future amounts for inflation generally (Shafir, Diamond, and Tversky 1997) suggest that nominal bonds will continue to be a major part of portfolios. Perceptions that those bonds are riskier than previously believed would then tend to decrease the required equity premium.

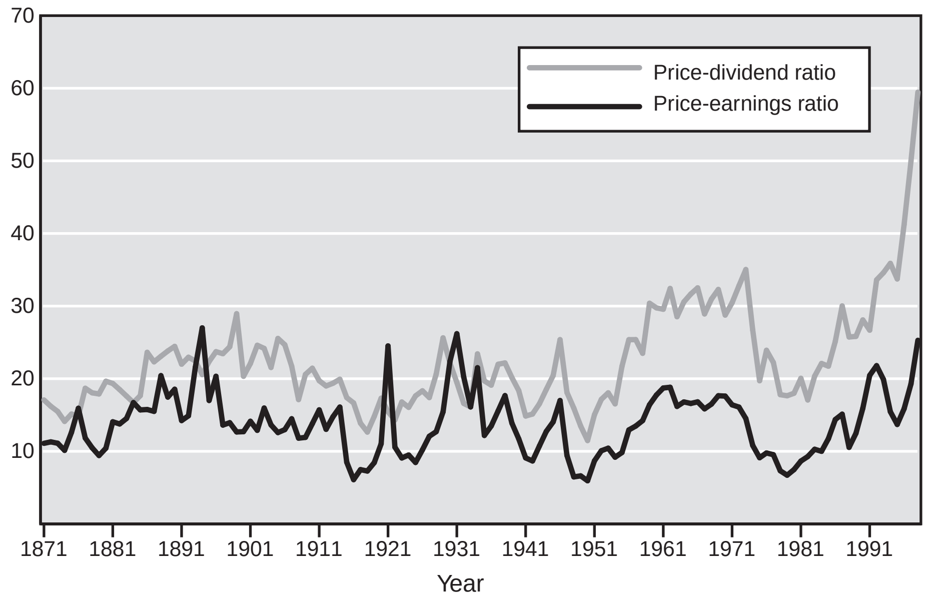
Popular perceptions may, however, be excessively influenced by recent events—both the high returns on equity and the low rates of inflation. Some evidence suggests that a segment of the public generally expects recent rates of increase in the prices of assets to continue, even when those rates seem highly implausible for a longer term (Case and Shiller 1988). The possibility of such extrapolative expectations is also connected with the historical link between stock prices and inflation. Historically, real stock prices have been adversely affected by inflation in the short run. Thus, the decline in inflation expectations over the past two decades would be associated with a rise in real stock prices if the historical pattern held. If investors and analysts fail to consider such a connection, they might expect robust growth in stock prices to continue without recognizing that further declines in inflation are unlikely. Sharpe (1999) reports evidence that stock analysts' forecasts of real growth in corporate earnings include extrapolations that may be implausibly high. If so, expectations of continuing rapid growth in stock prices suggest that the required equity premium may not have declined.

On balance, the continued growth and development of mutual funds and the broader participation in the stock market should contribute to a drop in future equity premiums relative to the historical premium, but the drop is limited.[26] Other factors, such as investors' time horizons and understanding, have less clear-cut implications for the equity premium.

Equity Premium and Current Market Values

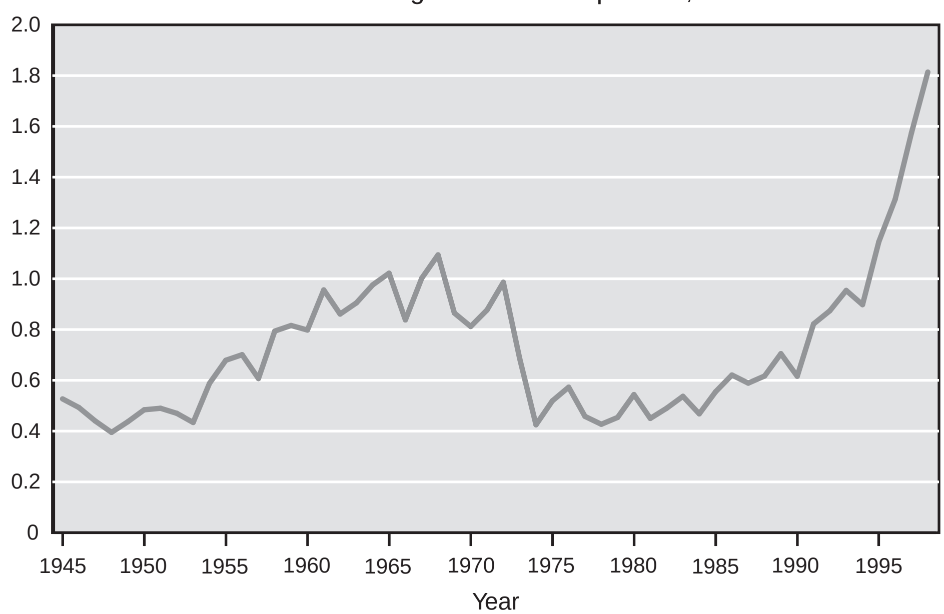
At present, stock prices are very high relative to a number of different indicators, such as earnings, dividends, book values, and gross domestic product (GDP) (Charts 1 and 2). Some

Chart 1.
Price-dividend ratio and price-earnings ratio, 1871-1998



Source: Robert Shiller, Yale University. Available at www.econ.yale.edu/~shiller/data/chapt26.html.
Note: These ratios are based on Standard and Poor's Composite Stock Price Index.

Chart 2.
Ratio of market value of stocks to gross domestic product,1945-1998



Source: Bureau of Economic Analysis data from the national income and product accounts and federal flow of funds.

critics, such as Baker (1998), argue that this high market value, combined with projected slow economic growth, is not consistent with a 7.0 percent return. Possible implications of the high prices have also been the subject of considerable discussion in the finance community (see, for example, Campbell and Shiller 1998; Cochrane 1997; Philips 1999; and Siegel 1999).

The inconsistency of current share prices and 7.0 percent real returns, given OCACT's assumptions for GDP growth, can be illustrated in two ways. The first way is to project the ratio of the stock market's value to GDP, starting with today's values and given assumptions about the future. The second way is to ask what must be true if today's values represent a steady state in the ratio of stock values to GDP.

The first calculation requires assumptions for stock returns, adjusted dividends (dividends plus net share repurchases),[27] and GDP growth. For stock returns, the 7.0 percent assumption is used. For GDP growth rates, OCACT's projections are used. For adjusted dividends, one approach is to assume that the ratio of the aggregate adjusted dividend to GDP would remain the same as the current level. However, as discussed in the accompanying box, the current ratio seems too low to use for projection purposes. Even adopting a higher, more plausible level of adjusted dividends, such as 2.5 percent or 3.0 percent, leads to an implausible rise in the ratio of stock value to GDP—in this case, a more than 20-fold increase over the next 75 years. The calculation derives each year's capital gains by subtracting projected adjusted dividends from the total cash flow to shareholders needed to return 7.0 percent on that year's share values. (See Appendix A for an alternative method of calculating this ratio using a continuous-time differential equation.)

A second way to consider the link between stock market value, stock returns, and GDP is to look at a steady-state relationship. The Gordon formula says that stock returns equal the ratio of adjusted dividends to prices (or the adjusted dividend yield) plus the growth rate of stock prices.[28] In a steady state,

Projecting Future Adjusted Dividends

This article uses the concept of adjusted dividends to estimate the dividend yield. The adjustment begins by adding the value of net share repurchases to actual dividends, since that also represents a cash flow to stockholders in aggregate. A further adjustment is then made to reflect the extent to which the current situation might not be typical of the relationship between dividends and gross domestic product (GDP) in the future. Three pieces of evidence suggest that the current ratio of dividends to GDP is abnormally low and therefore not appropriate to use for projection purposes.

First, dividends are currently very low relative to corporate earnings—roughly 40 percent of earnings compared with a historical average of 60 percent. Because dividends tend to be much more stable over time than earnings, the dividend-earnings ratio declines in a period of high growth of corporate earnings. If future earnings grow at the same rate as GDP, dividends will probably grow faster than GDP to move toward the historical ratio.[1] On the other hand, earnings, which are high relative to GDP, might grow more slowly than GDP. But then, corporate earnings, which have a sizable international component, might grow faster than GDP.

Second, corporations are repurchasing their outstanding shares at a high rate. Liang and Sharpe (1999) report on share repurchases by the 144 largest (nonbank) firms in the Standard and Poor's 500. From 1994 to 1998, approximately 2 percent of share value was repurchased, although Liang and Sharpe anticipate a lower value in the future. At the same time, those firms were issuing shares because employees were exercising stock options at prices below the share values, thus offsetting much of the increase in the number of shares outstanding. Such transfers of net wealth to employees presumably reflect past services. In addition, initial public offerings (IPOs) represent a negative cash flow from stockholders as a whole. Not only the amount paid for stocks but also the value of the shares held by insiders represents a dilution relative to a base for long-run returns on all stocks. As a result, some value needs to be added to the current dividend ratio to adjust for net share repurchases, but the exact amount is unclear. However, in part, the high rate of share repurchase may be just another reflection of the low level of dividends, making it inappropriate to both project much higher dividends in the near term and assume that all of the higher share repurchases will continue. Exactly how to project current numbers into the next decade is not clear.

Finally, projected slow GDP growth, which will plausibly lower investment levels, could be a reason for lower retained earnings in the future. A stable level of earnings relative to GDP and lower retained earnings would increase the ratio of adjusted dividends to GDP.[2]

In summary, the evidence suggests using an "adjusted" dividend yield that is larger than the current level. Therefore, the illustrative calculations in this article use adjusted dividend yields of 2.0 percent, 2.5 percent, 3.0 percent, and 3.5 percent. (The current level of dividends without adjustment for share repurchases is between 1.0 percent and 2.0 percent.)

---

[1] For example, Baker and Weisbrot (1999) appear to make no adjustment for share repurchases or for current dividends being low. However, they use a dividend payout of 2.0 percent, while Dudley and others (1999) report a current dividend yield on the Wilshire 5000 of 1.3 percent.

[2] Firms might change their overall financing package by changing the fraction of net earnings they retain. The implications of such a change would depend on why they were making it. A long-run decrease in retained earnings might merely be increases in dividends and borrowing, with investment held constant. That case, to a first approximation, is another application of the Modigliani-Miller theorem, and the total stock value would be expected to fall by the decrease in retained earnings. Alternatively, a change in retained earnings might signal a change in investment. Again, there is ambiguity. Firms might be retaining a smaller fraction of earnings because investment opportunities were less attractive or because investment had become more productive. These issues tie together two parts of the analysis in this article. If slower growth is associated with lower investment that leaves the return on capital relatively unchanged, then what financial behavior of corporations is required for consistency? Baker (1999b) makes such a calculation; it is not examined here.

the growth rate of prices can be assumed to equal that of GDP. Assuming an adjusted dividend yield of roughly 2.5 percent to 3.0 percent and projected GDP growth of 1.5 percent, the Gordon equation implies a stock return of roughly 4.0 percent to 4.5 percent, not 7.0 percent. Those lower values would imply an equity premium of 1.0 percent to 1.5 percent, given OCACT's assumption of a 3.0 percent yield on Treasury bonds. Making the equation work with a 7.0 percent stock return, assuming no change in projected GDP growth, would require an adjusted dividend yield of roughly 5.5 percent—about double today's level.[29]

For such a large jump in the dividend yield to occur, one of two things would have to happen—adjusted dividends would have to grow much more rapidly than the economy, or stock prices would have to grow much less rapidly than the economy (or even decline). But a consistent projection would take a very large jump in adjusted dividends, assuming that stock prices grew along with GDP starting at today's value. Estimates of recent values of the adjusted dividend yield range from 2.10 percent to 2.55 percent (Dudley and others 1999; Wadhwani 1998).[30]

Even with reasons for additional growth in the dividend yield, which are discussed in the box on projecting future dividends, an implausible growth of adjusted dividends is needed if the short- and long-term returns on stocks are to be 7.0 percent. Moreover, historically, very low values of the dividend yield and earnings-price ratio have been followed primarily by adjustments in stock prices, not in dividends and earnings (Campbell and Shiller 1998).

If the ratio of aggregate adjusted dividends to GDP is unlikely to change substantially, there are three ways out of the internal inconsistency between the market's current value and OCACT's assumptions for economic growth and stock returns. One can:

- Assume higher GDP growth, which would decrease the implausibility of the calculations described above for either the ratio of market value to GDP or the steady state under the Gordon equation. (The possibility of more rapid GDP growth is not explored further in this article.[31])
- Adopt a long-run stock return that is considerably less than 7.0 percent.
- Lower the rate of return during an intermediate period so that a 7.0 percent return could be applied to a lower market value base thereafter.

A combination of the latter two alternatives is also possible.

In considering the prospect of a near-term market decline, the Gordon equation can be used to compute the magnitude of the drop required over, for example, the next 10 years in order for stock returns to average 7.0 percent over the remaining 65 years of OCACT's projection period (see Appendix B). A long-run return of 7.0 percent would require a drop in real prices of between 21 percent and 55 percent, depending on the assumed value of adjusted dividends (Table 3).[32] That calculation is relatively sensitive to the assumed rate of return—for example, with a long-run return of 6.5 percent, the required drop in the market falls to a range of 13 percent to 51 percent.[33]

The two different ways of restoring consistency—a lower stock return in all years or a near-term decline followed by a return to the historical yield—have different implications for Social Security finances. To illustrate the difference, consider the contrast between a scenario with a steady yield of 4.25 percent derived by using current values for the Gordon equation as described above (the steady-state scenario) and a scenario in which stock prices drop by half immediately and the yield on stocks is 7.0 percent thereafter (the market-correction scenario).[34] First, dollars newly invested in the future (that is, after any drop in share prices) earn only 4.25 percent per year under the steady-state scenario, compared with 7.0 percent per year under the market-correction scenario. Second, even for dollars currently in the market, the long-run yield differs under the two scenarios when the returns on stocks are being reinvested. Under the steady-state scenario, the yield on dollars currently in the market is 4.25 percent per year over any projected time period; under the market-correction scenario, the annual rate of return depends on the time horizon used for the calculation.[35] After one year, the latter scenario has a rate of return of –46 percent. By the end of 10 years, the annual rate of return with the latter scenario is –0.2 percent; by the end of 35 years, 4.9 percent; and by the end of 75 years, 6.0 percent. Proposals for Social Security generally envision a gradual buildup of stock investments, which suggests that those investments would fare better under the market-correction scenario. The importance of the difference between scenarios depends also on the choice of additional changes to Social Security, which affect how long the money can stay invested until it is needed to pay benefits.

Given the different impacts of these scenarios, which one is more likely to occur? The key issue is whether the current stock

Table 3.
Required percentage decline in real stock prices over the next 10 years to justify a return of 7.0, 6.5, and 6.0 percent thereafter

Adjusted dividend yield	Percentage decline to justify a long-run return of—		
	7.0	6.5	6.0
2.0	55	51	45
2.5	44	38	31
3.0	33	26	18
3.5	21	13	4

Source: Author's calculations.

Note: Derived from the Gordon formula. Dividends are assumed to grow in line with gross domestic product (GDP), which the Office of the Chief Actuary (OCACT) assumes is 2.0 percent over the next 10 years. For long-run GDP growth, OCACT assumes 1.5 percent.

market is overvalued in the sense that rates of return are likely to be lower in the intermediate term than in the long run. Economists have divergent views on this issue.

One possible conclusion is that current stock prices signal a significant drop in the long-run required equity premium. For example, Glassman and Hassett (1999) have argued that the equity premium will be dramatically lower in the future than it has been in the past, so that the current market is not overvalued in the sense of signaling lower returns in the near term than in the long run.[36] Indeed, they even raise the possibility that the market is "undervalued" in the sense that the rate of return in the intermediate period will be higher than in the long run, reflecting a possible continuing decline in the required equity premium. If their view is right, then a 7.0 percent long-run return, together with a 4.0 percent equity premium, would be too high.

Others argue that the current stock market values include a significant price component that will disappear at some point, although no one can predict when or whether it will happen abruptly or slowly. Indeed, Campbell and Shiller (1998) and Cochrane (1997) have shown that when stock prices (normalized by earnings, dividends, or book values) have been far above historical ratios, the rate of return over the following decade has tended to be low, and the low return is associated primarily with the price of stocks, not the growth of dividends or earnings.[37] Thus, to project a steady rate of return in the future, one needs to argue that this historical pattern will not repeat itself. The values in Table 3 are in the range suggested by the historical relationship between future stock prices and current price-earnings and price-dividend ratios (see, for example, Campbell and Shiller 1998).

Therefore, either the stock market is overvalued and requires a correction to justify a 7.0 percent return thereafter, or it is correctly valued and the long-run return is substantially lower than 7.0 percent. (Some combination of the two is also possible.) Under either scenario, stock returns would be lower than 7.0 percent for at least a portion of the next 75 years. Some evidence suggests, however, that investors have not adequately considered that possibility.[38] The former view is more convincing, since accepting the "correctly valued" hypothesis implies an implausibly small long-run equity premium. Moreover, when stock values (compared with earnings or dividends) have been far above historical ratios, returns over the following decade have tended to be low. Since this discussion has no direct bearing on bond returns, assuming a lower return for stocks over the near or long term also means assuming a lower equity premium.

In short, given current stock values, a constant 7.0 percent return is not consistent with OCACT's projected GDP growth.[39] However, OCACT could assume lower returns for a decade, followed by a return equal to or about 7.0 percent.[40] In that case, OCACT could treat equity returns as it does Treasury rates, using different projection methods for the first 10 years and for the following 65. This conclusion is not meant to suggest that anyone is capable of predicting the timing of annual stock returns, but rather that this is an approach to financially consistent assumptions. Alternatively, OCACT could adopt a lower rate of return for the entire 75-year period.

Marginal Product of Capital and Slow Growth

In its long-term projections, OCACT assumes a slower rate of economic growth than the U.S. economy has experienced over an extended period. That projection reflects both the slowdown in labor force growth expected over the next few decades and the slowdown in productivity growth since 1973.[41] Some critics have suggested that slower growth implies lower projected rates of return on both stocks and bonds, since the returns to financial assets must reflect the returns on capital investment over the long run. That issue can be addressed by considering either the return to stocks directly, as discussed above, or the marginal product of capital in the context of a model of economic growth.[42]

For the long run, the returns to financial assets must reflect the returns on the physical assets that support the financial assets. Thus, the question is whether projecting slower economic growth is a reason to expect a lower marginal product of capital. As noted above, this argument speaks to rates of return generally, not necessarily to the equity premium.

The standard (Solow) model of economic growth implies that slower long-run economic growth with a constant savings rate will yield a lower marginal product of capital, and the relationship may be roughly point-for-point (see Appendix C). However, the evidence suggests that savings rates are not unaffected by growth rates. Indeed, growth may be more important for savings rates than savings are for growth rates. Bosworth and Burtless (1998) have observed that savings rates and long-term rates of income growth have a persistent positive association, both across countries and over time. That observation suggests that if future economic growth is slower than in the past, savings will also be lower. In the Solow model, low savings raise the marginal product of capital, with each percentage-point decrease in the savings rate increasing the marginal product by roughly one-half of a percentage point in the long run. Since growth has fluctuated in the past, the stability in real rates of return to stocks, as shown in Table 1, suggests an offsetting savings effect, preserving the stability in the rate of return.[43]

Focusing directly on demographic structure and the rate of return rather than on labor force growth and savings rates, Poterba (1998) does not find a robust relationship between demographic structure and asset returns. He does recognize the limited power of statistical tests based on the few "effective degrees of freedom" in the historical record. Poterba suggests that the connection between demography and returns is not simple and direct, although such a connection has been raised as a possible reason for high current stock values, as baby boomers save for retirement, and for projecting low future stock values, as they finance retirement consumption. Goyal (1999) estimates equity premium regressions and finds that changes in population age structure add significant explanatory power. Nevertheless, using a vector autoregression approach, his analysis predicts no significant increase in *average* outflows over the next 52 years. That occurs despite the retirement of baby boomers. Thus, both papers reach the same conclusion—that demography is not likely to effect large changes in the long-run rate of return.

Another factor to consider in assessing the connection between growth and rates of return is the increasing openness of the world economy. Currently, U.S. corporations earn income from production and trade abroad, and individual investors, while primarily investing at home, also invest abroad. It is not clear that putting the growth issue in a global context makes much difference. On the one hand, since other advanced economies are also aging, increased economic connections with other advanced countries do not alter the basic analysis. On the other hand, although investment in the less-developed countries may preserve higher rates, it is not clear either how much investment opportunities will increase or how to adjust for political risk. Increasing openness further weakens the argument for a significant drop in the marginal product of capital, but the opportunities abroad may or may not be realized as a better rate of return.

On balance, slower projected growth may reduce the return on capital, but the effect is probably considerably less than one-for-one. Moreover, this argument relates to the overall return to capital in an economy, not just stock returns. Any impact would therefore tend to affect returns on both stocks and bonds similarly, with no directly implied change in the equity premium.[44]

Other Issues

This paper has considered the gross rate of return to equities and the equity premium generally. Two additional issues arise in considering the prospect of equity investment for Social Security: how gross returns depend on investment strategy and how they differ from net returns; and the degree of risk associated with adding stock investments to a current all-bond portfolio.

Gross and Net Returns

A gross rate of return differs from a net return because it includes transactions costs such as brokerage charges, bid-ask spreads, and fees for asset management.[45]

If the Social Security trust fund invests directly in equities, the investment is likely to be in an index fund representing almost all of the equities outstanding in the United States. Thus, the analysis above holds for that type of investment. Although some critics have expressed concern that political influence might cause deviations from a broad-based indexing strategy, the evidence suggests that such considerations would have little impact on the expected rate of return (Munnell and Sundén 1999).

If the investment in stocks is made through individual accounts, then individuals may be given some choice either about the makeup of stock investment or about varying the mix of stocks and bonds over time. In order to consider the rate of return on stocks held in such individual accounts, one must consider the kind of portfolio choices individuals might make, both in the composition of the stock portfolio and in the timing of purchases and sales. Given the opportunity, many individuals would engage in numerous transactions, both among stocks and between stocks and other assets (attempts to time the market).

The evidence suggests that such transactions reduce gross returns relative to risks, even before factoring in transactions costs (Odean 1998). Therefore, both the presence of individual accounts with choice and the details of their regulation are likely to affect gross returns. On average, individual accounts with choice are likely to have lower gross returns from stocks than would direct trust fund investment.

Similarly, the cost of administration as a percentage of managed assets varies depending on whether there are individual accounts and how they are organized and regulated (National Academy of Social Insurance 1998; Diamond 2000). Estimates of that cost vary from 0.5 basis points for direct trust fund investment to 100 to150 basis points for individually organized individual accounts, with government-organized individual accounts somewhere in between.

Investment Risk of Stocks

The Office of the Chief Actuary's projections are projections of plausible long-run scenarios (ignoring fluctuations). As such, they are useful for identifying a sizable probability of future financial needs for Social Security. However, they do not address different probabilities for the trust fund's financial condition under different policies.[46] Nor are they sufficient for normative evaluation of policies that have different distributional or risk characteristics.

Although investment in stocks entails riskiness in the rate of return, investment in Treasury bonds also entails risk. Therefore, a comparison of those risks should consider the distribution of outcomes—concern about risk should not be separated from the compensation for bearing risk. That is, one needs to consider the probabilities of both doing better and doing worse as a result of holding some stocks. Merely observing that stocks are risky is an inadequate basis for policy evaluations. Indeed, studies of the historical pattern of returns show that portfolio risk decreases when some stocks are added to a portfolio consisting only of nominal bonds (Siegel 1998). Furthermore, many risks affect the financial future of Social Security, and investing a small portion of the trust fund in stocks is a small risk for the system as a whole relative to economic and demographic risks (Thompson 1998).

As long as the differences in risk and expected return are being determined in a market and reflect the risk aversion of market participants, the suitability of the trust fund's portfolio can be considered in terms of whether Social Security has more or less risk aversion than current investors. Of course, the "risk aversion" of Social Security is a derived concept, based on the risks to be borne by future beneficiaries and taxpayers, who will incur some risk whatever portfolio Social Security holds. Thus, the question is whether the balance of risks and returns looks better with one portfolio than with another. The answer is somewhat complex, since it depends on how policy changes in taxes and benefits respond to economic and demographic outcomes. Nevertheless, since individuals are normally advised to hold at least some stocks in their own portfolios, it seems appropriate for Social Security to also hold some stocks when investing on their behalf, at least in the long run, regardless of the rates of return used for projection purposes (Diamond and Geanakoplos 1999).[47]

Conclusion

Of the three main bases for criticizing OCACT's assumptions, by far the most important one is the argument that a constant 7.0 percent stock return is not consistent with the value of today's stock market and projected slow economic growth. The other two arguments—pertaining to developments in financial markets and the marginal product of capital—have merit, but neither suggests a dramatic change in the equity premium.

Given the high value of today's stock market and an expectation of slower economic growth in the future, OCACT could adjust its stock return projections in one of two ways. It could assume a decline in the stock market sometime over the next decade, followed by a 7.0 percent return for the remainder of the projection period. That approach would treat equity returns like Treasury rates, using different short- and long-run projection methods for the first 10 years and the following 65 years. Alternatively, OCACT could adopt a lower rate of return for the entire 75-year period. That approach may be more acceptable politically, but it obscures the expected pattern of returns and may produce misleading assessments of alternative financing proposals, since the appropriate uniform rate to use for projection purposes depends on the investment policy being evaluated.

Notes

Peter Diamond is Institute Professor at the Massachusetts Institute of Technology, where he has taught since 1966. He is a member of the Board of the National Academy of Social Insurance, where he has been President, Chair of the Board, and Chair of the Panel on Privatization of Social Security. He has written on public finance, macroeconomics, and the economics of uncertainty.

Acknowledgments: The author is grateful to John Campbell, Alicia Munnell, and Jim Poterba for extended discussions and to Andy Abel, Dean Baker, Olivier Blanchard, John Cochrane, Andy Eschtruth, Steve Goss, Joyce Manchester, Peter Orszag, Bernie Saffran, Jeremy Siegel, Tim Smeeding, Peter Temin, and Joe White for helpful comments. The views and remaining errors are those of the author.

[1] This 7.0 percent real rate of return is gross of administrative charges.

[2] To generate short-run returns on stocks, the Social Security Administration's Office of the Chief Actuary (OCACT) multiplied the ratio of one plus the ultimate yield on stocks to one plus the ultimate yield on bonds by the annual bond assumptions in the short run.

[3] An exception was the use of 6.75 percent for the President's proposal evaluated in a memorandum on January 26, 1999.

[4] This report is formally called the *1999 Annual Report of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.*

[5] For OCACT's short-run bond projections, see Table II.D.1 in the 1999 Social Security Trustees Report.

[6] This article was written in the summer of 1999 and uses numbers appropriate at the time. The 2000 Trustees Report uses the same assumptions of 6.3 percent for the nominal interest rate and 3.3 percent for the annual percentage change in the consumer price index. The real wage is assumed to grow at 1.0 percent, as opposed to 0.9 percent in the 1999 report.

[7] See, for example, Baker (1999a) and Baker and Weisbrot (1999). This article only considers return assumptions given economic growth assumptions and does not consider growth assumptions.

[8] This article does not analyze the policy issues related to stock market investment either by the trust fund or through individual accounts. Such an analysis needs to recognize that higher expected returns in the U.S. capital market come with higher risk. For the issues relevant for such a policy analysis, see National Academy of Social Insurance (1998).

[9] Ideally, one would want the yield on the special Treasury bonds held by Social Security. However, this article simply refers to published long-run bond rates.

[10] Because annual rates of return on stocks fluctuate so much, a wide band of uncertainty surrounds the best statistical estimate of the average rate of return. For example, Cochrane (1997) notes that over the 50 years from 1947 to 1996, the excess return of stocks over Treasury bills was 8 percent, but, assuming that annual returns are statistically independent, the standard statistical confidence interval extends from 3 percent to 13 percent. Using a data set covering a longer period lowers the size of the confidence interval, provided one is willing to assume that the stochastic process describing rates of return is stable for the longer period. This article is not concerned with that uncertainty, only with the appropriate rate of return to use for a central (or intermediate) projection. For policy purposes, one must also look at stochastic projections (see, for example, Copeland, VanDerhei, and Salisbury 1999; and Lee and Tuljapurkar 1998). Despite the value of stochastic projections, OCACT's central projection plays an important role in thinking about policy and in the political process. Nevertheless, when making a long-run projection, one must realize that great uncertainty surrounds any single projection and the relevance of returns in any short period of time.

[11] Table 2 also shows the equity premiums relative to Treasury bills. Those numbers are included only because they arise in other discussions; they are not referred to in this article.

[12] For determining the equity premium shown in Table 2, the rate of return is calculated assuming that a dollar is invested at the start of a period and the returns are reinvested until the end of the period. In contrast to that geometric average, an arithmetic average is the average of the annual rates of return for each of the years in a period. The arithmetic average is larger than the geometric average. Assume, for example, that a dollar doubles in value in year 1 and then halves in value from year 1 to year 2. The geometric average over the 2-year period is zero; the arithmetic average of +100 percent and –50 percent annual rates of return is +25 percent. For projection purposes, one looks for an estimated rate of return that is suitable for investment over a long period. Presumably the best approach would be to take the arithmetic average of the rates of return that were each the geometric average for different historical periods of the same length as

the average investment period within the projection period. That calculation would be close to the geometric average, since the variation in 35- or 40-year geometric rates of return, which is the source of the difference between arithmetic and geometric averages, would not be so large.

[13] In considering recent data, some adjustment should be made for bond rates being artificially low in the 1940s as a consequence of war and postwar policies.

[14] Also relevant is the fact that the real rate on 30-year Treasury bonds is currently above 3.0 percent.

[15] Finance theory relates the willingness to hold alternative assets to the expected risks and returns (in real terms) of the different assets, recognizing that expectations about risk and return are likely to vary with the time horizon of the investor. Indeed, time horizon is an oversimplification, since people are also uncertain about when they will want to have access to the proceeds of those investments. Thus, finance theory is primarily about the difference in returns to different assets (the equity premium) and needs to be supplemented by other analyses to consider the expected return to stocks.

[16] With Treasury bonds, investors can easily project future nominal returns (since default risk is taken to be virtually zero), although expected real returns depend on projected inflation outcomes given nominal yields. With inflation-protected Treasury bonds, investors can purchase bonds with a known real interest rate. Since those bonds were introduced only recently, they do not play a role in interpreting the historical record for projection purposes. Moreover, their importance in future portfolio choices is unclear.

[17] In theory, for determining asset prices at which markets clear, one wants to consider marginal investments. Those investments are made up of a mix of marginal portfolio allocations by all investors and by marginal investors who become participants (or nonparticipants) in the stock and/or bond markets.

[18] This conclusion does not contradict the Modigliani-Miller theorem. Different firms with the same total return distributions but different amounts of debt outstanding will have the same total value (stock plus bond) and so the same total expected return. A firm with more debt outstanding will have a higher expected return on its stock in order to preserve the total expected return.

[19] Consideration of equilibrium suggests an alternative approach to analyzing the historical record. Rather than looking at realized rates of return, one could construct estimates of expected rates of return and see how they have varied in the past. That approach has been taken by Blanchard (1993). He concluded that the equity premium (measured by expectations) was unusually high in the late 1930s and 1940s and, since the 1950s, has experienced a long decline from that unusually high level. The high realized rates of return over this period are, in part, a consequence of a decline in the equity premium needed for people to be willing to hold stocks. In addition, the real expected returns on bonds have risen since the 1950s, which should have moderated the impact of a declining equity premium on expected stock returns. Blanchard examines the importance of inflation expectations and attributes some of the recent trend to a decline in expected inflation. He concluded that the premium in 1993 appeared to be around 2 percent to 3 percent and would probably not move much if inflation expectations remain low. He also concluded that decreases in the equity premium were likely to involve both increases in expected bond rates and decreases in expected rates of return on stocks.

[20] If current cash returns to stockholders are expected to grow at rate g, with projected returns discounted at rate r, this fundamental value is the current return divided by $(r - g)$. If r is smaller, fluctuations in long-run projections of g result in larger fluctuations in the fundamental value.

[21] Several explanations have been put forth, including: (1) the United States has been lucky, compared with stock investment in other countries, and realized returns include a premium for the possibility that the U.S. experience might have been different; (2) returns to actual investors are considerably less than the returns on indexes that have been used in analyses; and (3) individual preferences are different from the simple models that have been used in examining the puzzle.

[22] The timing of realized returns that are higher than required returns is somewhat more complicated, since recognizing and projecting such a trend will tend to boost the price of equities when the trend is recognized, not when it is realized.

[23] Nonprofit institutions, such as universities, and defined benefit plans for public employees now hold more stock than in the past. Attributing the risk associated with that portfolio to the beneficiaries of those institutions would further expand the pool sharing in the risk.

[24] More generally, the equity premium depends on the investment strategies being followed by investors.

[25] This tendency, known as mean reversion, implies that a short period of above-average stock returns is likely to be followed by a period of below-average returns.

[26] To quantify the importance of these developments, one would want to model corporate behavior as well as investor behavior. A decline in the equity premium reflects a drop to corporations in the "cost of risk" in the process of acquiring funds for risky investment. If the "price per unit of risk" goes down, corporations might respond by selecting riskier investments (those with a higher expected return), thereby somewhat restoring the equity premium associated with investing in corporations.

[27] In considering the return to an individual from investing in stocks, the return is made up of dividends and a (possible) capital gain from a rise in the value of the shares purchased. When considering the return to all investment in stocks, one needs to consider the entire cash flow to stockholders, including dividends and net share repurchases by the firms. That suggests two methods of examining the consistency of any assumed rate of return on stocks. One is to consider the value of all stocks outstanding. If one assumes that the value of all stocks outstanding grows at the same rate as the economy (in the long run), then the return to all stocks outstanding is that rate of growth plus the sum of dividends and net share repurchases, relative to total share value. Alternatively, one can consider ownership of a single share. The assumed rate of return minus the rate of dividend payment then implies a rate of capital gain on the single share. However, the relationship between the growth of value of a single share and the growth of the economy depends on the rate of share repurchase. As shares are being repurchased, remaining shares should grow in value relative to the growth of the economy. Either approach can be calculated in a consistent manner. What must be avoided is an inconsistent mix, considering only dividends and also assuming that the value of a single share grows at the same rate as the economy.

[28] Gordon (1962). For an exposition, see Campbell, Lo, and MacKinlay (1997).

[29] The implausibility refers to total stock values, not the value of single shares—thus, the relevance of net share repurchases. For example, Dudley and others (1999) view a steady equity premium in the range of 1.0 percent to 3.0 percent as consistent with current stock prices and their projections. They assume 3.0 percent GDP growth and a 3.5 percent real bond return, both higher than the assumptions used by OCACT. Wadhwani (1998) finds that if the S&P 500 is correctly valued, he has to assume a negative risk premium. He considers various adjustments that lead to a higher premium, with his "best guess" estimate being 1.6 percent. That still seems too low.

[30] Dudley and others (1999) report a current dividend yield on the Wilshire 5000 of 1.3 percent. They then make an adjustment that is equivalent to adding 80 basis points to that rate for share repurchases, for which they cite Campbell and Shiller (1998). Wadhwani (1998) finds a current expected dividend yield of 1.65 percent for the S&P 500, which he adjusts to 2.55 percent to account for share repurchases. For a discussion of share repurchases, see Cole, Helwege, and Laster (1996).

[31] Stock prices reflect investors' assumptions about economic growth. If their assumptions differ from those used by OCACT, then it becomes difficult to have a consistent projection that does not assume that investors will be surprised.

[32] In considering these values, note the observation that a fall of 20 percent to 30 percent in advance of recessions is typical for the U.S. stock market (Wadhwani 1998). With OCACT assuming a 27 percent rise in the price level over the next decade, a 21 percent decline in real stock prices would yield the same nominal prices as at present.

[33] The importance of the assumed growth rate of GDP can be seen by redoing the calculations in Table 3 for a growth rate that is one-half of a percent larger in both the short and long runs. Compared with the original calculations, such a change would increase the ratios by 16 percent.

[34] Both scenarios are consistent with the Gordon formula, assuming a 2.75 percent adjusted dividend yield (without a drop in share prices) and a growth of dividends of 1.5 percent per year.

[35] With the steady-state scenario, a dollar in the market at the start of the steady state is worth 1.0425^t dollars t years later, if the returns are continuously reinvested. In contrast, under the market-correction scenario, a dollar in the market at the time of the drop in prices is worth $(1/2)(1.07^t)$ dollars t years later.

[36] The authors appear to assume that the Treasury rate will not change significantly, so that changes in the equity premium and in the return to stocks are similar.

[37] One could use equations estimated on historical prices to check the plausibility of intermediate-run stock values with the intermediate-run values needed for plausibility for the long-run assumptions. Such a calculation is not considered in this article. Another approach is to consider the value of stocks relative to the replacement cost of the capital that corporations hold, referred to as Tobin's q. That ratio has fluctuated considerably and is currently unusually high. Robertson and Wright (1998) have analyzed the ratio and concluded that a cumulative real decline in the stock market over the first decades of the 21st century has a high probability.

[38] As Wadhwani (1998, p. 36) notes, "Surveys of individual investors in the United States regularly suggest that they expect returns above 20 percent, which is obviously unsustainable. For example, in a survey conducted by Montgomery Asset Management in 1997, the typical mutual fund investor expected annual returns from the stock market of 34 percent over the next 10 years! Most U.S. pension funds operate under actuarial assumptions of equity returns in the 8-10 percent area, which, with a dividend yield under 2 percent and nominal GNP growth unlikely to exceed 5 percent, is again, unsustainably high."

[39] There is no necessary connection between the rate of return on stocks and the rate of growth of the economy. There is a connection among the rate of return on stocks, the current stock prices, dividends relative to GDP, and the rate of growth of the economy.

[40] The impact of such a change in assumptions on actuarial balance depends on the amount that is invested in stocks in the short term relative to the amount invested in the long term. The levels of holdings at different times depend on both the speed of initial investment and whether stock holdings are sold before very long (as would happen with no other policy changes) or whether, instead, additional policies are adopted that result in a longer holding period, possibly including a sustained sizable portfolio of stocks. Such an outcome would follow if Social Security switched to a sustained level of funding in excess of the historical long-run target of just a contingency reserve equal to a single year's expenditures.

[41] "The annual rate of growth in total labor force decreased from an average of about 2.0 percent per year during the 1970s and 1980s to about 1.1 percent from 1990 to 1998. After 1998 the labor force is projected to increase about 0.9 percent per year, on average, through 2008, and to increase much more slowly after that, ultimately reaching 0.1 percent toward the end of the 75-year projection period" (Social Security Trustees Report, p. 55). "The Trustees assume an intermediate trend growth rate of labor productivity of 1.3 percent per year, roughly in line with the average rate of growth of productivity over the last 30 years" (Social Security Trustees Report, p. 55).

[42] Two approaches are available to answer this question. Since the Gordon formula, given above, shows that the return to stocks equals the adjusted dividend yield plus the growth of stock prices, one needs to consider how the dividend yield is affected by slower growth. In turn, that relationship will depend on investment levels relative to corporate earnings. Baker (1999b) makes such a calculation, which is not examined here. Another approach is to consider the return on physical capital directly, which is the one examined in this article.

[43] Using the Granger test of causation (Granger 1969), Carroll and Weil (1994) find that growth causes saving but saving does not cause growth. That is, changes in growth rates tend to precede changes in savings rates but not vice versa. For a recent discussion of savings and growth, see Carroll, Overland, and Weil (2000).

[44] One can also ask how a change in policy designed to build and maintain a larger trust fund in a way that significantly increases national saving might affect future returns. Such a change would plausibly tend to lower rates of return. The size of that effect depends on the size of investment increases relative to available investment opportunities, both in the United States and worldwide. Moreover, it depends on the response of private saving to the policy, including the effect that would come through any change in the rate of return. There is plausibly an effect here, although this article does not explore it. Again, the argument speaks to the level of rates of return generally, not to the equity premium.

[45] One can also ask how changed policies might affect future returns. A change in portfolio policy that included stocks (whether in the trust fund or in individual accounts) would plausibly lower the equity premium somewhat. That effect could come about through a combination of a rise in the Treasury rate (thereby requiring a change

in tax and/or expenditure policy) and a fall in expected returns on stocks. The latter depends on both the underlying technology of available returns to real investments and the effect of portfolio policy on national saving. At this time, research on this issue has been limited, although it is plausible that the effect is not large (Bohn 1998; Abel 1999; Diamond and Geanakoplos 1999).

[46] For stochastic projections, see Copeland, VanDerhei, and Salisbury (1999); and Lee and Tuljapurkar (1998). OCACT generally provides sensitivity analysis by doing projections with several different rates of return on stocks.

[47] Cochrane (1997, p. 32) reaches a similar conclusion relative to individual investment: "We could interpret the recent run-up in the market as the result of people finally figuring out how good an investment stocks have been for the last century, and building institutions that allow wise participation in the stock market. If so, future returns are likely to be much lower, but there is not much one can do about it but sigh and join the parade."

References

Abel, Andrew B. 1999. *The Social Security Trust Fund, The Riskless Interest Rate, and Capital Accumulation.* NBER Working Paper 6991. Cambridge, Mass.: National Bureau of Economic Research.

Advisory Council on Social Security. 1997. *Report of the 1994-1996 Advisory Council on Social Security*. Vol. 1, *Findings and Recommendations*. Washington, D.C.

Baker, Dean. 1998. "Saving Social Security in Three Steps." Briefing Paper, Economic Policy Institute, Washington, D.C.

_______. 1999a. "Two Cheers for Clinton's Social Security Plan." *The American Prospect,* Vol. 10, No. 44, pp. 82-85.

_______.1999b. Letter to Martin Feldstein, May 15. Available at www.preamble.org.

Baker, Dean, and Mark Weisbrot. 1999. *Social Security: The Phony Crisis*. Chicago, Ill.: University of Chicago Press.

Blanchard, Olivier J. 1993. "Movements in the Equity Premium." *Brookings Papers on Economic Activity*, No. 2, pp. 75-138.

Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds. 1999. *1999 Annual Report.* Washington, D.C.: U.S. Government Printing Office.

Bohn, Henning. 1998. "Social Security Reform and Financial Markets." In Federal Reserve Bank of Boston, *Social Security Reform: Links to Saving, Investment, and Growth*. Boston, Mass.: Federal Reserve Bank of Boston.

Bosworth, Barry, and Gary Burtless. 1998. "Social Security Reform in a Global Context." In Federal Reserve Bank of Boston, *Social Security Reform: Links to Saving, Investment, and Growth*. Boston, Mass.: Federal Reserve Bank of Boston.

Brav, Alon, and Christopher C. Geczy. 1996. "An Empirical Resurrection of the Simple Consumption CAPM with Power Utility." Working Paper, University of Chicago.

Campbell, John Y.; Andrew W. Lo; and A. Craig MacKinlay. 1997. *The Econometrics of Financial Markets*. Princeton, N.J.: Princeton University Press.

Campbell, John Y., and Robert J. Shiller. 1998. "Valuation Ratios and the Long-Run Stock Market Outlook." *Journal of Portfolio Management*, Vol. 24, No. 2, pp. 11-26.

Campbell, John Y., and Luis M. Viceira. 1999. "Consumption and Portfolio Decisions When Expected Returns Are Time Varying." *Quarterly Journal of Economics* (May), pp. 433-495.

Carroll, Christopher D.; Jody Overland; and David N. Weil. 2000. "Saving and Growth with Habit Formation." *American Economic Review*, Vol. 90, No. 3, pp. 341-355.

Carroll, Christopher D., and David N. Weil. 1994. "Saving and Growth: A Reinterpretation." *Carnegie-Rochester Conference Series on Public Policy,* No. 40, pp. 133-192.

Case, Karl E., and Robert J. Shiller. 1988. "The Behavior of Home Buyers in Boom and Post-Boom Markets." *New England Economic Review* (November/December), pp. 29-46.

Cochrane, John H. 1997. "Where is the Market Going? Uncertain Facts and Novel Theories." *Economic Perspectives,* Vol. 21, Federal Reserve Bank of Chicago. (Also published as NBER Working Paper 6207, Cambridge, Mass.: National Bureau of Economic Research.)

Cole, Kevin; Jean Helwege; and David Laster. 1996. "Stock Market Valuation Indicators: Is This Time Different?" *Financial Analysts Journal,* Vol. 52, No. 3, pp. 56-64.

Copeland, Craig; Jack VanDerhei; and Dallas L. Salisbury. 1999. *Social Security Reform: Evaluating Current Proposals: Latest Results of the EBRI SSASIM2 Policy Simulation Model*. EBRI Issue Brief 210. Washington, D.C.: Employee Benefit Research Institute.

Diamond, Peter. 1999. *Issues in Privatizing Social Security*. Cambridge, Mass.: MIT Press.

_______. 2000. "Administrative Costs and Equilibrium Charges with Individual Accounts." In *Administrative Aspects of Investment-Based Social Security Reform*, edited by John B. Shoven. Chicago and London: University of Chicago Press, pp. 137-162.

Diamond, Peter, and John Geanakoplos. 1999. *Social Security Investment in Equities I: Linear Case*. Working Paper 2. Center for Retirement Research at Boston College, Chestnut Hill, Mass.

Dudley, William; Jan Hatzius; Joanne Hill; Rajib Pal; and Maria Tsu. 1999. "An Analysis of Social Security Trust Fund Investment into Equities." Goldman Sachs Global Economics Paper 18, New York.

Glassman, James K., and Kevin A. Hassett. 1999. *Dow 36,000*. New York: Times Business.

Gordon, Myron. 1962. *The Investment, Financing, and Valuation of the Corporation*. Homewood, Ill.: Irwin.

Gordon, Roger. 1985. "Taxation of Corporate Capital Income: Tax Revenues versus Tax Distributions." *Quarterly Journal of Economics,* Vol. 100, No. 1, pp. 1-27.

Goyal, Amit. 1999. "Demographics, Stock Market Flows, and Stock Returns." Anderson Graduate School of Management, University of California at Los Angeles. Unpublished.

Granger, C.W.J. 1969. "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." *Econometrica*, Vol. 37, No. 3, pp. 424-438.

Heaton, John, and Deborah Lucas. 2000. "Stock Prices and Fundamentals." In *NBER Macroeconomics Annual,* edited by Ben S. Bernanke and Julio Rotemberg. Cambridge, Mass.: MIT Press.

Investment Company Institute. 1999. *Mutual Fund Fact Book* (May), p. 22. Available at www.ici.org.

Kaplow, Louis. 1994. “Taxation and Risk-taking: A General Equilibrium Perspective.” *National Tax Journal,* Vol. 47, pp. 789-798.

Kennickell, Arthur B.; Martha Starr-McCluer; and Annika Sundén. 1997. “Family Finances in the U.S.: Recent Evidence from the Survey of Consumer Finances.” *Federal Reserve Bulletin,* Vol. 83 (June), pp. 1-24.

Lee, Ronald, and Shripad Tuljapurkar. 1998. “Stochastic Forecasts for Social Security.” In *Frontiers in the Economics of Aging*, edited by David Wise. Chicago, Ill.: University of Chicago Press, pp. 393-420.

Liang, J. Nellie, and Steven A. Sharpe. 1999. “Share Repurchases and Employee Stock Options and Their Implications for S&P 500 Share Retirements and Expected Returns.” Division of Research and Statistics, Federal Reserve Board, Washington, D.C. Unpublished.

Mehra, Rajnish, and Edward Prescott. 1985. “The Equity Premium Puzzle.” *Journal of Monetary Economics*, Vol. 15, pp. 145-161.

Mankiw, N.G., and S. Zeldes. 1991. “The Consumption of Stockholders and Non-Stockholders.” *Journal of Financial Economics*, Vol.17, pp. 211-219.

Munnell, Alicia, and Annika Sundén (with the assistance of Cynthia Perry and Ryan Kling). 1999. *Investment Practices of State and Local Pension Funds: Implications for Social Security*. Working Paper 1. Center for Retirement Research at Boston College, Chestnut Hill, Mass.

National Academy of Social Insurance. 1998. *Evaluating Issues in Privatizing Social Security.* Report of the Panel on Privatization of Social Security. Available at www.nasi.org and as Diamond 1999.

Odean, Terrance. 1998. “Do Investors Trade Too Much?” Graduate School of Management,University of California, Davis. Unpublished.

Philips, Thomas K. 1999. “Why Do Valuation Ratios Forecast Long-Run Equity Returns?” *Journal of Portfolio Management*, Vol. 25, No. 3, pp. 39-44.

Poterba, James M. 1998. “Population Age Structure and Asset Returns: An Empirical Investigation.” Department of Economics, Massachusetts Institute of Technology, Cambridge, Mass. Unpublished.

Poterba, James M., and Andrew A. Samwick. 1995. “Stock Ownership Patterns, Stock Market Fluctuations, and Consumption.” *Brookings Papers on Economic Activity*, No. 2, pp. 295-372.

Rea, John D., and Brian K. Reid. 1998. *Trends in the Ownership Cost of Equity Mutual Funds*. Perspective, Vol. 4, No. 3. Washington, D.C.: Investment Company Institute.

_______.1999. *Total Shareholder Cost of Bond and Money Market Mutual Funds*. Perspective, Vol. 5, No. 3. Washington, D.C.: Investment Company Institute.

Robertson, Donald, and Stephen Wright. 1998. “The Good News and the Bad News about Long-Run Stock Market Returns.” Faculty of Economics and Politics, University of Cambridge. Unpublished.

Samuelson, Paul A. 1989. “A Case at Last for Age-Phased Reduction in Equity.” *Proceedings of the National Academy of Sciences*, Vol. 86, pp. 9048-9051.

Shafir, Eldar; Peter Diamond; and Amos Tversky. 1997. “Money Illusion.” *Quarterly Journal of Economics*, Vol. 112, No. 2, pp. 341-374.

Sharpe, Steven. 1999. “Stock Prices, Expected Returns, and Inflation.” Federal Reserve Board, Washington, D.C. Unpublished.

Siegel, Jeremy J. 1998. *Stocks for the Long Run*. New York: McGraw Hill, p. 318.

_______. 1999. “The Shrinking Equity Premium: Historical Facts and Future Forecasts.” Wharton School, University of Pennsylvania, Philadelphia, Pa. Unpublished.

Thompson, Lawrence. 1998. *Older and Wiser: The Economics of Public Pensions*. Washington, D.C.: Urban Institute Press.

Vissing-Jorgensen, Annette. 1997. “Limited Stock Market Participation.” Department of Economics, Massachusetts Institute of Technology, Cambridge, Mass. Unpublished.

Wadhwani, Sushil B. 1998. “The U.S. Stock Market and the Global Economic Crisis.” Tudor Investment Corporation, New York. Unpublished.

Appendix A: Alternative Method for Determining the Ratio of Stock Value to GDP

Variables

r …… rate of return on stocks

g …… rate of growth of both GDP and dividends

a …… adjusted dividend yield at time 0

$P(t)$ … aggregate stock value at time t

$Y(t)$ … GDP at time t

$D(t)$ … dividends at time t

Equations

$$Y(t) = Y(0)e^{gt}$$

$$D(t) = D(0)e^{gt} = aP(0)e^{gt}$$

$$dP(t)/dt = rP - D(t) = rP - aP(0)e^{gt}$$

Solving the differential equation, we have:

$$\begin{aligned} P(t) &= P(0)\{(r-g-a)e^{rt} + ae^{gt}\}/(r-g) \\ &= P(0)\{e^{rt} - (a/(r-g))(e^{rt} - e^{gt})\} \end{aligned}$$

Taking the ratio of prices to GDP, we have:

$$\begin{aligned} P(t)/Y(t) &= \{P(0)/Y(0)\}\{(r-g-a)e^{(r-g)t} + a\}/(r-g) \\ &= \{P(0)/Y(0)\}\{(e^{(r-g)t} - (a/(r-g))(e^{(r-g)t} - 1)\} \end{aligned}$$

Consistent with the Gordon formula, a constant ratio of P/Y (that is, a steady state) follows from $r = g + a$. As a non-steady-state example—with values of .07 for r, .015 for g, and .03 for a—$P(75)/Y(75) = 28.7P(0)/Y(0)$.

Appendix B: Calculation Using the Gordon Equation

In discrete time, once we are in a steady state, the Gordon growth model relates a stock price *P* at time *t* to the expected dividend *D* in the following period, the rate of growth of dividends *G*, and the rate of return on the stock *R*. Therefore, we have:

$$P_t = D_{t+1}/(R-G) = (1+G)D_t/(R-G)$$

We denote values after a decade (when we are assumed to be in a steady state) by *P*' and *D*' and use an "adjusted" initial dividend that starts at a ratio *X* times current stock prices. Thus, we assume that dividends grow at the rate *G* from the "adjusted" current value for 10 years, where *G* coincides with GDP growth over the decade. We assume that dividends grow at *G*' thereafter, which coincides with long-run GDP growth. Thus, we have:

$$\begin{aligned} P'/P &= (1+G')D'/((R-G')P) \\ &= (1+G')D(1+G)^{10}/((R-G')P) \\ &= X(1+G')(1+G)^{10}/(R-G') \end{aligned}$$

For the basic calculation, we assume that *R* is .07, *G* is .02, *G*' is .015. In this case, we have:

$$P'/P = 22.5X$$

Thus, for initial ratios of adjusted dividends to stock prices of .02, .025, .03, and .035, *P*'/*P* equals .45, .56, .67 and .79, respectively. Subtracting those numbers from 1 yields the required decline in the real value of stock prices as shown in the first column of Table 3. Converting them into nominal values by multiplying by 1.27, we have values of .57, .71, and .86. If the long-run stock return is assumed to be 6.5 percent instead of 7.0 percent, the ratio *P*'/*P* is higher and the required decline is smaller. Increasing GDP growth also reduces the required decline. Note that the required declines in stock values in Table 3 is the decline in real values; the decline in nominal terms would be less.

Appendix C: A Cobb-Douglas Solow Growth Model in Steady State

Variables

Y ……… output
K ……… capital
L ………. labor
a ………. growth rate of Solow residual
g ………. growth rate of both *K* and *Y*
n ………. growth rate of labor
b ………. share of labor
s ………. savings rate
c ………. depreciation rate
MP(*K*) … marginal product of capital

Equations

$\log[Y] = at + b\log[L] + (1-b)\log[K]$

$(dL/dt)/L = n$

$(dY/dt)/Y = (dK/dt)/K = g$

$dK/dt = sY - cK$

$(dK/dt)/K = sY/K - c$

$Y/K = (g+c)/s$

$MP(K) = (1-b)Y/K = (1-b)(g+c)/s$

$g = a + bn + (1-b)g$

$g = (a+bn)/b$

$MP(K) = (1-b)\{(a+bn)/(bs) + c/s\}$

$dMP(K)/da = (1-b)/(bs)$

$dg/da = 1/b$

Assume that the share of labor is .75 and the gross savings rate is .2. Then the change in the marginal product of capital from a change in the growth rate is:

$$dMP(K)/dg = (dMP(K)/da)/(dg/da) = (1-b)/s == .25/.2$$

(Note that these are gross savings, not net savings. But the corporate income tax reduces the return to savers relative to the return to corporate capital, so the derivative should be multiplied by roughly 2/3.)

Similarly, we can consider the effect of a slowdown in labor force growth on the marginal product of capital:

$$dMP(K)/dn = (1-b)/s$$

$$dg/dn = 1$$

$$dMP(K)/dg = (dMP(K)/dn)/(dg/dn) = (1-b)/s == .25/.2$$

(This is the same expression as when the slowdown in economic growth comes from a drop in technical progress.)

Turning to the effects of changes in the savings rate, we have:

$$dMP(K)/ds = -MP(K)/s == .5$$

Thus, the savings rate has a large impact on the marginal product of capital as well.

Both of these effects are attenuated to the extent that the economy is open and rates of return in the United States change less because some of the effect occurs abroad.

Tracking the Slowdown in Long-Run GDP Growth*

Juan Antolin-Diaz — Fulcrum Asset Management

Thomas Drechsel — LSE, CFM

Ivan Petrella — Bank of England, Birkbeck, CEPR

This draft: January 25, 2016 First draft: October 24, 2014

Abstract

Using a dynamic factor model that allows for changes in both the long-run growth rate of output and the volatility of business cycles, we document a significant decline in long-run output growth in the United States. Our evidence supports the view that most of this slowdown occurred prior to the Great Recession. We show how to use the model to decompose changes in long-run growth into its underlying drivers. At low frequencies, a decline in the growth rate of labor productivity appears to be behind the recent slowdown in GDP growth for both the US and other advanced economies. When applied to real-time data, the proposed model is capable of detecting shifts in long-run growth in a timely and reliable manner.

Keywords: Long-run growth; Business cycles; Productivity; Dynamic factor models; Real-time data.

JEL Classification Numbers: E32, E23, O47, C32, E01.

---

*Antolin-Diaz: Department of Macroeconomic Research, Fulcrum Asset Management, London W1H 5BT, UK; E-Mail: juan.antolin-diaz@fulcrumasset.com. Drechsel: Department of Economics and Centre for Macroeconomics, LSE, London, WC2A 2AE, UK; E-Mail: t.a.drechsel@lse.ac.uk. Petrella: Structural Economic Analysis Division, Bank of England, London EC2R 8AH, UK; E-Mail: Ivan.Petrella@bankofengland.co.uk.

The views expressed in this paper are solely the responsibility of the authors and cannot be taken to represent those of Fulcrum Asset Management or the Bank of England, or to state Bank of England policy. This paper should not be reported as representing the views of the Bank of England or members of the Monetary Policy Committee or Financial Policy Committee. We thank Neele Balke, James Bell, Gavyn Davies, Davide Delle Monache, Wouter den Haan, Robert Gordon, Andrew Harvey, Ed Hoyle, Dimitris Korobilis, Juan Rubio-Ramirez, Ron Smith, James Stock, Paolo Surico, Jon Temple, Silvana Tenreyro, Joaquin Vespignani and the seminar participants at Banca d'Italia, Birkbeck Centre for Applied Macro, ECARES, European Central Bank, BBVA Research, the Applied Econometrics Workshop at the Federal Reserve Bank of St. Louis, the World Congress of the Econometric Society, Norges Bank and LSE for useful comments and suggestions. Alberto D'Onofrio provided excellent research assistance.

1 Introduction

> *"The global recovery has been disappointing (...) Year after year we have had to explain from mid-year on why the global growth rate has been lower than predicted as little as two quarters back".* Stanley Fischer, August 2014.

The slow pace of the recovery from the Great Recession of 2007-2009 has prompted questions about whether the long-run growth rate of GDP in advanced economies is lower now than it has been on average over the past decades (see e.g. Fernald, 2014, Gordon, 2014b, Summers, 2014). Indeed, forecasts of US and global real GDP growth have been persistently too optimistic for the last six years.[1] As emphasized by Orphanides (2003), real-time misperceptions about the long-run growth of the economy can play a large role in monetary policy mistakes. Moreover, small changes in assumptions about the long-run growth rate of output can have large implications on fiscal sustainability calculations (Auerbach, 2011). This calls for a framework that takes the uncertainty about long-run growth seriously and can inform decision-making in real time. In this paper, we present a dynamic factor model (DFM) which allows for gradual changes in the mean and the variance of real output growth. By incorporating a broad panel of economic activity indicators, DFMs are capable of precisely estimating the cyclical comovement in macroeconomic data in a real-time setting. Our model exploits this to track changes in the long-run growth rate of real GDP in a timely and reliable manner, separating them from their cyclical counterpart.[2]

The evidence of a decline in long-run US growth is accumulating, as documented

[1] For instance, Federal Open Market Committee (FOMC) projections since 2009 expected US growth to accelerate substantially, only to downgrade the forecast back to 2% throughout the course of the subsequent year. An analysis of forecasts produced by international organizations and private sector economists reveals the same pattern, see Pain et al. (2014) for a retrospective.

[2] Throughout this paper, our concept of the long run refers to changes in growth that are permanent in nature, i.e. do not mean-revert, as in Beveridge and Nelson (1981). In practice this should be thought of as frequencies lower than the business cycle.

by the recent growth literature such as Fernald and Jones (2014). Lawrence Summers and Robert Gordon have articulated a particularly pessimistic view of long-run growth which contrasts with the optimism prevailing before the Great Recession (see Jorgenson et al., 2006). To complement this evidence, we start our analysis by presenting the results of two popular structural break tests proposed by Nyblom (1989) and Bai and Perron (1998). Both suggest that a possible shift in the mean of US real GDP growth exists, the latter approach suggesting that a break probably occurred in the early part of the 2000's.[3] However, sequential testing using real-time data reveals that the break would not have been detected at conventional significance levels until as late as mid-2014, highlighting the problems of conventional break tests for real-time analysis (see also Benati, 2007). To address this issue, we introduce two novel features into an otherwise standard DFM of real activity data. First, we allow the mean of real GDP growth, and possibly other series, to drift gradually over time. As emphasized by Cogley (2005), if the long-run output growth rate is not constant, it is optimal to give more weight to recent data when estimating its current state. By taking a Bayesian approach, we can combine our prior beliefs about the rate at which the past information should be discounted with the information contained in the data. We also characterize the uncertainty around estimates of long-run growth taking into account both filtering and parameter uncertainty. Second, we allow for stochastic volatility (SV) in the innovations to both factors and idiosyncratic components. Given our interest in studying the entire postwar period, the inclusion of SV is essential to capture the substantial changes in the volatility of output that have taken place in this sample, such as the "Great Moderation" first reported by Kim and Nelson (1999a) and McConnell and Perez-Quiros (2000), as well as the cyclicality of macroeconomic volatility as documented by Jurado et al. (2014).

[3]This finding is consistent with the analysis of US real GDP by Luo and Startz (2014), as well as Fernald (2014), who applies the Bai and Perron (1998) test to US labor productivity.

When applied to US data, our model concludes that long-run GDP growth declined meaningfully during the 2000's and currently stands at about 2%, more than one percentage point lower than the postwar average. The results are supportive of a gradual decline rather than a discrete break. Since in-sample results obtained with revised data often underestimate the uncertainty faced by policymakers in real time, we repeat the exercise using real-time vintages of data. The model detects the fall from the beginning of the 2000's onwards, and by the summer of 2010 it reaches the significant conclusion that a decline in long-run growth is behind the slow recovery, well before the structural break tests become conclusive.

We also investigate the performance of the model in "nowcasting" short-term developments in GDP. Since the seminal contributions of Evans (2005) and Giannone et al. (2008) DFMs have become the standard tool for this purpose.[4] Interestingly, our analysis shows that standard DFM forecasts revert very quickly to the unconditional mean of GDP, so taking into account the variation in long-run GDP growth substantially improves point and density GDP forecasts even at very short horizons.

Finally, we extend our model in order to disentangle the drivers of secular fluctuations of GDP growth. Edge et al. (2007) emphasize the relevance as well as the difficulty of tracking permanent shifts in productivity growth in real time. In our framework, long-run output growth can be decomposed into labor productivity and labor input trends. The results of this decomposition exercise point to a slowdown in labor productivity as the main driver of recent weakness in GDP growth. Applying the model to other advanced economies, we provide evidence that the weakening in labor productivity appears to be a global phenomenon.

Our work is closely related to two strands of literature. The first one encompasses papers that allow for structural changes within the DFM framework. Del Negro and

[4] An extensive survey of the nowcasting literature is provided by Banbura et al. (2012), who also demonstrate, in a real-time context, the good out-of-sample performance of DFM nowcasts.

Otrok (2008) model time variation in factor loadings and volatilities, while Marcellino et al. (2014) show that the addition of SV improves the performance of the model for short-term forecasting of euro area GDP.[5] Acknowledging the importance of allowing for time-variation in the means of the variables, Stock and Watson (2012) pre-filter their data set in order to remove any low-frequency trends from the resulting growth rates using a biweight local mean. In his comment to their paper, Sims (2012) suggests to explicitly model, rather than filter out, these long-run trends, and emphasizes the importance of evolving volatilities for describing and understanding macroeconomic data. We see the present paper as extending the DFM literature, and in particular its application to tracking GDP, in the direction suggested by Chris Sims. The second strand of related literature takes a similar approach to decomposing long-run GDP growth into its drivers, in particular Gordon (2010, 2014a) and Reifschneider et al. (2013). Relative to these studies, we emphasize the importance of using a broader information set, as well as a Bayesian approach, which allows to use priors to inform the estimate of long-run growth, and to characterize the uncertainty around the estimate stemming both from filtering and parameter uncertainty.

The remainder of this paper is organized as follows. Section 2 presents preliminary evidence of a slowdown in long-run US GDP growth. Section 3 discusses the implications of time-varying long-run output growth and volatility for DFMs and presents our model. Section 4 applies the model to US data and documents the decline in long-run growth. The implications for tracking GDP in real time as well as the key advantages of our methodology are discussed. Section 5 decomposes the changes in long-run output growth into its underlying drivers. Section 6 concludes.

[5]While the model of Del Negro and Otrok (2008) includes time-varying factor loadings, the means of the observable variables are still treated as constant.

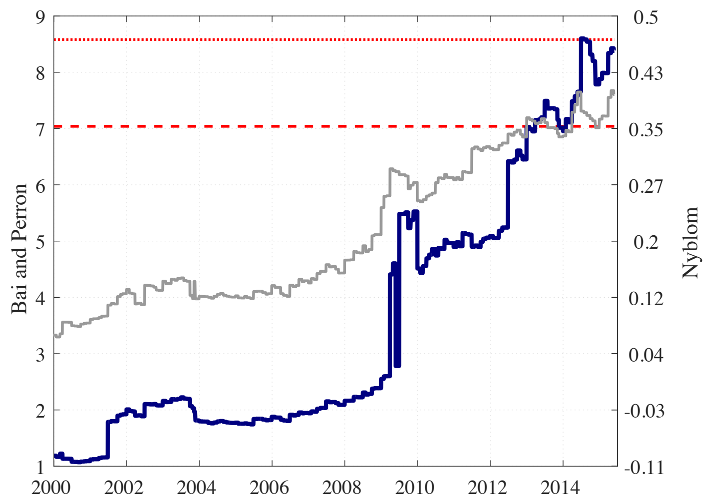
2 Preliminary Evidence

The literature on economic growth favors a view of the long-run growth rate as a process that evolves over time. It is by now widely accepted that a slowdown in productivity and long-run output growth occurred in the early 1970's, and that accelerating productivity in the IT sector led to a boom in the late 1990's.[6] In contrast, in the context of econometric modeling the possibility that long-run growth is time-varying is the source of a long-standing controversy. In their seminal contribution, Nelson and Plosser (1982) model the (log) level of real GDP as a random walk with drift. This implies that after first-differencing, the resulting growth rate fluctuates around a constant mean, an assumption still embedded in many econometric models. After the slowdown in productivity became apparent in the 1970's, many researchers such as Clark (1987) modeled the drift term as an additional random walk, implying that the level of GDP is integrated of order two. The latter assumption would also be consistent with the local linear trend model of Harvey (1985), the Hodrick and Prescott (1997) filter, and Stock and Watson (2012)'s practice of removing a local biweight mean from the growth rates before estimating a DFM. The $I(2)$ assumption is nevertheless controversial since it implies that the growth rate of output can drift without bound. Consequently, papers such as Perron and Wada (2009), have modeled the growth rate of GDP as stationary around a trend with one large break around 1973.

Ever since the Great Recession of 2007-2009 US real GDP has grown well below its postwar average, once again raising the question whether its mean may have declined. There are two popular strategies that could be followed from a frequentist perspective to detect parameter instability or the presence of breaks in the mean growth rate. The first one is Nyblom's (1989) L-test as described in Hansen (1992), which tests

[6] For a retrospective on the productivity slowdown, see Nordhaus (2004). Oliner and Sichel (2000) provide evidence on the role of the IT sector in the acceleration of the late 1990's.

Figure 1: Real-Time Test Statistics of the Nyblom and Bai-Perron Tests



Note: The gray and blue lines are the values of the test statistics obtained from sequentially re-applying the Nyblom (1989) and Bai and Perron (1998) tests in real time as new National Accounts vintages are being published. In both cases, the sample starts in 1947:Q2 and the test is re-applied for every new data release occurring after the beginning of 2000. The dotted and dashed red lines represent the 5% and 10% critical values corresponding to the two tests.

the null hypothesis of constant parameters against the alternative that the parameters follow a martingale. Modeling real GDP growth as an AR(1) over the sample 1947-2015 this test rejects the stability of the constant term at the 10% significance level.[7] The second commonly used approach, which can determine the number and timing of multiple discrete breaks, is the Bai and Perron (1998) test. This test finds evidence in favor of a single break in the mean of US real GDP growth at the 10%-level. The most likely break date is in the second quarter of 2000. In related research, Fernald (2014) provides evidence for breaks in labor productivity in 1973:Q2, 1995:Q3, and 2003:Q1, and links the latter two to developments in the IT sector. From a Bayesian perspective, Luo and Startz (2014) calculate the posterior probability of a single break and find the most likely break date to be 2006:Q1 for the full postwar sample and 1973:Q1 for a

[7]The same result holds for an AR(2) specification. In both cases, stability of the autoregressive coefficients cannot be rejected, whereas stability of the variance is rejected at the 1%-level. Appendix B provides the full results of both tests applied in this section. The appendix to the paper is available at: `http://personal.lse.ac.uk/drechsel/papers/ADP_appendix.pdf`.

sample excluding the 2000's.

The above results indicate that substantial evidence for a recent change in the mean of US GDP growth has built up. However, the strategy of applying conventional tests and introducing deterministic breaks into econometric models is not satisfactory for the purposes of real-time decision making. In fact, the detection of change in the mean of GDP growth can arrive with substantial delay. To demonstrate this, a sequential application of the Nyblom (1989) and Bai and Perron (1998) tests using real-time data is presented in Figure 1. The evolution of the test statistics in real-time reveals that a break would not have been detected at the 10% significance levels until as late as mid-2012, which is more than ten years later than the actual break date suggested by the Bai and Perron (1998) procedure. The Nyblom (1989) test, which is designed to detect gradual change rather than a discrete break, becomes significant roughly at the same time. This lack of timeliness highlights the importance of an econometric framework capable of quickly adapting to changes in long-run growth as new information arrives.

3 Econometric Framework

DFMs in the spirit of Geweke (1977), Stock and Watson (2002) and Forni et al. (2009) capture the idea that a small number of unobserved factors drives the comovement of a possibly large number of macroeconomic time series, each of which may be contaminated by measurement error or other sources of idiosyncratic variation. Their theoretical appeal (see e.g. Sargent and Sims, 1977 or Giannone et al., 2006), as well as their ability to parsimoniously model large data sets, have made them a workhorse of empirical macroeconomics. Giannone et al. (2008) and Banbura et al. (2012) have pioneered the use of DFMs to produce current-quarter forecasts ("nowcasts") of GDP growth by exploiting more timely monthly indicators and the factor structure of the data. Given the widespread use of DFMs to track GDP in real time, this paper aims to make these models robust to changes in long-run growth. We do so by introducing two novel features into the DFM framework. First, we allow the long-run growth rate of real GDP growth, and possibly other series, to vary over time. Second, we allow for stochastic volatility (SV) in the innovations to both factors and idiosyncratic components, given our interest in studying the entire postwar period for which drastic changes in volatility have been documented. With these changes, the DFM proves to be a powerful tool to detect changes in long-run growth. The information contained in a broad panel of activity indicators facilitates the timely decomposition of real GDP growth into persistent long-run movements, cyclical fluctuations and short-lived noise.

3.1 The Model

Let $\mathbf{y}_t$ be an $n \times 1$ vector of observable macroeconomic time series, and let $\mathbf{f}_t$ denote a $k \times 1$ vector of latent common factors. It is assumed that $n >> k$, i.e. the number of observables is much larger than the number of factors. Formally,

$$\mathbf{y}_t = \mathbf{c}_t + \mathbf{\Lambda}\mathbf{f}_t + \mathbf{u}_t, \tag{1}$$

where $\mathbf{\Lambda}$ contains the loadings on the common factors and $\mathbf{u}_t$ is a vector of idiosyncratic components.[8] Shifts in the long-run mean of $\mathbf{y}_t$ are captured by time-variation in $\mathbf{c}_t$. In principle one could allow time-varying intercepts in all or a subset of the variables in the system. Moreover, time variation in a given series could be shared by other series.

[8]The model can be easily extended to include lags of the factor in the measurement equation. In the latter case, it is sensible to avoid overfitting by choosing priors that shrink the additional lag coefficients towards zero (see D'Agostino et al., 2015, and Luciani and Ricci, 2014). We consider this possibility when we explore robustness of our results to using larger data panels in Section 4.6.

$\mathbf{c}_t$ is therefore flexibly specified as

$$\mathbf{c}_t = \begin{bmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{c} \end{bmatrix} \begin{bmatrix} \mathbf{a}_t \\ 1 \end{bmatrix}, \tag{2}$$

where $\mathbf{a}_t$ is an $r \times 1$ vector of time-varying means, $\mathbf{B}$ is an $m \times r$ matrix which governs how the time-variation affects the corresponding observables, and $\mathbf{c}$ is an $(n - m) \times 1$ vector of constants. In our baseline specification, $\mathbf{a}_t$ will be a scalar capturing time-variation in long-run real GDP growth, which is shared by real consumption growth, so that $r = 1, m = 2$. A detailed discussion of this and additional specifications of $\mathbf{c}_t$ will be provided in Section 3.2.

Throughout the paper, we focus on the case of a single dynamic factor by setting $k = 1$ (i.e. $\mathbf{f}_t = f_t$).[9] The laws of motion of the latent factor and the idiosyncratic components are

$$\begin{aligned} (1 - \phi(L))f_t &= \sigma_{\varepsilon_t}\varepsilon_t, && (3) \\ (1 - \rho_i(L))u_{i,t} &= \sigma_{\eta_{i,t}}\eta_{i,t}, \quad i = 1, \dots, n && (4) \end{aligned}$$

where $\phi(L)$ and $\rho_i(L)$ denote polynomials in the lag operator of order p and q, respectively. The idiosyncratic components are cross-sectionally orthogonal and are assumed to be uncorrelated with the common factor at all leads and lags, i.e. $\varepsilon_t \overset{iid}{\sim} N(0, 1)$ and $\eta_{i,t} \overset{iid}{\sim} N(0, 1)$.

Finally, the dynamics of the model's time-varying parameters are specified to follow

[9] For the purpose of tracking real GDP with a large number of closely related activity indicators, the use of one factor is appropriate, which is explained in more detail in Sections 4.1 and 4.2. Also note that we order real GDP growth as the first element of $\mathbf{y}_t$, and normalize the loading for GDP to unity. This serves as an identifying restriction in our estimation algorithm. Bai and Wang (2015) discuss minimal identifying assumptions for DFMs.

driftless random walks:

$$a_{j,t} = a_{j,t-1} + v_{a_{j,t}}, \qquad v_{a_{j,t}} \overset{iid}{\sim} N(0, \omega_{a,j}^2) \quad j = 1, \ldots, r \tag{5}$$

$$\log \sigma_{\varepsilon_t} = \log \sigma_{\varepsilon_{t-1}} + v_{\varepsilon,t}, \qquad v_{\varepsilon,t} \overset{iid}{\sim} N(0, \omega_{\varepsilon}^2) \tag{6}$$

$$\log \sigma_{\eta_{i,t}} = \log \sigma_{\eta_{i,t-1}} + v_{\eta_{i,t}}, \qquad v_{\eta_{i,t}} \overset{iid}{\sim} N(0, \omega_{\eta,i}^2) \quad i = 1, \ldots, n \tag{7}$$

where $a_{j,t}$ are the r time-varying elements in $\mathbf{a}_t$, and σ_{ε_t} and $\sigma_{\eta_{i,t}}$ capture the SV of the innovations to factor and idiosyncratic components. Our motivation for specifying the time-varying parameters as random walks is similar to Primiceri (2005). While in principle it is unrealistic model real GDP growth as a process that could wander in an unbounded way, as long as the variance of the process is small and the drift is considered to be operating for a finite period of time, the assumption is innocuous. Moreover, modeling a trend as a random walk is more robust to misspecification when the actual process is instead characterized by discrete breaks, whereas models with discrete breaks might not be robust to the true process being a random walk.[10] Finally, the random walk assumption also has the desirable feature that, unlike stationary models, confidence bands around forecasts of real GDP growth increase with the forecast horizon, reflecting uncertainty about the possibility of future breaks or drifts in long-run growth.

Note that a standard DFM is usually specified under two assumptions. First, the original data have been differenced appropriately so that both the factor and the idiosyncratic components can be assumed to be stationary. Second, it is assumed that the innovations in the idiosyncratic and common components are iid. In equations (1)-(7) we have relaxed these assumptions to allow for two novel features, a stochastic

[10]We demonstrate this point with the use of Monte Carlo simulations, showing that a random walk trend in real GDP growth 'learns' quickly about a discrete break once it has occurred. On the other hand, the random walk does not detect a drift when there is not one, despite the presence of a large cyclical component. Appendix C provides a discussion and the full results of these simulations.

trend in the mean of selected series, and SV. By shutting down these features, we can recover the specifications previously proposed in the literature, which are nested in our framework. We obtain the DFM with SV of Marcellino et al. (2014) if we shut down time-variation in the intercepts of the observables, i.e. set $r = m = 0$ and $\mathbf{c}_t = \mathbf{c}$. If we further shut down the SV, i.e. set $\omega^2_{a,j} = \omega^2_{\epsilon} = \omega^2_{\eta,i} = 0$, we obtain the specification of Banbura and Modugno (2014) and Banbura et al. (2012).

3.2 A Baseline Specification for Long-Run Growth

Equations (1) and (2) allow for stochastic trends in the mean of all or a subset of selected observables in $\mathbf{y}_t$. This paper focuses on tracking changes in the long-run growth rate of real GDP. For this purpose, the simplest specification of $\mathbf{c}_t$ is to include a time-varying intercept only in GDP and to set $\mathbf{B} = 1$. However, a number of empirical studies (e.g. Harvey and Stock, 1988, Cochrane, 1994, and Cogley, 2005) argue that incorporating information about consumption is informative about the permanent component in GDP as predicted by the permanent income hypothesis. The theory predicts that consumers, smoothing consumption throughout their lifetime, should react more strongly to permanent, as opposed to transitory, changes in income. As a consequence, looking at GDP and consumption data together will help separating growth into long-run and cyclical fluctuations.[11] Therefore, our baseline specification imposes that consumption and output grow at the same rate g_t in the long-run.

Formally, ordering real GDP and consumption growth first, and setting $m = 2$ and $r = 1$, this is represented as

$$\mathbf{a}_t = g_t, \quad \mathbf{B} = [1\ 1]' \tag{8}$$

[11] While a strict interpretation of the permanent income hypothesis is rejected in the data, from an econometric point of view the statement applies as long as permanent changes are the main driver of consumption. See Cochrane (1994) for a very similar discussion.

Note that in this baseline specification we model time-variation only in the intercept for GDP and consumption while leaving it constant for the other observables. Of course it may be the case that some of the remaining $n - m$ series in $\mathbf{y}_t$ feature low frequency variation in their means. The key question is whether leaving this variation in other series unspecified will affect the estimate of the long-run growth rate of GDP, which is our main object of interest. We ensure that this is not the case by allowing for persistence (and, in particular, we do not rule out unit roots) in the idiosyncratic components. If a series does feature a unit root which is not included in $\mathbf{a}_t$, its trend component will be absorbed by the idiosyncratic component. The choice of which elements to include in $\mathbf{a}_t$ therefore reflects the focus of a particular application.[12] Of course, if two series share the same underlying low-frequency component, and this is known with certainty, explicitly accounting for the shared low frequency variation will improve the precision of the estimation, but the risk of incorrectly including the trend is much larger than the risk of incorrectly excluding it. Therefore, in our baseline specification we include in $\mathbf{a}_t$ the intercept for GDP and consumption, while leaving any possible low-frequency variation in other series to be captured by the respective idiosyncratic components.[13]

An extension to include additional time-varying intercepts is straightforward through the flexible construction of $\mathbf{c}_t$ in equation (2). In fact, in Section 5 we explore how interest in the low frequency movements of additional series leads to alternative

[12]In principle, these unmodeled trends could still be recovered from our specification by applying a Beveridge-Nelson decomposition to its estimated idiosyncratic component. In practice, any low-frequency variation in the idiosyncratic component is likely to be obscured by a large amount of high frequency noise in the data and as result the extracted Beveridge-Nelson trend component will be imprecisely estimated, and as Morley et al. (2003) show, will not be smooth. In our specification, the elements of $\mathbf{a}_t$ are instead extracted directly, so that we are able to improve the extraction by imposing additional assumptions (e.g. smoothness) and prior beliefs (e.g. low variability) on its properties.

[13]We confirm this line of reasoning with a series of Monte Carlo experiments, in which data is generated from a system that features low frequency movements in more series, which are left unmodeled in the estimation. Both in the case of series with independent trends and the case of series which share the trend of interest, the fact that they are left unmodeled has little impact on the estimate of the latter. Appendix C presents further discussion and the full results of these simulations.

choices for $\mathbf{a}_t$ and $\mathbf{B}$.[14]

3.3 Dealing with Mixed Frequencies and Missing Data

Tracking activity in real time requires a model that can efficiently incorporate information from series measured at different frequencies. In particular, it must include both quarterly variables, such as the growth rate of real GDP, as well as more timely monthly indicators of real activity. Therefore, the model is specified at monthly frequency, and following Mariano and Murasawa (2003), the (observed) quarterly growth rates of a generic quarterly variable, y_t^q, can be related to the (unobserved) monthly growth rate y_t^m and its lags using a weighted mean. Specifically,

$$y_t^q = \frac{1}{3}y_t^m + \frac{2}{3}y_{t-1}^m + y_{t-2}^m + \frac{2}{3}y_{t-3}^m + \frac{1}{3}y_{t-4}^m, \tag{9}$$

and only every third observation of y_t^q is actually observed. Substituting the corresponding line of (1) into (9) yields a representation in which the quarterly variable depends on the factor and its lags. The presence of mixed frequencies is thus reduced to a problem of missing data in a monthly model.

Besides mixed frequencies, additional sources of missing data in the panel include: the "ragged edge" at the end of the sample, which stems from the non-synchronicity of data releases; missing data at the beginning of the sample, since some data series have been created or collected more recently than others; and missing observations due to outliers and data collection errors. Our Bayesian estimation method exploits the state space representation of the DFM and jointly estimates the latent factors,

[14] Note that the limiting case explicitly models time-varying intercept in all indicators, so that $m = r = n$ and $\mathbf{B} = \mathbf{I}_n$, i.e. an identity matrix of dimension n. See Creal et al. (2010) and Fleischman and Roberts (2011) for similar approaches. This setup would imply that the number of state variables increases with the number of observables, which severely increases the computational burden of the estimation, while offering little additional evidence with respect to the focus of this paper.

the parameters, and the missing data points using the Kalman filter (see Durbin and Koopman, 2012, for a textbook treatment).

3.4 State Space Representation and Estimation

The model features autocorrelated idiosyncratic components (see equation (4)). In order to cast it in state-space form, we include the idiosyncratic components of the quarterly variables in the state vector, and we redefine the system for the monthly indicators in terms of quasi-differences (see e.g. Kim and Nelson, 1999b, pp. 198-199, and Bai and Wang, 2015).[15] The model is estimated with Bayesian methods simulating the posterior distribution of parameters and factors using a Markov Chain Monte Carlo (MCMC) algorithm. We closely follow the Gibbs-sampling algorithm for DFMs proposed by Bai and Wang (2015), but extend it to include mixed frequencies, the time-varying intercept, and SV. The SVs are sampled using the approximation of Kim et al. (1998), which is considerably faster than the exact Metropolis-Hastings algorithm of Jacquier et al. (2002). Our complete sampling algorithm together with the details of the state space representation can be found in Appendix D.

4 Results for US Data

4.1 Data Selection

Our data set includes four key business cycle variables measured at quarterly frequency (output, consumption, investment and aggregate hours worked), as well as a set

[15]Since the quarterly variables are observed only every third month, we cannot take the quasi-difference for their idiosyncratic components, which are instead added as an additional state with the corresponding transition dynamics. Banbura and Modugno (2014) suggest including all of the idiosyncratic components as additional elements of the state vector. Our solution has the desirable feature that the number of state variables will increase with the number of quarterly variables, rather than the total number of variables, leading to a gain of computational efficiency.

of 24 monthly indicators which are intended to provide additional information about cyclical developments in a timely manner.

The included quarterly variables are strongly procyclical and are considered key indicators of the business cycle (see e.g. Stock and Watson, 1999). Furthermore, theory predicts that they will be useful in disentangling low frequency movements from cyclical fluctuations in output growth. Indeed, as discussed in Section 3.2, the permanent income hypothesis predicts that consumption data will be particularly useful for the estimation of the long-run growth component, g_t.[16] On the other hand, investment and hours worked are very sensitive to cyclical fluctuations, and thus will be particularly informative for the estimation of the cyclical factor, f_t.[17]

The additional monthly indicators are crucial to our objective of disentangling in real time the cyclical and long-run components of GDP growth, since the quarterly variables are only available with substantial delay. In principle, a large number of candidate series are available to inform the estimate of f_t, and indirectly, of g_t. In practice, however, macroeconomic data series are typically clustered in a small number of broad categories (such as production, employment, or income) for which disaggregated series are available along various dimensions (such as economic sectors, demographic characteristics, or expenditure categories). The choice of which available series to include for estimation can therefore be broken into, first, a choice of which broad categories to

[16]Due to the presence of faster technological change in the durable goods sector there is a downward trend in the relative price of durable goods. As a consequence, measured consumption grows faster than overall GDP. Following a long tradition in the literature (see e.g. Whelan, 2003), we construct a Fisher index of non-durables and services and use its growth rate as an observable variable in the panel. It can be verified that the ratio of consumption defined in this manner to real GDP displays no trend in the data, unlike the trend observed in the ratio of overall consumption to GDP.

[17]We define investment as a chain-linked aggregate of business fixed investment and consumption of durable goods, which is consistent with our treatment of consumption. In order to obtain a measure of hours for the total economy, we follow the methodology of Ohanian and Raffo (2012) and benchmark the quarterly series of hours in the non-farm business sector provided by the BLS to the annual estimates of hours in the total economy compiled by the Conference Boards Total Economy Database (TED). The TED series has the advantage of being comparable across countries (Ohanian and Raffo, 2012), which will be useful for our international results in Section 5.

include, and second, to which level and along which dimensions of disaggregation.

With regards to which broad categories of data to include, previous studies agree that prices, monetary and financial indicators are uninformative for the purpose of tracking real GDP, and argue for extracting a single common factor that captures real economic activity.[18] As for the possible inclusion of disaggregated series within each category, Boivin and Ng (2006) argue that the presence of strong correlation in the idiosyncratic components of disaggregated series of the same category will be a source of misspecification that can worsen the performance of the model in terms of in-sample fit and out-of-sample forecasting of key series.[19] Alvarez et al. (2012) investigate the trade-off between DFMs with very few indicators, where the good large-sample properties of factor models are unlikely to hold, and those with a very large amount of indicators, where the problems above are likely to arise. They conclude that using a medium-sized panel with representative indicators of each category yields the best forecasting results.

The above considerations lead us to select 24 monthly indicators that include the high-level aggregates for all of the available broad categories that capture real activity, without overweighting any particular of these categories. The complete list of variables contained in our data set is presented in Table 1. As the table shows, we include representative series of expenditure and income, the labor market, production and sales, foreign trade, housing and business and consumer confidence.[20] The inclusion of all the available monthly surveys is particularly important. Apart from being the most timely series available, these are stationary by construction, and have a high signal-to-

[18]Giannone et al. (2005) conclude that that prices and monetary indicators do not contribute to the precision of GDP nowcasts. Banbura et al. (2012), Forni et al. (2003) and Stock and Watson (2003) find at best mixed results for financial variables.

[19]This problem is exacerbated by the fact that more detailed disaggregation levels and dimensions are available for certain categories of data, such as employment, meaning that the disaggregation will automatically 'tilt' the factor estimates towards that category.

[20]Whenever there are multiple candidates for the higher level aggregate of a particular category, we include them both. For example, we include employment as measured both by the establishment and household surveys, and consumer confidence as measured both by the Conference Board and the University of Michigan surveys.

noise ratio, providing a clean signal to separate the cyclical component of GDP growth from its long-run counterpart. In Section 4.6 we explore sensitivity of our results to the size and the composition of the data panel used.

Our panel spans the period January 1947 to March 2015. The start of our sample coincides with the year for which quarterly national accounts data are available from the Bureau of Economic Analysis. This enables us to study the evolution of long-run growth over the entire postwar period.[21]

4.2 Model Settings and Priors

The choice of the data set justifies the single-factor structure of the model. f_t can in this case be interpreted as a coincident indicator of real economic activity (see e.g. Stock and Watson, 1989, and Mariano and Murasawa, 2003). The number of lags in the polynomials $\phi(L)$ and $\rho(L)$ is set to $p = 2$ and $q = 2$ as in Stock and Watson (1989).

We wish to impose as little prior information as possible, so we use uninformative priors for the factor loadings and the autoregressive coefficients of factors and idiosyncratic components. The variances of the innovations to the time-varying parameters, namely ω_a^2, ω_ε^2 and $\omega_{\eta,i}^2$ in equations (5)-(7) are however difficult to identify from the information contained in the likelihood alone. As the literature on Bayesian VARs documents, attempts to use non-informative priors for these parameters will in many cases produce posterior estimates which imply a relatively large amount of time-variation. While this will tend to improve the in-sample fit of the model it is also likely to worsen

[21]We take full advantage of the Kalman filter's ability to deal with missing observations at any point in the sample, and we are able to incorporate series that become available substantially later than 1947, up to as late as 2007. Note that for consumption expenditures, monthly data became available in 1959, whereas quarterly data is available from 1947. In order to use all available data, we apply the polynomial in Equation (9) to the monthly data and treat the series as quarterly, with available observations for the last month of the quarter for 1947-1958 and for all months since 1959.

Table 1: DATA SERIES USED IN EMPIRICAL ANALYSIS

	Type	Start Date	Transformation	Publ. Lag
QUARTERLY TIME SERIES				
Real GDP	Expenditure & Income	Q2:1947	% QoQ Ann.	26
Real Consumption (excl. durables)	Expenditure & Income	Q2:1947	% QoQ Ann.	26
Real Investment (incl. durable cons.)	Expenditure & Income	Q2:1947	% QoQ Ann.	26
Total Hours Worked	Labor Market	Q2:1948	% QoQ Ann.	28
MONTHLY INDICATORS				
Real Personal Income less Trans. Paym.	Expenditure & Income	Feb 59	% MoM	27
Industrial Production	Production & Sales	Jan 47	% MoM	15
New Orders of Capital Goods	Production & Sales	Mar 68	% MoM	25
Real Retail Sales & Food Services	Production & Sales	Feb 47	% MoM	15
Light Weight Vehicle Sales	Production & Sales	Feb 67	% MoM	1
Real Exports of Goods	Foreign Trade	Feb 68	% MoM	35
Real Imports of Goods	Foreign Trade	Feb 69	% MoM	35
Building Permits	Housing	Feb 60	% MoM	19
Housing Starts	Housing	Feb 59	% MoM	26
New Home Sales	Housing	Feb 63	% MoM	26
Payroll Empl. (Establishment Survey)	Labor Market	Jan 47	% MoM	5
Civilian Empl. (Household Survey)	Labor Market	Feb 48	% MoM	5
Unemployed	Labor Market	Feb 48	% MoM	5
Initial Claims for Unempl. Insurance	Labor Market	Feb 48	% MoM	4
MONTHLY INDICATORS (SOFT)				
Markit Manufacturing PMI	Business Confidence	May 07	-	-7
ISM Manufacturing PMI	Business Confidence	Jan 48	-	1
ISM Non-manufacturing PMI	Business Confidence	Jul 97	-	3
NFIB: Small Business Optimism Index	Business Confidence	Oct 75	Diff 12 M.	15
U. of Michigan: Consumer Sentiment	Consumer Confidence	May 60	Diff 12 M.	-15
Conf. Board: Consumer Confidence	Consumer Confidence	Feb 68	Diff 12 M.	-5
Empire State Manufacturing Survey	Business (Regional)	Jul 01	-	-15
Richmond Fed Mfg Survey	Business (Regional)	Nov 93	-	-5
Chicago PMI	Business (Regional)	Feb 67	-	0
Philadelphia Fed Business Outlook	Business (Regional)	May 68	-	0

Notes: % QoQ Ann. refers to the quarter on quarter annualized growth rate, % MoM refers to $(y_t - y_{t-1})/y_{t-1}$ while Diff 12 M. refers to $y_t - y_{t-12}$. The last column shows the average publication lag, i.e. the number of days elapsed from the end of the period that the data point refers to until its publication by the statistical agency. All series were obtained from the Haver Analytics database.

out-of-sample forecast performance. We therefore use priors to shrink these variances towards zero, i.e. towards the standard DFM which excludes time-varying long-run GDP growth and SV. In particular, for ω_a^2 we set an inverse gamma prior with one degree of freedom and scale equal to 0.001.[22] For ω_ϵ^2 and $\omega_{\eta,i}^2$ we set an inverse gamma prior with one degree of freedom and scale equal to 0.0001, closely following Cogley and Sargent (2005) and Primiceri (2005).[23]

[22]To gain an intuition about this prior, note that over a period of ten years, this would imply that the posterior mean of the long-run growth rate is expected to vary with a standard deviation of around 0.4 percentage points in annualized terms, which is a fairly conservative prior.

[23]We provide further explanations and address robustness to the choice of priors in Appendix F.

We estimate the model with 7000 replications of the Gibbs-sampling algorithm, of which the first 2000 are discarded as burn-in draws and the remaining ones are kept for inference.[24]

4.3 In-Sample Results

Panel (a) of Figure 2 plots the posterior median, together with the 68% and 90% posterior credible intervals of the long-run growth rate of real GDP. This estimate is conditional on the entire sample and accounts for both filtering and parameter uncertainty. Several features of our estimate of long-run growth are worth noting. While the growth rate is stable between 3% and 4% during the first decades of the postwar period, a slowdown is clearly visible from around the late 1960's through the 1970's, consistent with the "productivity slowdown" (Nordhaus, 2004). The acceleration of the late 1990's and early 2000's associated with the productivity boom in the IT sector (Oliner and Sichel, 2000) is also visible. Thus, until the middle of the decade of the 2000's, our estimate conforms well to the generally accepted narrative about fluctuations in potential growth.[25] More recently, after peaking at about 3.5% in 2000, the median estimate of the long-run growth rate has fallen to about 2% in early 2015, a more substantial decline than the one observed after the productivity slowdown of the 1970's. Moreover, the slowdown appears to have happened gradually since the start of the 2000's, with most of the decline having occurred before the Great Recession.[26]

[24]Thanks to the efficient state space representation discussed above, the improvements in the simulation smoother proposed by Bai and Wang (2015), and other computational improvements we implemented, the estimation is very fast. Convergence is achieved after only 1500 iterations, which take less than 20 minutes in MATLAB using an Intel 3.6 GHz computer with 16GB of DDR3 Ram.

[25]Appendix G provides a comparison of our estimate with the Congressional Budget Office (CBO) measure of potential growth, with some additional discussion.

[26]In principle, it is possible that our choice of modeling long-run GDP growth as a random walk is hard-wiring into our results the conclusion that the decline happened in a gradual way. In experiments with simulated data, presented in Appendix C, we show that if changes in long-run growth occur in the form of discrete breaks rather than evolving gradually, the (one-sided) filtered estimates will exhibit a discrete jump at the moment of the break. Instead, for US data the filtered estimates of the long-run

Interestingly, a small rebound is visible at the end of the sample, but long-run growth stands far below its postwar average.

Panel (b) plots the time series of quarterly real GDP growth, together with the median posterior estimates of the common cyclical factor, aligned with the mean of real GDP growth. This plot highlights how the cyclical factor captures the bulk of business-cycle frequency variation in output growth, while higher frequency, quarter-to-quarter variation is attributed to the idiosyncratic component of GDP growth. In the latter part of the sample, GDP growth is visibly below the cyclical factor, reflecting the decline in long-run growth.

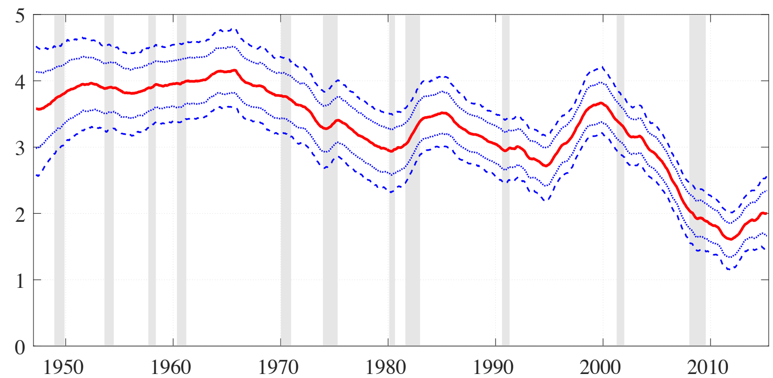
The posterior estimate of the SV of the common factor is presented in Panel (c). It is clearly visible that volatility declines over the sample. The late 1940's and 1950's were extremely volatile, with a first large drop in volatility in the early 1960's. The Great Moderation is also clearly visible, with the average volatility pre-1985 being much larger than the average of the post-1985 sample. Notwithstanding the large increase in volatility during the Great Recession, our estimate of the common factor volatility since then remains consistent with the Great Moderation still being in place. This confirms the early evidence reported by Gadea-Rivas et al. (2014). It is clear from the figure that volatility spikes during recessions, a feature that brings our estimates close to the recent findings of Jurado et al. (2014) and Bloom (2014) relating to business-cycle uncertainty.[27] It appears that the random walk specification is flexible enough to capture cyclical changes in volatility as well as permanent phenomena such as the Great Moderation. Appendix A contains analogous charts for the estimated volatilities of the idiosyncratic components of selected data series. Similar to the volatility of the

growth component also decline in a gradual manner (see Figure A.1 in Appendix A).

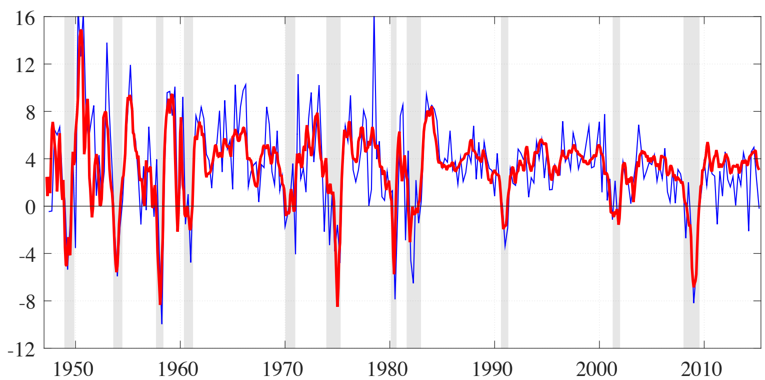
[27]It is interesting to note that while in our model the innovations to the level of the common factor and its volatility are uncorrelated, the fact that increases in volatility are observed during recessions indicate the presence of negative correlation between the first and second moments, implying negative skewness in the distribution of the common factor. We believe a more explicit model of this feature is an important priority for future research.

Figure 2: Trend, cycle and volatility: 1947-2015 (% Ann. Growth Rate)

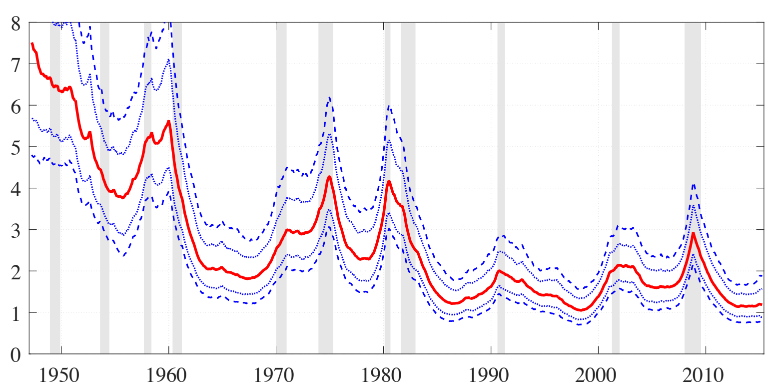
(a) Posterior estimate of long-run growth



(b) Posterior estimate and in-sample fit of cyclical factor



(c) Posterior estimate of common factor volatility



Note: Panel (a) displays the posterior median (solid red), together with the 68% and 90% (dotted and dashed blue) posterior credible intervals of long-run real GDP growth. Panel (b) plots actual real GDP growth (thin blue) against the posterior median estimate of the cyclical factor (thick red). Panel (c) presents the median (red), the 68% and the 90% (dotted and dashed blue) posterior credible intervals of the volatility of the common factor, i.e the square root of $\text{var}(f_t) = \sigma^2_{\varepsilon,t}(1-\phi_2)/[(1+\phi_2)((1-\phi_2)^2-\phi_1^2)]$. Shaded areas represent NBER recessions.

common factor, many of the idiosyncratic volatilities present sharp increases during recessions.

The above results provide evidence that a significant decline in long-run US real GDP growth occurred over the last decade, and are consistent with a relatively gradual decline since the early 2000's. Our estimates show that the bulk of the slowdown from the elevated levels of growth at the turn of the century occurred before the Great Recession, which is consistent with the narrative of Fernald (2014) on the fading of the IT productivity boom. This recent decline is the largest movement in long-run growth observed in the postwar period.

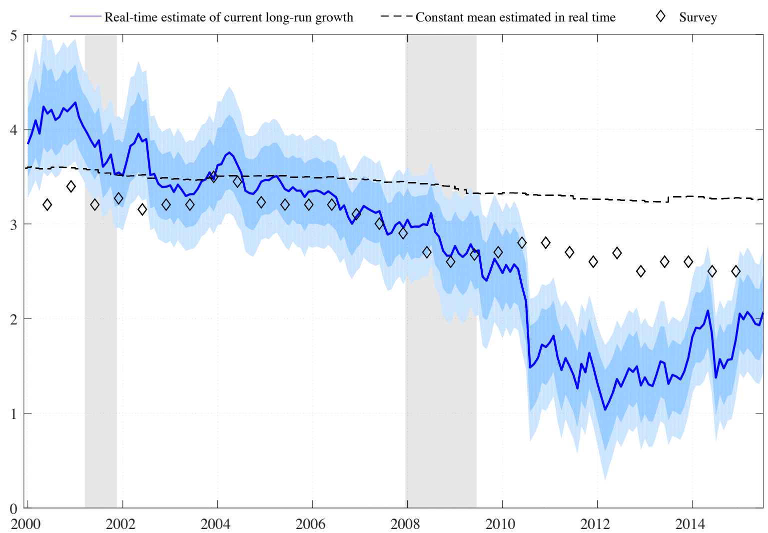
4.4 Real-Time Results

As emphasized by Orphanides (2003), macroeconomic time series are heavily revised over time and in many cases these revisions contain valuable information that was not available at initial release. Therefore, it is important to assess, using the data available at each point in time, when the model detected the slowdown in long-run growth. For this purpose, we reconstruct our data set using vintages of data available from the Federal Reserve Bank of St. Louis ALFRED data base. Our aim is to replicate as closely as possible the situation of a decision-maker which would have applied our model in real time. We fix the start of our sample in 1947:Q1 and use an expanding out-of-sample window which starts on 11 January 2000 and ends on 30 June 2015. This is the longest possible window for which we are able to include the entire panel in Table 1 using fully real-time data. We then proceed by re-estimating the model each day in which new data are released.[28]

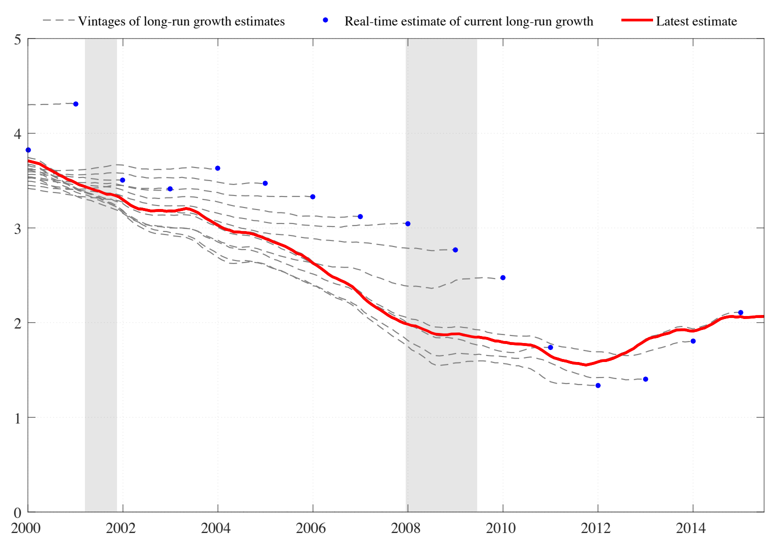
[28] In a few cases new indicators were developed after January 2000. For example, the Markit Manufacturing PMI survey is currently one of the most timely and widely followed indicators, but it started being conducted in 2007. In those cases, we append to the panel, in real time, the vintages of the new indicators as soon sufficient history is available. In the example of the PMI, this is the case since mid-2012. By implication, the number of indicators in our data panel grows when new indicators

Figure 3: Long-Run GDP Growth Estimates in Real Time

(a) Evolution of the current assessment of long-run growth



(b) Selected vintages of long-run growth estimates using real-time data



Note: The figure presents results from re-estimating the model using the vintage of data available at each point in time from January 2000 to March 2015. The start of the estimation sample is fixed at Q1:1947. Panel (a) plots the median real-time estimate of current long-run growth over time. This is the locus traced by the end points of all vintages. The blue shaded areas represent the 68th and 90th percentiles. The dashed line is the contemporaneous estimate of the historical average of real GDP growth. The diamonds are the median response to the Philadelphia Fed Livingston Survey of Professional Forecasters on the average growth rate for the next 10 years. Panel (b) displays the median estimate of long-run GDP growth for various vintages of data (dashed gray lines). The estimate of the latest vintage is shown in solid red. Gray shaded areas represent NBER recessions.

Figure 3 looks at the model's real-time assessment of the posterior distribution of long-run growth at various points in time. Panel (a) plots the real-time estimate of current long-run growth, with 68% and 90% uncertainty bands. For comparison, the panel also displays the median response to the Philadelphia Feds Livingston Survey of Professional Forecasters (SPF) on the average growth rate for the next 10 years, and the estimate of long-run growth from a model with a constant intercept for GDP growth. The latter estimate is also updated as new information arrives, but weighs all points of the sample equally. Panel (b) displays several vintages of the median long-run growth estimate, using information available up to July of each year. The locus traced by the end point of each vintage corresponds to the current real-time estimate of Panel (a).

The evolution of the baseline model's estimate of long-run growth when estimated in real time declines gradually from a peak of about 4% in early 2000 to around 2.5% just after the end of the Great Recession. From this time, the constant estimate shown in panel (a) is always outside of the 90% posterior bands. There is a sharp reassessment of long-run growth around July 2010, coinciding with the publication by the Bureau of Economic Analysis of the annual revisions to the National Accounts, which each year incorporate previously unavailable information for the previous three years. The revisions implied a substantial downgrade, in particular, to the growth of consumption in the first year of the recovery, from 2.5% to 1.6%, and instead allocated much of the growth in GDP during the recovery to inventory accumulation.[29] Reflecting the role of consumption as the most persistent and forward looking component of GDP, the estimate of long-run growth is downgraded sharply. Panel (b) shows how the 2010 revisions in fact trigger a re-interpretation of the years leading to the Great

appear. Full details about the construction of the vintage database are available in Appendix E.

[29]See Appendix I for additional figures on the National Accounts revisions during this period.

Recession. With the revised information, the bulk of the slowdown in long-run growth is now estimated to have occurred *before* the recession.[30] From 2010 onward, the model predicts a recovery that is extremely slow by historical standards. This is four years before the structural break test detected a statistically significant decline.[31] In the latest part of the sample, the estimate of long-run growth has recovered slightly to about 2%. Interestingly, this has been triggered by improvements in incoming data, rather than revisions to past vintages.

With regards to the SPF, it is noticeable that from 2003 to about 2010, the survey is remarkably similar to the model, but since then, the SPF forecast has continued to drift down only very slowly, standing at 2.5% as of mid-2015. It is noteworthy that, as pointed out by Stanley Fischer in the speech quoted in the introduction, during that period both private and institutional forecasters systematically overestimated growth.

4.5 Implications for Nowcasting GDP

The standard DFM with constant long-run growth and constant volatility has been successfully applied to produce current quarter nowcasts of GDP (see Banbura et al., 2010, for a survey). Using our real-time US database, we carefully evaluate whether the introduction of time-varying long-run growth and SV into the DFM framework also improves the performance of the model along this dimension. We find that over the evaluation window 2000-2015 the model is at least as accurate at point forecasting, and significantly better at density forecasting than the benchmark DFM. We find that most of the improvement in density forecasting comes from correctly assessing the center and the right tail of the distribution, implying that the time-invariant DFM

[30]Indeed, the (one-sided) filtered estimate based on the latest vintage, which ignores the effect of data revisions, displays a more gradual pattern of decline (see Figure A.1 in Appendix A).

[31]A simpler specification that does not use consumption to inform the trend would detect the decline in long-run growth one year later, with additional revisions to past GDP in July 2011.

is assigning excessive probability to a strong recovery. In an evaluation sub-sample spanning the post-recession period, the relative performance of both point and density forecasts improves substantially, coinciding with the significant downward revision of the model's assessment of long-run growth. In fact, ignoring the variation in long-run GDP growth would have resulted in being on average around 1 percentage point too optimistic from 2009 to 2015.[32]

To sum up, the addition of the time-varying components not only provides a tool for decision-makers to update their knowledge about the state of long-run growth in real time. It also brings about a substantial improvement in short-run forecasting performance when the trend is shifting, without worsening the forecasts when the latter is relatively stable. The proposed model therefore provides a robust and timely methodology to track GDP when long-run growth is uncertain.

4.6 Inspecting the Role of Data Set Size and Composition

In this paper we argue that the rich multivariate framework of a DFM will facilitate the extraction of the long-run growth component of GDP. The DFM will exploit the cross-sectional dimension, and not just the time series dimension in separating cycle from trend. It is interesting to quantify the advantage that the DFM provides over traditional trend-cycle decompositions, and to investigate the robustness of our main conclusions to alternative datasets of varying size and composition. In order to do so, we consider (1) a bivariate model with GDP and unemployment only (labeled "Okun"), (2) an intermediate model with GDP and the four additional variables often included in the construction of coincident indicators, see Mariano and Murasawa (2003) and Stock and Watson (1989) (labeled "MM03"), (3) our "Baseline" specification with 28 variables, and (4) an "Extended" model that uses disaggregated data for many of the

[32] Appendix H provides the full details of the forecast evaluation exercise.

headline series included in the baseline specification, totaling 155 variables.[33] Moreover, in order to investigate the gains associated with imposing additional structure to long-run GDP growth, for the last two specifications we also consider a version of the model that does not impose common long-run growth in GDP and consumption.

The top panel of Table 2 reports the mean point-estimates for each of the specifications over selected subsamples of the data.[34] In all cases, the results are consistent with a decline in the long-run growth rate in the last part of the sample. Quantitatively, most specifications are very close to the baseline, with the specifications that impose common long-run growth in GDP and consumption finding an earlier and sharper decline. The exception is the "Okun" specification which instead estimates a smaller increase in the mid 1990s as well as a larger decline in long-run growth in the past decade. It is noteworthy that the mean estimate of the extended specification is very close to that of the baseline.

The lower panel of Table 2 instead investigates the uncertainty around the mean estimates. The uncertainty around the long-run growth estimate declines as we move from the bivariate to the multivariate specifications, with most of the reduction happening once a handful of variables are included. On the other hand, when the panel is extended to include a large number of disaggregated series, the uncertainty increases.[35]

[33] As we argue in Section 4.1, the introduction of a large number of disaggregated series, even if related to real activity, is likely to lead to model misspecification whenever the sectoral data are not contemporaneously related. For the extended specification, we consider a solution to this problem which allows to maintain the parsimonious one factor structure. By extending the model to include lags of the factor in the transition equation, each variable can display heterogeneous responses to the common factor, and correlation between idiosyncratic components is reduced. Given that the extended model is heavily parameterized, we follow D'Agostino et al. (2015) in choosing priors that shrink the model towards the contemporaneous-only specification, which is nested in the extended case. Full details and the composition of the data set and the changes to the estimation in case of the extended model are provided in Appendix J.

[34] See Figure J.1 in Appendix J for a comparison of the results of each alternative specification with the baseline results over the entire sample.

[35] The variance of the common factor declines further in the extended specification. However, as more disaggregated variables are added which the common factor must explain, more variability of GDP is left to be explained by the long-run component. As a result, the long-run component tends to overfit higher frequency movements in GDP and consumption.

While including a few key series, such as the ones in the specification of Mariano and Murasawa (2003) seems to already achieve the bulk of the reduction in uncertainty, it should be taken into account that those variables are available only with a relatively long publication lag, and subject to considerable revisions over time. Our proposed strategy of using an intermediate number of indicators, including the more timely and accurate surveys, is likely to lead to more satisfactory results in a real-time setting. Furthermore, the inclusion of the surveys is helpful in identifying the long-run growth rate, as those variables do not display a time-varying long-run mean by construction.

Overall this exercise highlights that the finding of a substantial decline in the long-run growth rate is confirmed across different specifications that use data sets of varying size and composition. The baseline specification, which uses an intermediate number of series including both hard data and surveys, leads to the lowest uncertainty around the long-run growth estimate, supporting the baseline choice of data set size and composition proposed in Section 4.1. Our results have important implications for trend-cycle decompositions of output, which usually include only a few cyclical indicators, generally inflation or variables that are direct inputs to the production function (see e.g. Gordon, 2014a or Reifschneider et al., 2013). As we show, greater precision of the trend component can be achieved by exploiting the common cyclical features of additional macroeconomic variables.[36]

[36] Basistha and Startz (2008) make a similar point, arguing that the inclusion of indicators that are informative about common cycles can help reduce the uncertainty around Kalman filter estimates of the long-run rate of unemployment (NAIRU).

Table 2: Comparison of results for alternative data sets and specifications

	Okun	*MM03*	Baseline *GDP only*	Baseline ***GDP + C***	Extended *GDP only*	Extended *GDP + C*
Long-run growth						
1947-1972	3.9	3.5	3.6	**3.9**	3.7	3.9
1973-1995	3.2	3.4	3.1	**3.2**	3.2	3.2
1996-2007	2.6	3.2	3.0	**3.1**	3.0	3.1
2008-2015	1.5	2.5	2.5	**1.8**	2.1	1.7
End of Sample	1.2	2.4	2.3	**2.0**	2.0	2.1
Uncertainty: Long run						
Filtered	0.84	0.62	0.63	**0.53**	0.70	0.62
Smoothed	0.44	0.36	0.35	**0.32**	0.41	0.39
Uncertainty: Cycle						
Filtered	2.15	1.48	0.79	**0.78**	0.23	0.18
Smoothed	1.98	1.32	0.62	**0.61**	0.25	0.17

Notes: Each column presents the estimation results corresponding to the alternative models (data sets) considered in this section. The upper panel displays the posterior means of the long-run growth rate of real GDP, over selected subsamples. In the lower panel, the posterior uncertainty corresponding to both the long-run growth rate of real GDP, as well as the cyclical factor are displayed.

5 Decomposing Movements in Long-Run Growth

In this section, we show how our model can be used to decompose the long-run growth rate of output into long-run movements in labor productivity and labor input. By doing this, we exploit the ability of the model to filter away cyclical variation and idiosyncratic noise and obtain clean estimates of underlying long-run trends. We see this exercise as a step towards giving an economically more meaningful interpretation to the movements in long-run real GDP growth detected by our model.

GDP growth is by identity the sum of growth in output per hour and growth in total hours worked. It is therefore possible to split the long-run growth trend in our model into two orthogonal components such that this identity is satisfied in the long run. Here we make use of our flexible definition of $\mathbf{c}_t$ in equation (2). In particular, ordering the growth rates of real GDP, real consumption and total hours as the first three variables in $\mathbf{y}_t$, we define

$$\mathbf{a}_t = \begin{bmatrix} z_t \\ h_t \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}, \tag{10}$$

so that the model is specified with two time-varying components, the first of which loads output and consumption but not hours, and the second loads all three series. The first component is then by construction the long-run growth rate of labor productivity, while the second one captures low-frequency movements in labor input independent of productivity.[37] Given the relation in (10), the two components add up to the time-varying intercept in the baseline specification, i.e. $g_t = z_t + h_t$.[38] It follows from standard growth theory that our estimate of the long-run growth rate of labor productivity will capture both technological factors and other factors, such as capital deepening and labor quality.[39]

Figure 4 presents the results of the decomposition exercise for the US. Panel (a) plots the median posterior estimate of long-run real GDP growth and its labor productivity and total hours components. The posterior bands for long-run real GDP growth are included. The time series evolution conforms very closely to the narrative of Fernald (2014), with a pronounced boom in labor productivity in the mid-1990's and a subsequent fall in the 2000's clearly visible. The decline in the 2000's is relatively

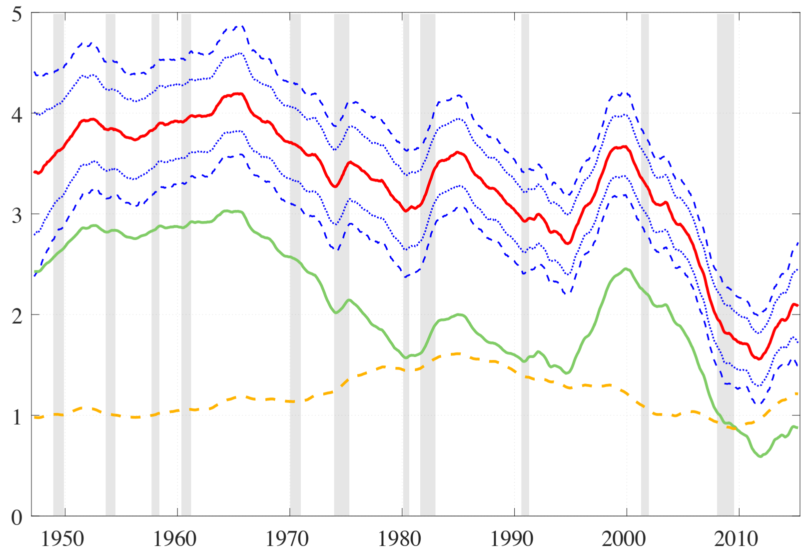
[37] z_t and h_t jointly follow random walks with diagonal covariance matrix as defined by equation (7). While the orthogonality assumption is not required for identification, imposing it allows us to interpret the innovations to the trends as exogenous shocks to the long-run growth rates of the variables. The hours trend is therefore interpreted as those low-frequency movements in hours which are uncorrelated with labor productivity. Allowing for a full covariance matrix leaves the results broadly unchanged.

[38] Since z_t and h_t are independent and add up to g_t, we set the prior on the scale of their variances to half of the one set in Section 4.2 on g_t. In addition, note that the cyclical movement in labor productivity is given by $(1 - \lambda_3)f_t$.

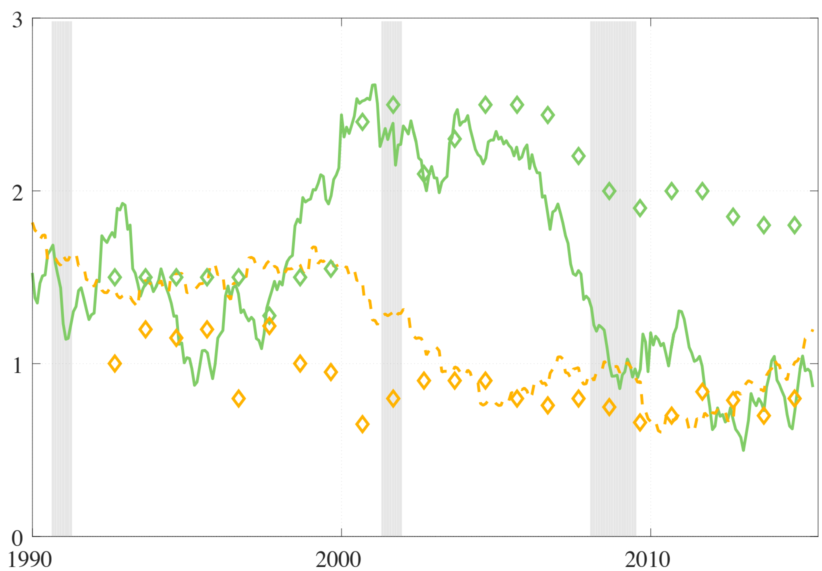
[39] Further decomposing z_t into technology and non-technology movements requires additional information to separately identify these components. One possibility, which we explore in Appendix K, is to use an independent measure of TFP to isolate technological factors. Note, however, that reliable data on capital input, labor quality, or estimates of TFP are not available in real time, making the focus on long-run labor productivity more appealing in a real-time setting.

Figure 4: Decomposition of Long-run US Output Growth

(a) Posterior median estimates of decomposition



(b) Filtered estimates of long-run growth components



Note: Panel (a) plots the posterior median (solid red), together with the 68% and 90% (doted and dashed blue) posterior credible intervals of long-run GDP growth and the posterior median of both long-run labor productivity growth and long-run total hours growth (solid green and dashed orange). Panel (b) plots the filtered estimates of these two components, i.e. $\hat{z}_{t|t}$ and $\hat{h}_{t|t}$, since 1990. For comparison, the corresponding forecasts from the SPF are plotted. The SPF forecast for total hours is obtained as the difference between the forecasts for real GDP and labor productivity.

sudden while the 1970's slowdown appears as a more gradual phenomenon starting in the late 1960's. Furthermore, the results reveal that during the 1970's and 1980's the impact of the productivity slowdown on output growth was partly masked by a secular increase in hours, probably reflecting increases in the working-age population as well as labor force participation (see e.g. Goldin, 2006). Focusing on the period since 2000, labor productivity accounts for almost the entire decline.[40] This contrasts explanations by which slow labor force growth has been a drag on GDP growth. When taking away the cyclical component of hours and focusing solely on its long-run component, the contribution of hours has, if anything, accelerated since the Great Recession. Panel (b) presents the filtered estimates of the two components, i.e. the output of the Kalman Filter which uses data only up to each point in time. For comparison, the corresponding SPF forecasts are included. Most notably, this plot reveals that starting around 2005 a relatively sharp revision to labor productivity drives the decline in long-run output growth.[41] Interestingly, the professional forecasters have been very slow in incorporating the productivity slowdown into their long-run forecasts. This delay explains their persistent overestimation of GDP growth since the recession.

It is interesting to compare the results of our decomposition exercise to similar approaches in the recent literature, in particular Gordon (2010, 2014a) and Reifschneider et al. (2013). Like us, they specify a state space model with a common cyclical component and use the 'output identity' to decompose the long-run growth rate of GDP into underlying drivers. A key difference resides in the Bayesian estimation of the model, which enables us to impose a conservative prior on the variance of the long-run growth

[40] In Appendix K, where we extend this analysis to further decompose the labor productivity trend into long-run TFP and non-technological forces, we find that movements in TFP account for virtually all of the slowdown.

[41] In an additional figure, provided in Appendix A, we plot 5,000 draws from the joint posterior distribution of the variances of the innovations to the labor productivity and hours components. This analysis confirms the conclusion from the discussion here that changes in labor productivity, rather than in labor input, are the key driver of low frequency movements in real GDP growth.

component that helps avoiding over-fitting the data. Furthermore, the inclusion of SV in the cyclical component helps to prevent unusually large cyclical movements from contaminating the long-run estimate. Another important difference is that we use a larger amount of information, including key cyclical indicators like industrial production, sales, and business surveys, which are generally not included in a production function approach. This allows us to retrieve a timely and precise estimate of the cyclical component and, as a consequence, to reduce the uncertainty that is inherent to any trend-cycle decomposition of the data, as discussed in Section 4.6. As a result, we obtain a substantially less pessimistic estimate of the long-run growth of GDP than these studies in the latest part of the sample. For instance, Gordon (2014a) reports a long-run GDP growth estimate below 1% for the end of the sample, whereas our median estimate stands at around 2%.[42]

5.1 International Evidence

To gain an international perspective on our results, we estimate the DFM for the other G7 economies and perform the decomposition exercise for each of them.[43] The median posterior estimates of the labor productivity and labor input trends are displayed in Figure 5. Labor productivity, displayed in Panel (a), plays again the key role in determining movements in long-run growth. In the Western European economies and Japan, the elevated growth rates of labor productivity prior to the 1970's reflect the

[42]The results for a bivariate model of GDP and unemployment, which we have discussed in Section 4.6 show that the current long-run growth estimate is 1.2%, close to Gordon (2014a).

[43]Details on the specific data series used for each country are available in Appendix E. For hours, we again follow the methodology of Ohanian and Raffo (2012). In the particular case of the UK, the quarterly series for hours displays a drastic change in its stochastic properties in the early 1990's owing to a methodological change in the construction by the ONS, as confirmed by the ONS LFS manual. We address this issue by using directly the annual series from the TED, which requires an appropriate extension of equation (9) to annual variables (see Banbura et al. 2012). To avoid weak identification of h_t for the UK, we truncate our prior on its variance to discard values which are larger than twice the maximum posterior draw of the case of the other countries.

rebuilding of the capital stock from the destruction from World War II, and ended as these economies converged towards US levels of output per capita. The labor productivity profile of Canada broadly follows that of the US, with a slowdown in the 1970's and a temporary mild boom during the late 1990's. Interestingly, this acceleration in the 1990's did not occur in Western Europe and Japan.[44] The UK displays a decline in labor productivity similar to the US. This "productivity puzzle" has been debated extensively in the UK (see e.g. Pessoa and Van Reenen, 2014). It is interesting to note that the two countries which experienced a more severe financial crisis, the US and the UK, appear to be the ones with greatest declines in productivity since the early 2000's, similar to the evidence documented in Reinhart and Rogoff (2009).

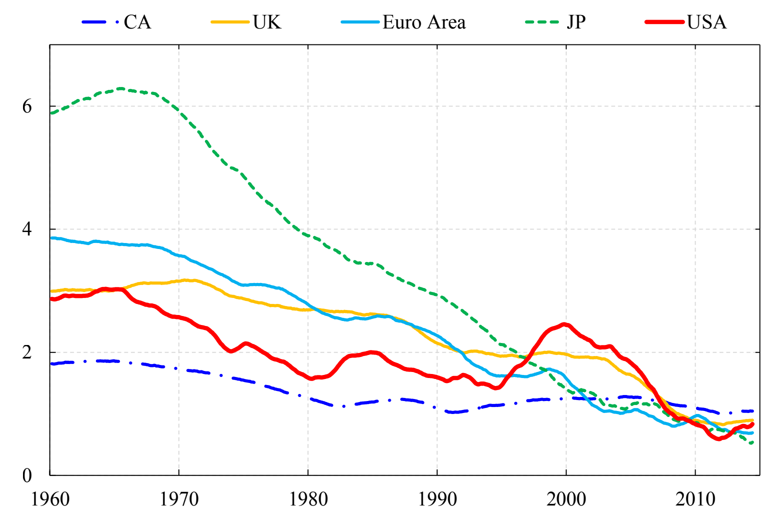
Panel (b) displays the movements in long-run hours worked identified by equation (10). The contribution of this component to overall long-run output growth varies considerably across countries. However, within each country it is more stable over time than the productivity component, which is in line with our findings for the US. Indeed, the extracted long-run trend in total hours includes various potentially offsetting forces that can lead to changes in long-run output growth. In any case, the results of our decomposition exercise indicate that after using the DFM to remove business-cycle variation in hours and output, the decline in long-run GDP growth that has been observed in the advanced economies since the early 2000's is entirely accounted for by a decline in the labor productivity trend. Finally, it is interesting to note that for the countries in the sample long-run productivity growth appears to converge in the cross section, while there is no evidence of convergence in the long-run growth of hours.[45]

[44]On the lost decade in Japan, see Hayashi and Prescott (2002). Gordon (2004) examines the absence of the IT boom in Europe.

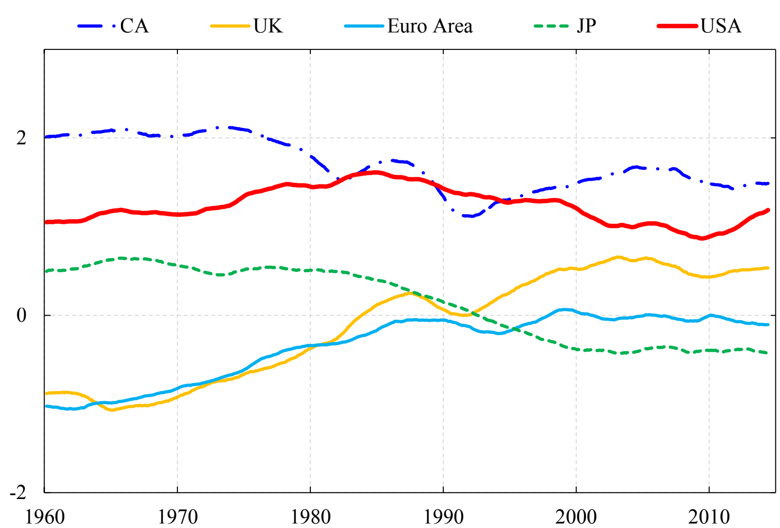
[45]Similar evidence for emerging economies has been recently presented by Pritchett and Summers (2014). Their evidence refers to convergence of overall GDP growth rates, whereas ours indicates that convergence in productivity growth appears to be the dominant source of convergence.

Figure 5: Decomposition for Other Advanced Economies

(a) Long-run Labor Productivity



(b) Long-run Labor Input



Note: Panel (a) displays the posterior median of long-run labor productivity across advanced economies. Panel (b) plots the corresponding estimates of long-run total hours worked. In both panels, 'Euro Area' represents a weighted average of Germany, Italy and France.

6 Concluding Remarks

The sluggish recovery from the Great Recession has raised the question whether the long-run growth rate of US real GDP is now lower than it has been on average over the postwar period. We have presented a dynamic factor model that allows for both changes in long-run GDP growth and stochastic volatility. Estimating the model with Bayesian methods, we provide evidence that long-run growth of US GDP displays a gradual decline after the turn of the century, moving from its peak of 3.5% to about 2% in 2015. Using real-time vintages of data we demonstrate the model's ability to track GDP in a timely and reliable manner. By the summer of 2010 the model would have concluded that a significant decline in long-run growth was behind the slow recovery, therefore substantially improving the real-time tracking of GDP by explicitly taking into account the uncertainty surrounding long-run growth. Finally, we discuss the drivers of movements in long-run output growth through the lens of our model by decomposing it into the long-run growth rates of labor productivity and labor input. Using data for both the US and other advanced economies our model points to a global slowdown in labor productivity as the main driver of weak growth in recent years, extending the narrative of Fernald (2014) to other economies. Studying the deep causes of the secular decline in growth is an important priority for future research.

References

Alvarez, R., Camacho, M., and Perez-Quiros, G. (2012). Finite sample performance of small versus large scale dynamic factor models. CEPR Discussion Papers 8867, C.E.P.R. Discussion Papers.

Auerbach, A. J. (2011). Long-term fiscal sustainability in major economies. BIS Working Papers 361, Bank for International Settlements.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–68.

Bai, J. and Wang, P. (2015). Identification and Bayesian Estimation of Dynamic Factor Models. *Journal of Business & Economic Statistics*, 33(2):221–240.

Banbura, M., Giannone, D., Modugno, M., and Reichlin, L. (2012). Now-casting and the real-time data flow. Working Papers ECARES 2012-026, ULB.

Banbura, M., Giannone, D., and Reichlin, L. (2010). Nowcasting. CEPR Discussion Papers 7883, C.E.P.R. Discussion Papers.

Banbura, M. and Modugno, M. (2014). Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, 29(1):133–160.

Basistha, A. and Startz, R. (2008). Measuring the NAIRU with Reduced Uncertainty: A Multiple-Indicator Common-Cycle Approach. *The Review of Economics and Statistics*, 90(4):805–811.

Benati, L. (2007). Drift and breaks in labor productivity. *Journal of Economic Dynamics and Control*, 31(8):2847–2877.

Beveridge, S. and Nelson, C. R. (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle. *Journal of Monetary Economics*, 7(2):151–174.

Bloom, N. (2014). Fluctuations in Uncertainty. *Journal of Economic Perspectives*, 28(2):153–76.

Boivin, J. and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132(1):169–194.

Clark, P. K. (1987). The Cyclical Component of U.S. Economic Activity. *The Quarterly Journal of Economics*, 102(4):797–814.

Cochrane, J. H. (1994). Permanent and Transitory Components of GNP and Stock Prices. *The Quarterly Journal of Economics*, 109(1):241–65.

Cogley, T. (2005). How fast can the new economy grow? A Bayesian analysis of the evolution of trend growth. *Journal of Macroeconomics*, 27(2):179–207.

Cogley, T. and Sargent, T. J. (2005). Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics*, 8(2):262–302.

Creal, D., Koopman, S. J., and Zivot, E. (2010). Extracting a robust US business cycle using a time-varying multivariate model-based bandpass filter. *Journal of Applied Econometrics*, 25(4):695–719.

D'Agostino, A., Giannone, D., Lenza, M., and Modugno, M. (2015). Nowcasting business cycles: A bayesian approach to dynamic heterogeneous factor models. Technical report.

Del Negro, M. and Otrok, C. (2008). Dynamic factor models with time-varying parameters:measuring changes in international business cycles. NYFed Staff Rep. 326.

Durbin, J. and Koopman, S. J. (2012). *Time Series Analysis by State Space Methods: Second Edition*. Oxford University Press.

Edge, R. M., Laubach, T., and Williams, J. C. (2007). Learning and shifts in long-run productivity growth. *Journal of Monetary Economics*, 54(8):2421–2438.

Evans, M. D. D. (2005). Where Are We Now? Real-Time Estimates of the Macroeconomy. *International Journal of Central Banking*, 1(2).

Fernald, J. (2014). Productivity and potential output before, during, and after the great recession. *NBER Macroeconomics Annual 2014*, 29.

Fernald, J. G. and Jones, C. I. (2014). The Future of US Economic Growth. *American Economic Review*, 104(5):44–49.

Fleischman, C. A. and Roberts, J. M. (2011). From many series, one cycle: improved estimates of the business cycle from a multivariate unobserved components model. Finance and Economics Discussion Series 2011-46, Fed Board of Governors.

Forni, M., Giannone, D., Lippi, M., and Reichlin, L. (2009). Opening The Black Box: Structural Factor Models With Large Cross Sections. *Econometric Theory*, 25(05):1319–1347.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2003). Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, 50(6):1243–1255.

Gadea-Rivas, M. D., Gomez-Loscos, A., and Perez-Quiros, G. (2014). The Two Greatest. Great Recession vs. Great Moderation. WP Series 1423, Bank of Spain.

Geweke, J. (1977). The dynamic factor analysis of economic time series. In *Latent Variables in Socio-Economic Models*. North-Holland.

Giannone, D., Reichlin, L., and Sala, L. (2006). Vars, common factors and the empirical validation of equilibrium business cycle models. *Journal of Econometrics*, 132(1):257–279.

Giannone, D., Reichlin, L., and Small, D. (2005). Nowcasting gdp and inflation: the real-time informational content of macroeconomic data releases. Technical report.

Giannone, D., Reichlin, L., and Small, D. (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics*, 55(4):665–676.

Goldin, C. (2006). The quiet revolution that transformed women's employment, education, and family. *AEA Papers and Proceedings*, May 2006.

Gordon, R. J. (2004). Why was europe left at the station when america's productivity locomotive departed? Working Paper 10661, National Bureau of Economic Research.

Gordon, R. J. (2010). Okun's Law and Productivity Innovations. *American Economic Review*, 100(2):11–15.

Gordon, R. J. (2014a). A new method of estimating potential real gdp growth: Implications for the labor market and the debt/gdp ratio. NBER Working Paper 20423.

Gordon, R. J. (2014b). The Demise of U.S. Economic Growth: Restatement, Rebuttal, and Reflections. NBER Working Papers 19895.

Hansen, B. E. (1992). Testing for parameter instability in linear models. *Journal of Policy Modeling*, 14(4):517–533.

Harvey, A. C. (1985). Trends and Cycles in Macroeconomic Time Series. *Journal of Business & Economic Statistics*, 3(3):216–27.

Harvey, A. C. and Stock, J. H. (1988). Continuous time autoregressive models with common stochastic trends. *Journal of Economic Dynamics and Control*, 12(2-3):365–384.

Hayashi, F. and Prescott, E. C. (2002). The 1990s in japan: A lost decade. *Review of Economic Dynamics*, 5(1):206 – 235.

Hodrick, R. J. and Prescott, E. C. (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, 29(1):1–16.

Jacquier, E., Polson, N. G., and Rossi, P. E. (2002). Bayesian Analysis of Stochastic Volatility Models. *Journal of Business & Economic Statistics*, 20(1):69–87.

Jorgenson, D. W., Ho, M. S., and Stiroh, K. J. (2006). Potential Growth of the U.S. Economy: Will the Productivity Resurgence Continue? *Business Economics*, 41(1):7–16.

Jurado, K., Ludvigson, S. C., and Ng, S. (2014). Measuring uncertainty. *American Economic Review, Forthcoming*.

Kim, C.-J. and Nelson, C. R. (1999a). Has The U.S. Economy Become More Stable? A Bayesian Approach Based On A Markov-Switching Model Of The Business Cycle. *The Review of Economics and Statistics*, 81(4):608–616.

Kim, C.-J. and Nelson, C. R. (1999b). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. The MIT Press.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies*, 65(3):361–93.

Luciani, M. and Ricci, L. (2014). Nowcasting Norway. *International Journal of Central Banking, Forthcoming*.

Luo, S. and Startz, R. (2014). Is it one break or on going permanent shocks that explains u.s. real gdp? *Journal of Monetary Economics*, 66:155–163.

Marcellino, M., Porqueddu, M., and Venditti, F. (2014). Short-term gdp forecasting with a mixed frequency dynamic factor model with stochastic volatility. *Journal of Business & Economic Statistics, Forthcoming.*

Mariano, R. S. and Murasawa, Y. (2003). A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics*, 18(4):427–443.

McConnell, M. M. and Perez-Quiros, G. (2000). Output fluctuations in the united states: What has changed since the early 1980's? *American Economic Review*, 90(5):1464–1476.

Morley, J. C., Nelson, C., and Zivot, E. (2003). Why are the beveridge-nelson and unobserved-components decompositions of gdp so different? *The Review of Economics and Statistics*, 85(2):235–243.

Nelson, C. R. and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, 10(2):139–162.

Nordhaus, W. (2004). Retrospective on the 1970s Productivity Slowdown. Nber working papers, National Bureau of Economic Research, Inc.

Nyblom, J. (1989). Testing for the constancy of parameters over time. *Journal of the American Statistical Association*, 84(405):223–230.

Ohanian, L. E. and Raffo, A. (2012). Aggregate hours worked in OECD countries: New measurement and implications for business cycles. *Journal of Monetary Economics*, 59(1):40–56.

Oliner, S. D. and Sichel, D. E. (2000). The Resurgence of Growth in the Late 1990s: Is Information Technology the Story? *Journal of Economic Perspectives*, 14(4):3–22.

Orphanides, A. (2003). The quest for prosperity without inflation. *Journal of Monetary Economics*, 50(3):633–663.

Pain, N., Lewis, C., Dang, T.-T., Jin, Y., and Richardson, P. (2014). OECD Forecasts During and After the Financial Crisis: A Post Mortem. OECD Economics Department Working Papers 1107, OECD Publishing.

Perron, P. and Wada, T. (2009). Let's take a break: trends and cycles in us real gdp. *Journal of Monetary Economics*, 56:749–765.

Pessoa, J. P. and Van Reenen, J. (2014). The uk productivity and jobs puzzle: Does the answer lie in wage flexibility? *The Economic Journal*, 124(576):433–452.

Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies*, 72(3):821–852.

Pritchett, L. and Summers, L. H. (2014). Asiaphoria Meets Regression to the Mean. NBER Working Papers 20573, National Bureau of Economic Research, Inc.

Reifschneider, D., Wascher, W., and Wilcox, D. (2013). Aggregate supply in the United States: recent developments and implications for the conduct of monetary policy. Finance and Economics Discussion Series 2013-77, Board of Governors of the Federal Reserve System.

Reinhart, C. M. and Rogoff, K. S. (2009). The Aftermath of Financial Crises. *American Economic Review*, 99(2):466–72.

Sargent, T. J. and Sims, C. A. (1977). Business cycle modeling without pretending to have too much a priori economic theory. Working Papers 55, Federal Reserve Bank of Minneapolis.

Sims, C. A. (2012). Comment to stock and watson (2012). *Brookings Papers on Economic Activity*, Spring:81–156.

Stock, J. and Watson, M. (2012). Disentangling the channels of the 2007-09 recession. *Brookings Papers on Economic Activity*, Spring:81–156.

Stock, J. H. and Watson, M. W. (1989). New Indexes of Coincident and Leading Economic Indicators. In *NBER Macroeconomics Annual 1989, Volume 4*, NBER Chapters, pages 351–409. National Bureau of Economic Research, Inc.

Stock, J. H. and Watson, M. W. (1999). Business cycle fluctuations in us macroeconomic time series. In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics*, volume 1 of *Handbook of Macroeconomics*, chapter 1, pages 3–64. Elsevier.

Stock, J. H. and Watson, M. W. (2002). Forecasting Using Principal Components From a Large Number of Predictors. *Journal of the American Statistical Association*, 97:1167–1179.

Stock, J. H. and Watson, M. W. (2003). Has the business cycle changed and why? *NBER Macroeconomics Annual 2002*, 17:159–230.

Summers, L. (2014). Secular stagnation. IMF Economic Forum: Policy Responses to Crises, speech delivered at the IMF Annual Research Conference, November 8th.

Whelan, K. (2003). A Two-Sector Approach to Modeling U.S. NIPA Data. *Journal of Money, Credit and Banking*, 35(4):627–56.

Comments invited

THE WORLDWIDE EQUITY PREMIUM: A SMALLER PUZZLE

Elroy Dimson, Paul Marsh, and Mike Staunton*

London Business School

Revised 7 April 2006

Abstract: We use a new database of long-run stock, bond, bill, inflation, and currency returns to estimate the equity risk premium for 17 countries and a world index over a 106-year interval. Taking U.S. Treasury bills (government bonds) as the risk-free asset, the annualised equity premium for the world index was 4.7% (4.0%). We report the historical equity premium for each market in local currency and US dollars, and decompose the premium into dividend growth, multiple expansion, the dividend yield, and changes in the real exchange rate. We infer that investors expect a premium on the world index of around 3–3½% on a geometric mean basis, or approximately 4½–5% on an arithmetic basis.

JEL classifications: G12, G15, G23, G31, N20.

Keywords: Equity risk premium; long run returns; survivor bias; financial history; stocks, bonds, bills, inflation.

---

* London Business School, Regents Park, London NW1 4SA, United Kingdom. Tel: +44 (0)20 7262 5050. Email: edimson@london.edu, pmarsh@london.edu, and mstaunton@london.edu. We are grateful to Rajnish Mehra and an anonymous referee, participants at over 40 seminars, and the 37 individuals who contributed the datasets described in Appendix 2.

THE WORLDWIDE EQUITY PREMIUM: A SMALLER PUZZLE

Abstract: We use a new database of long-run stock, bond, bill, inflation, and currency returns to estimate the equity risk premium for 17 countries and a world index over a 106-year interval. Taking U.S. Treasury bills (government bonds) as the risk-free asset, the annualised equity premium for the world index was 4.7% (4.0%). We report the historical equity premium for each market in local currency and US dollars, and decompose the premium into dividend growth, multiple expansion, the dividend yield, and changes in the real exchange rate. We infer that investors expect a premium on the world index of around 3–3½% on a geometric mean basis, or approximately 4½–5% on an arithmetic basis.

In their seminal paper on the equity premium puzzle, Mehra and Prescott (1985) showed that the historical equity premium in the United States—measured as the excess return on stocks relative to the return on relatively risk-free Treasury bills—was much larger than could be justified as a risk premium on the basis of standard theory. Using the accepted neoclassical paradigms of financial economics, combined with estimates of the mean, variance and auto-correlation of annual consumption growth in the U.S. economy and plausible estimates of the coefficient of risk aversion and time preference, they argued that stocks should provide at most a 0.35% annual risk premium over bills. Even by stretching the parameter estimates, they concluded that the premium should be no more than 1% (Mehra and Prescott (2003)). This contrasted starkly with their historical mean annual equity premium estimate of 6.2%.

The equity premium puzzle is thus a quantitative puzzle about the magnitude, rather than the sign, of the risk premium. Ironically, since Mehra and Prescott wrote their paper, this puzzle has grown yet more quantitatively puzzling. Over the 27 years from the end of the period they examined to the date of completing this contribution, namely over 1979–2005, the mean annual U.S. equity premium relative to bills using Mehra-Prescott's definition and data sources was 8.1%.

Logically, there are two possible resolutions to the puzzle: either the standard models are wrong, or else the historical premium is misleading and we should expect a lower premium in the future. Over the last two decades, researchers have tried to resolve the puzzle by generalising and adapting the Mehra-Prescott (1985) model. Their efforts have focused on alternative assumptions about preferences, including risk aversion, state separability, leisure, habit formation and precautionary saving; incomplete markets and uninsurable income shocks; modified probability distributions to admit rare, disastrous events; market imperfections, such as borrowing constraints and transactions costs; models of limited participation of consumers in the stock market, and behavioural explanations. There are several excellent surveys of this work, including Kocherlakota (1996), Cochrane (1997), Mehra and Prescott (2003), and most recently, Mehra and Prescott (2006).

While some of these models have the potential to resolve the puzzle, as Cochrane (1997) points out, the most promising of them involve "deep modifications to the standard models" and "every quantitatively successful current story…still requires astonishingly high risk aversion". This leads us back to the second possible resolution to the puzzle, namely, that the historical premium may be misleading. Perhaps U.S. equity investors simply enjoyed good fortune and the twentieth century for them represented the "triumph of the optimists" (Dimson, Marsh, and Staunton (2002)). As Cochrane (1997) puts it, maybe it was simply "100 years of good luck"—the opposite of the old joke about Soviet agriculture being the result of "100 years of bad luck."

This good luck story may also be accentuated by country selection bias, making the historical data even more misleading. To illustrate this, consider the parallel with selection bias in the choice of stocks, and the task facing a researcher who wished to estimate the required risk premium and expected return on the common stock of Microsoft. It would be foolish to extrapolate from Microsoft's stellar past performance. Its success and survival makes it non-typical of companies as a whole. Moreover, in its core business Microsoft has a market share above 50%. Since, by definition, no competitor can equal this accomplishment, we should not extrapolate expected returns from this one example of success. The past performance of individual stocks is anyway largely uninformative about their future returns, but when there is *ex post* selection bias based on past success, historical mean returns will provide an upward biased estimate of future expected returns. That is one reason why equity premium projections are usually based on the performance of the entire market, including unsuccessful as well as successful stocks.[1]

For similar reasons, we should also be uncomfortable about extrapolating from a stock market that has survived and been successful, and gained a market share of above 50%. Organized trading in marketable securities began in Amsterdam in 1602 and London in 1698, but did not commence in New York until 1792. Since then, the U.S. share of the global stock market as measured by the percentage of overall world equity market capitalization has risen from zero to around 50% (see Dimson, Marsh, and Staunton (2004)). This reflects the superior performance of the U.S. economy, as evidenced by a large volume of initial public offerings (IPOs) and seasoned equity offerings (SEOs) that enlarged the U.S. equity market, and the substantial returns from U.S. common stocks after they had gained a listing. No other market can rival this long-term accomplishment.

Mehra and Prescott's initial focus on the United States and the ready availability of U.S. data has ensured that much of the subsequent research prompted by their paper has investigated the premium within the context of the U.S. market. The theoretical work usually starts with the assumption that the equity premium is of the magnitude that has been observed historically in the United States, and seeks to show why the Mehra-Prescott observations are not (quite so much of) a puzzle. Some empirical work has looked beyond the United States, including Jorion and Goetzmann (1999) and Mehra and Prescott (2003). However, researchers have hitherto been hampered by the paucity of long-run equity returns data for other countries. Most research seeking to resolve the equity premium puzzle has thus focused on empirical evidence for the United States. In emphasizing the U.S.—a country that must be a relative outlier—this body of work may be starting from the wrong set of beliefs about the past.

The historically measured equity premium could also be misleading if the risk premium has been non-stationary. This could have arisen if, over the measurement interval, there have been changes in risk, or the risk attitude of investors, or investors' diversification opportunities. If, for example, these have caused a reduction in the risk premium, this fall in the discount rate will

[1] Another key reason is that equilibrium asset pricing theories such as the CAPM or CCAPM assign a special role to the value weighted market portfolio. However, our argument for looking beyond the United States is not dependent on the assumption that the market portfolio should necessarily be the world portfolio. Instead, we are simply pointing out that if one selects a country which is known after the event to have been unusually successful, then its past equity returns are likely to be an upward biased estimate of future returns.

have led to re-pricing of stocks, thus adding to the magnitude of historical returns. The historical mean equity premium will then overstate the prospective risk premium, not only because the premium has fallen over time, but also because historical returns are inflated by past repricings that were triggered by a *reduction* in the risk premium.

In this paper, we therefore revisit two fundamental questions: How large has the equity premium been historically, and how big is it likely to be in the future? To answer these questions, we extend our horizon beyond just the United States and use a new source of long-run returns, the Dimson-Marsh-Staunton (2006) database, to examine capital market history in 17 countries over the 106-year period from 1900 to 2005. Initially, we use the DMS database to estimate the historical equity premium around the world on the assumption that the premium was stationary. We then analyse the components of the premium to provide insights into the impact on historical returns of (i) luck and (ii) repricing resulting from changes in the underlying risk premium. This then enables us to make inferences about the likely future long-run premium.

Our paper is organized as follows. The next section reviews previous estimates and beliefs about the size of the equity premium. Section 3 describes the new DMS global database and explains why it represents a significant advance over previous data. Section 4 utilizes the database to present summary data on long-run returns, and to illustrate why we need long-run histories to estimate premiums with any precision—even if the underlying processes are non-stationary. Section 5 presents new evidence on the historical equity premium around the world, assuming stationarity. Section 6 decomposes historical equity premiums into several elements, documenting the contribution of each to historical returns. Section 7 uses this decomposition to infer expectations of the equity premium, discusses why these are lower than the historical realizations, and provides a summary and conclusion. There are two appendices, one formalising the methodology behind our decomposition, and the other documenting our data sources.

2. PRIOR ESTIMATES OF THE EQUITY PREMIUM

Prior estimates of the historical equity premium draw heavily on the United States, with most researchers and textbooks citing just the American experience. The most widely cited source is Ibbotson Associates whose U.S. database starts in 1926. At the turn of the millennium, Ibbotson's estimate of the U.S. arithmetic mean equity premium from 1926–1999 was 9.2%. In addition, before the DMS database became available, researchers such as Mehra and Prescott (2003), Siegel (2002), and Jorion and Goetzmann (1999) used the Barclays Capital (1999) and Credit Suisse First Boston (CSFB) (1999) data for the United Kingdom. In 1999, both Barclays and CSFB were using identical U.K. equity and Treasury bill indexes that started in 1919 and gave rise to an arithmetic mean equity premium of 8.8%.

In recent years, a growing appreciation of the equity premium puzzle made academics and practitioners increasingly concerned that these widely cited estimates were too high. This distrust proved justified for the historical numbers for the U.K., which were wrong. The former Barclays/CSFB index was retrospectively constructed, and from 1919–35, was based on a sample of 30 stocks chosen from the largest companies (and sectors) in 1935. As we show in Dimson, Marsh and Staunton (2001), the index thereby suffered from *ex post* bias. It represented a potential investment strategy only for investors with perfect foresight in 1919 about which companies were destined to survive (survivorship bias). Even more seriously, it incorporated hindsight on which stocks and sectors were destined in 1919 subsequently to perform well and grow large (success bias).[2]

After correcting for this *ex post* selection bias, the arithmetic mean equity premium from 1919–35 fell from 10.6% to 5.2%. The returns on this index were also flattered by the choice of start-date. By starting in 1919, it captured the post-World War I recovery, while omitting wartime losses and the lower pre-war returns. Adding in these earlier years gave an arithmetic mean U.K. equity premium over the entire twentieth century of 6.6%, some 2¼% lower than might have been inferred from the earlier, incorrect data for 1919–99.

The data used by Ibbotson Associates to compute the historical U.S. equity premium is of higher quality and does not suffer from the problems that afflicted the old U.K. indexes. Those believing that the premium is "too good to be true" have therefore pointed their finger of suspicion mainly at success bias—a choice of market that was influenced by that country's record of success. Bodie (2002) argued that high U.S. and U.K. premiums are likely to be anomalous, and underlined the need for comparative international evidence. He pointed out that long-run studies are almost always of U.S. or U.K. premiums: "There were 36 active stock markets in 1900, so why do we only look at two? I can tell you—because many of the others don't have a 100-year history, for a variety of reasons."

There are indeed relatively few studies extending beyond the United States and the United Kingdom. Mehra and Prescott (2003) report comparative premiums for France, Japan, and Germany. They find a similar pattern to the United States, but their premiums are based on post-1970 data and periods of 30 years or less. Ibbotson Associates (2005) compute equity premiums for 16 countries, but only from 1970. Siegel (2002) reports premiums for Germany and Japan since 1926, finding magnitudes similar to those in the United States. Jorion and Goetzmann (1999) provide the most comprehensive long-run global study by assembling a database of capital gain indexes for 39 markets, 11 of which started as early as 1921. However, they were able to identify only four markets, apart from the United States and the United Kingdom, with pre-1970 dividend information. They concluded that, "the high equity premium obtained for U.S. equities appears to be the exception rather than the rule." But in the absence of reliable dividend information, this assertion must be treated with caution. We therefore return to this question using comprehensive total returns data in section 5 below.

Expert Opinion

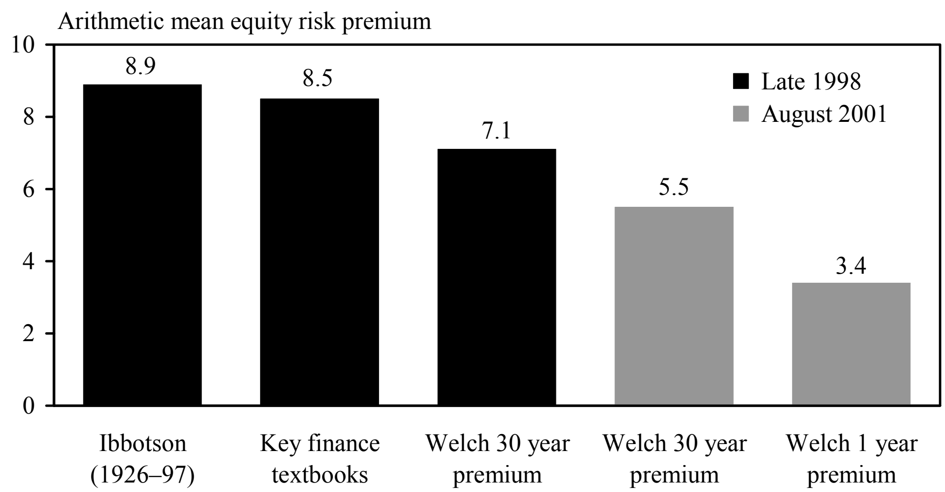
The equity premium has thus been a source of controversy, even among experts. Welch (2000) studied the opinions of 226 financial economists who were asked to forecast the average annual equity premium over the next 30 years. Their forecasts ranged from 1% to 15%, with a mean and median of 7%. No clear consensus emerged: the cross-sectional dispersion of the forecasts was as large as the standard error of the mean historical equity premium.

---

[2] After becoming aware of our research, Barclays Capital (but not CSFB) corrected their pre-1955 estimates of U.K. equity returns for bias and extended their index series back to 1900.

Most respondents to the Welch survey would have viewed the Ibbotson Associates Yearbook as the definitive study of the historical U.S. equity premium. At that time, the most recent Yearbook was the 1998 edition, covering 1926–1997. The first bar of **Figure 1** shows that the arithmetic mean equity premium based on the Yearbook data was 8.9% per annum.[3] The second bar shows that the key finance textbooks were on average suggesting a slightly lower premium of 8.5%. This may have been based on earlier, slightly lower, Ibbotson estimates, or perhaps the authors were shading the estimates down. The Welch survey mean is in turn lower than the textbook figures, but since the respondents claimed to lower their forecasts when the equity market rises, this may reflect the market's strong performance in the 1990s.

Figure 1: Estimated Arithmetic Equity Premiums Relative to Bills, 1998 and 2001



At the time of this survey, academics' forecasts of the long-run premium thus seemed strongly influenced by the historical record. Certainly, leading textbooks advocated the use of the historical mean, including Bodie, Kane, and Marcus (1999) and Brealey and Myers (2000). The latter states, "Many financial managers and economists believe that long-run historical returns are the best measure available." This was supported by researchers such as Goyal and Welch (2006) who could not identify a single predictive variable that would have been of robust use for forecasting the equity premium, and recommended "assuming that the equity premium is 'like it always has been'." Even Mehra and Prescott (2003) state, "…over the long horizon the equity premium is likely to be similar to what it has been in the past and the returns to investment in equity will continue to dominate that in T-bills for investors with a long planning horizon."

The survey and textbook figures shown in the second and third bars of Figure 1 indicate what was being taught at the end of the 1990s in the world's top business schools and economics departments. But by 2001, longer-term estimates were gaining publicity. Our own estimate (Dimson, Marsh, and Staunton (2000)) of the U.S. arithmetic mean premium over the entire twentieth century of 7.7% was 1.2% lower than Ibbotson's estimate of 8.9% for 1926–1997.

[3] This is the arithmetic mean of the one-year geometric risk premiums. The arithmetic mean of the one-year arithmetic risk premiums, i.e., the average annual difference between the equity return and the Treasury bill return, was slightly higher at 9.1%.

In August 2001, Welch (2001) updated his survey, receiving 510 responses. Respondents had revised their estimates downward by an average of 1.6%. They now estimated an equity premium averaging 5.5% over a 30-year horizon, and 3.4% over a one-year horizon (see Figure 1). Those taking part for the first time estimated the same mean premiums as those who had participated in the earlier survey. While respondents to the earlier survey had indicated that, on average, a bear market would raise their equity premium forecast, Welch reports that "this is in contrast with the observed findings: it appears as if the recent bear market correlates with lower equity premium forecasts, not higher equity premium forecasts."

The academic consensus now appears to be lower still (e.g., see Jagannathan, McGrattan and Scherbina (2000) and Fama and French (2002)). Investment practitioners typically agree (see Arnott and Ryan (2001) and Arnott and Bernstein (2002), and the latest editions of many textbooks have reduced their equity premium estimates (for a summary of textbook prescriptions, see Fernandez (2004)). Meanwhile, surveys by Graham and Harvey (2005) indicate that U.S. CFOs have reduced their forecasts of the equity premium from 4.65% in September 2000 to 2.93% by September 2005. Yet predictions of the long-term premium should not be so sensitive to short-term market fluctuations. Over this period, the long-run historical mean premium—which just a few years earlier had been the anchor of beliefs—has fallen only modestly, as adding in the years 2000–05 reduces the long-run mean by just 0.4%, despite the bear market of 2000–02. The sharp lowering of the consensus view about the future premium must therefore reflect more than this, such as new ways of interpreting the past, new approaches to forecasting the premium, or new facts about global long-term performance, such as evidence that the U.S. premium was higher than in most other countries.

3. LONG-RUN INTERNATIONAL DATA

We have seen that previous research has been hampered by the quality and availability of long-run global data. The main problems were the short time-series available and hence the focus on recent data, the absence of dividends, *ex post* selection bias, and emphasizing data that is "easy" to access.

Historically, the most widely used database for international stock market research has been the Morgan Stanley Capital International (MSCI) index series, but the MSCI data files start only in 1970. This provides a rather short history for estimating equity premiums, and spans a period when equities mostly performed well, so premiums inevitably appear large. Researchers interested in longer-term data have found no shortage of earlier stock price indexes but, as is apparent in Jorion and Goetzmann (1999), they have encountered problems over dividend availability. We show in section 6 that this is a serious drawback, because the contribution of dividends to equity returns is of the same order of magnitude as the equity premium itself, and since there have been considerable cross-country differences in average dividend yield. The absence of dividends makes it hard to generate meaningful estimates of equity premiums.

Even for countries where long-run total returns series were available, we have seen that they sometimes suffered from *ex post* selection bias, as had been the case in the U.K. Finally, the data sources that pre-dated the DMS database often suffered from "easy data" bias. This refers to the tendency of researchers to use data that is easy to obtain, excludes traumatic intervals such as wars and their aftermath, and typically relates to more recent periods. Dimson, Marsh, and Staunton (2002) identify the most widely cited prior data source for each of 16 countries and show that equity returns over the periods covered are higher than the 1900–2000 returns from the DMS database by an average of 3% per year. Easy data bias almost certainly led researchers to believe that equity returns over the twentieth century were higher than was really the case.

The DMS Global Database: Composition and Start-date

These deficiencies in existing data provided the motivation for the DMS global database. This contains annual returns on stocks, bonds, bills, inflation, and currencies for 17 countries from 1900–2005, and is described in Dimson, Marsh, and Staunton (2006a and 2006b). The countries include the United States and Canada, seven markets from what is now the Euro currency area, the United Kingdom and three other European markets that have not embraced the Euro, two Asia-Pacific markets, and one African market. Together, they made up 91% of total world equity market capitalization at the start of 2006, and we estimate that they constituted 90% by value at the start of our period in 1900 (see section 5 for more details).

The DMS database also includes four "world" indexes based on the countries included in the DMS dataset. There is, first, a World equity index: a 17-country index denominated in a common currency, here taken as U.S. dollars, in which each country is weighted by its starting-year equity market capitalization or, in years before capitalizations were available, by its GDP. Second, there is an analogous 16-country worldwide equity index that excludes the United States ("World ex-U.S."). Third and fourth, we compute a World bond index and a World ex-U.S. bond index, both of which are constructed in the same way, but with each country weighted by its GDP.

The DMS series all commence in 1900, and this common start-date aids international comparisons. The choice of start-date was dictated by data availability and quality. At first sight, it appears feasible to start earlier. Jorion and Goetzmann (1999) note that, by 1900, stock exchanges existed in at least 33 of today's nations, with markets in seven countries dating back another 100 years to 1800. An earlier start-date would in principle be desirable, as a very long series of stationary returns is needed to estimate the equity premium with any precision. Even with non-stationary returns, a long time-series is still helpful,[4] and it would anyway be interesting to compare nineteenth century premiums with those from later years. Indeed, some researchers report very low premiums for the nineteenth century. Mehra and Prescott (2003) report a U.S. equity premium of zero over 1802–62, based on Schwert's (1990) equity series and Siegel's (2002) risk free rate estimates, while Hwang and Song (2004) claim there was no U.K. equity premium puzzle in the nineteenth century, since bonds outperformed stocks.

These inferences, however, are unreliable due to the poor quality of nineteenth century data. The equity series used by Hwang and Song omits dividends, and before 1871, suffers from *ex post*

[4] Pástor and Stambaugh (2001) show that a long return history is useful in estimating the current equity premium even if the historical distribution has experienced structural breaks. The long series helps not only if the timing of breaks is uncertain but also if one believes that large shifts in the premium are unlikely or that the premium is associated, in part, with volatility.

bias and poor coverage. From 1871–1913, they use a broader index (Grossman (2002)), but this has problems with capital changes, omitted data, and stocks disappearing. Within the range of likely assumptions about these disappearances, Grossman shows that he can obtain a 1913 end-value of anywhere between 400 and 1700 (1871=100). Mehra and Prescott (2003) list similar weaknesses in Schwert's 1802–71 U.S. data, such as the lack of dividends, tiny number of stocks, frequent reliance on single sectors, and likelihood of *ex post* bias. These flaws undermine the reliability of equity premium estimates for the nineteenth century.

Unfortunately, better nineteenth century U.K. equity indexes do not exist, and, until recently, Schwert's series was the only source of pre-1871 U.S. data. However, most recently, Goetzmann and Ibbotson (2006) employ a new NYSE database for 1815–1925 (see Goetzmann, Ibbotson, and Peng (2001)) to estimate the nineteenth century U.S. equity premium. But they highlight two problems. First, dividend data is absent pre-1825, and incomplete from 1825–71. Equity returns for 1825–71 are thus estimated in two ways based on different assumptions about dividends, producing two widely divergent estimates of the mean annual return, namely, 6.1% and 11.5%, which are then averaged. Second, since Treasury bills or their equivalents did not yet exist, the risk free rate proves even more problematic and has to be estimated from risky bonds. These two factors make it hard to judge the efficacy of their nineteenth century equity premium estimates.

Returning to the question of the start-date for the DMS database, it is clear that, even for the United States, the world's best-documented capital market, pre-1871 data is still problematic. Wilson and Jones (2002) observe that after 1871, U.S. equity returns are of higher quality; but while a few other DMS countries also have acceptable series over this period, most, including the United Kingdom, have no suitable data prior to 1900. Before then, there are virtually no stock indexes to use as a starting point, and creating new nineteenth century indexes would be a major task, requiring hand collection of stock data from archives.[5] For practical purposes, 1900 is thus the earliest plausible common start-date for a comparative international database.

The DMS Global Database: General Methodology and Guiding Principles

The DMS database comprises annual returns, and is based on the best quality capital appreciation and income series available for each country, drawing on previous studies and other sources. Where possible, data were taken from peer-reviewed academic papers, or highly rated professional studies. From the end point of these studies, the returns series are linked into the best, most comprehensive, commercial returns indexes available. The DMS database is updated annually (see Dimson, Marsh, and Staunton (2006a and 2006b)). Appendix 2 lists the data sources used for each country.

To span the entire period from 1900 we link multiple index series. The best index is chosen for each period, switching when feasible to better alternatives, as they become available. Other factors equal, we have chosen equity indexes that afford the broadest coverage of their market.

---

[5] The Dow Jones Industrial Average was, we believe, the first index ever published. It began in 1884 with 11 constituents. Charles Dow had neither computer nor calculator, hence his limited coverage. While today, computation is trivial, creating indexes more than 100 years after the event poses a major data challenge. While it is often fairly easy to identify hard copy sources of stock prices, the real problems lie in identifying (i) the full population, including births, name changes, and deaths and their outcome, and (ii) data on dividends, capital changes, shares outstanding, and so on. Archive sources tend to be poorer, or non-existent, the further back one goes in time.

The evolution of the U.S. equity series illustrates these principles. From 1900–25, we use the capitalization weighted Cowles Index of all NYSE stocks (as modified by Wilson and Jones (2002)); from 1926–61, we use the capitalization weighted CRSP Index of all NYSE stocks; from 1962–70, we employ the extended CRSP Index, which over this period also includes Amex stocks; and from 1971 on, we utilize the Wilshire 5000 Index, which contains over 7,000 U.S. stocks, including those listed on Nasdaq.

The creation of the DMS database was in large part an investigative and assembly operation. Most of the series needed already existed, but some were long forgotten, unpublished, or came from research in progress. In other cases, the task was to estimate total returns by linking dividends to existing capital gains indexes. But for several countries, there were periods for which no adequate series existed. For example, U.K. indexes were of poor quality before 1962, and far from comprehensive thereafter. To remedy this, we compiled an index spanning the entire U.K. equity market for 1955–2005 (Dimson and Marsh (2001)), while for 1900–1955, we built a 100-stock index by painstaking data collection from archives. Similarly, we used archive data to span missing sub-periods for Canada, Ireland, Norway, Switzerland, and South Africa.

Virtually all of the DMS countries experienced trading breaks at some point in their history, often in wartime. Jorion and Goetzmann (1999) provide a list and discuss the origins of these interruptions. In assembling our database, we needed to span these gaps. The U.K. and European exchanges, and even the NYSE, closed at the start of World War I, but typically reopened 4–6 months later. Similarly, the Danish, Norwegian, Belgian, Dutch and French markets were closed for short periods when Germany invaded in 1940, and even the Swiss market closed from May to July 1940 for mobilization. There were other temporary closures, notably in Japan after the Great Tokyo Earthquake of 1923. These relatively brief breaks were easy to bridge.[6] But three longer stock exchange closures proved more difficult: Germany and Japan from towards the end of World War II, and Spain during the Civil War. We were able to bridge these gaps,[7] but as markets were closed or prices were controlled, the end-year index levels recorded for Germany for 1943–47, Japan for 1945, and Spain for 1936–38 cannot be regarded as market-determined values. This needs to be borne in mind when reviewing arithmetic means, standard deviations, and other statistics relating to annual returns computed using these values. Over each of these stock exchange closures, more reliance can be placed on the starting and ending values than on the intermediate index levels. We are therefore still able to compute changes in investors' wealth and geometric mean returns over periods spanning these closures.

Finally, there was one unbridgeable discontinuity, namely, bond and bill (but not equity) returns in

---

[6] Since the DMS database records annual returns, trading breaks pose problems only when they span a calendar year boundary. For example, at the start of World War I, the NYSE was closed from 31 July until 11 December 1914, so it was still possible to calculate equity and bond returns for 1914. However, the London Stock Exchange closed in July 1914 and did not reopen until 5 January 1915, so prices for the latter date were used as the closing prices for 1914 and the opening prices for 1915. A similar approach was adopted for French returns during the closure of the Paris Exchange from June 1940 until April 1941.

[7] Wartime share dealing in Germany and Japan was subject to strict controls. In Germany, stock prices were effectively fixed after January 1943; the market closed in 1944 with the Allied invasion, and did not reopen until July 1948. Both Gielen (1944) and Ronge (2002) provide data that bridges the gap between 1943 and 1948. In Japan, stock market trading was suspended in August 1945, and although it did not officially reopen until May 1949, over-the-counter trading resumed in May 1946, and the Oriental Economist Index provides relevant stock return data. In Spain, trading was suspended during the Civil War from July 1936 to April 1939, and the Madrid exchange remained closed through February 1940; over the closure we assume a zero change in nominal stock prices and zero dividends.

Germany during the hyperinflation of 1922–23, when German bond and bill investors suffered a total loss of –100%. This episode serves as a stark reminder that, under extreme circumstances, bonds and bills can become riskier than equities. When reporting equity premiums for Germany, whether relative to bonds or bills, we thus have no alternative but to exclude the years 1922–23.

All DMS index returns are computed as the arithmetic average of the individual security returns, and not as geometric averages (an inappropriate method encountered in certain older indexes); and all the DMS security returns include reinvested gross (pre-tax) income as well as capital gains. Income reinvestment is especially important, since, as we saw above, many early equity indexes measure just capital gains and ignore dividends, thus introducing a serious downward bias. Similarly, many early bond indexes record only yields, ignoring price movements. Virtually all DMS equity indexes are capitalization weighted, and are calculated from year-end stock prices, but in the early years, for a few countries, we were forced to use equally weighted indexes or indexes based on average- or mid-December prices (see Appendix 2).

Our guiding principle was to avoid survivorship, success, look-ahead, or any other form of *ex post* selection bias. The criterion was that each index should follow an investment policy that was specifiable in advance, so that an investor could have replicated the performance of the index (before dealing costs) using information that would have been available at the time. The DMS database and its world indexes do, however, suffer from survivorship bias, in the sense that all 17 countries have a full 106-year history. In 1900, an investor could not have known which markets were destined to survive. Certainly, in some markets that existed in 1900, such as Russia and China, domestic equity and bond investors later experienced total losses. In section 5 below, we assess the likely impact of this survivorship bias on our worldwide equity premium estimates.

The DMS inflation rates are derived from each country's consumer price index (CPI), although for Canada (1900–10), Japan (1900), and Spain (1900–14) the wholesale price index is used, as no CPI was available. The exchange rates are year-end rates from *The Financial Times* (1907–2005) and *The Investors' Review* (1899–1906). Where appropriate, market or unofficial rates are substituted for official rates during wartime or the aftermath of World War II. DMS bill returns are in general treasury bill returns, but where these instruments did not exist, we used the closest equivalent, namely, a measure of the short-term interest rate with the lowest possible credit risk.

The DMS bond indexes are based on government bonds. They are usually equally weighted, with constituents chosen to fall within the desired maturity range. For the United States and United Kingdom, they are designed to have a maturity of 20 years, although from 1900−55, the U.K. bond index is based on perpetuals, since there were no 20-year bonds in 1900, and perpetuals dominated the market in terms of liquidity until the 1950s. For all other countries, 20-year bonds are targeted, but where these are not available, either perpetuals (usually for earlier periods) or shorter maturity bonds are used. Further details are given in Appendix 2.

In summary, the DMS database is more comprehensive and accurate than the data sources used in previous research and it spans a longer period. This allows us to set the U.S. equity premium alongside comparable 106-year premiums for 16 other countries and the world indexes, thereby helping us to put the U.S. experience in perspective.

4. LONG-RUN HISTORICAL RATES OF RETURN

In this section we use the DMS dataset to examine real equity market returns around the world. In Table **1**, we compare U.S. returns with those in 16 other countries, and long run returns with recent performance, to help show why we need long time series when analyzing equity returns.

The second column of Table 1 reports annualized real returns over the early years of the twenty-first century, from 2000–2005, the most recent 6-year period at the time of writing. It shows that real equity returns were negative in seven of the seventeen countries and that the return on the world index was -1.25%. Equities underperformed bonds and bills (not shown here) in twelve of the seventeen countries. Inferring the expected equity premium from returns over such a short period would be nonsense: investors cannot have required or expected a negative return for assuming risk. This was simply a disappointing period for equities.

It would be just as misleading to project the future equity premium from data for the previous decade. Column three of Table 1 shows that, with the exception of one country, namely, Japan, which we discuss below, real equity returns between 1990 and 1999 were typically high. Over this period, U.S. equity investors achieved a total real return of 14.2% per annum, increasing their initial stake five-fold. This was a golden age for stocks, and golden ages are, by definition, untypical, providing a poor basis for future projections.

Table 1: Real Equity Returns in 17 Countries, 1900–2005

	Annualized Returns (% p.a.)			**Properties of Annual (%) Real Returns, 1900–2005**					
Country	**2000 to 2005**	**1990 to 1999**	**1900 to 2005**	**Arith. Mean**	**Std. Error**	**Std. Devn.**	**Skew-ness**	**Kurt-osis**	**Serial Corr.**
Belgium	3.99	9.13	**2.40**	4.58	2.15	22.10	0.95	2.33	0.23
Italy	-0.73	6.42	**2.46**	6.49	2.82	29.07	0.76	2.43	0.03
Germany	-4.08	9.89	**3.09**	8.21	3.16	32.53	1.47	5.65	-0.12
France	-1.64	12.53	**3.60**	6.08	2.25	23.16	0.41	-0.27	0.19
Spain	2.48	12.16	**3.74**	5.90	2.12	21.88	0.80	2.17	0.32
Norway	10.91	8.25	**4.28**	7.08	2.62	26.96	2.37	11.69	-0.06
Switzerland	1.11	13.95	**4.48**	6.28	1.92	19.73	0.42	0.38	0.18
Japan	0.64	-5.23	**4.51**	9.26	2.92	30.05	0.49	2.36	0.19
Ireland	5.14	11.79	**4.79**	7.02	2.15	22.10	0.60	0.81	-0.04
World ex-U.S (USD)	0.11	3.41	**5.23**	7.02	1.92	19.79	0.58	1.41	0.25
Denmark	9.41	7.52	**5.25**	6.91	1.97	20.26	1.83	6.71	-0.13
Netherlands	-5.41	17.79	**5.26**	7.22	2.07	21.29	1.06	3.18	0.09
United Kingdom	-1.34	11.16	**5.50**	7.36	1.94	19.96	0.66	3.69	-0.06
World (USD)	-1.25	7.87	**5.75**	7.16	1.67	17.23	0.13	1.05	0.15
Canada	4.32	8.28	**6.24**	7.56	1.63	16.77	0.09	-0.13	0.16
United States	-2.74	14.24	**6.52**	8.50	1.96	20.19	-0.14	-0.35	0.00
South Africa	11.05	4.61	**7.25**	9.46	2.19	22.57	0.94	2.58	0.05
Australia	7.78	8.98	**7.70**	9.21	1.71	17.64	-0.25	0.06	-0.02
Sweden	-0.70	15.02	**7.80**	10.07	2.20	22.62	0.55	0.92	0.11

Extremes of History

While the 1990s and early 2000s were not typical, they are not unique. The top panel of Table **2** highlights other noteworthy episodes of world political and economic history since 1900. It shows real equity returns over the five worst episodes for equity investors, and over four "golden ages" for the world indexes and the world's five largest markets. These five markets are of interest not just because of their economic importance, but also because they experienced the most extreme returns out of all 17 countries in our database.

The five worst episodes for equity investors comprise the two World Wars and the three great bear markets—the Wall Street Crash and Great Depression, the first oil shock and recession of 1973–74, and the 2000–02 bear market after the internet bubble. While the World Wars were in

Table 2: Real Equity Returns in Key Markets over Selected Periods

		Real Rate of Return (%) over the Period						
Period	**Description**	**U.S.**	**U.K.**	**France**	**Germany**	**Japan**	**World**	**World ex-US**
Selected Episodes								
1914–18:	World War I	-18	-36	-50	-66	66	-20	-21
1919–28	Post-WWI recovery	372	234	171	18	30	209	107
1929–31	Wall Street Crash	-60	-31	-44	-59	11	-54	-47
1939–48	World War II	24	34	-41	-88	-96	-13	-47
1949–59	Post-WWII recovery	426	212	269	4094	1565	517	670
1973–74	Oil shock/recession	-52	-71	-35	-26	-49	-47	-37
1980–89	Expansionary 80s	184	319	318	272	431	255	326
1990–99	90s tech boom	279	188	226	157	-42	113	40
2000–02	Internet 'bust'	-42	-40	-46	-57	-49	-44	-46
Periods with Highest Returns								
1-year	Return	57	97	66	155	121	70	79
periods	*Period*	*1933*	*1975*	*1954*	*1949*	*1952*	*1933*	*1933*
2-year	Return	90	107	123	186	245	92	134
periods	*Period*	*1927–28*	*1958–59*	*1927–28*	*1958–59*	*1951–52*	*1932–33*	*1985–86*
5-year	Return	233	176	310	652	576	174	268
periods	*Period*	*1924–28*	*1921–25*	*1982–86*	*1949–53*	*1948–52*	*1985–89*	*1985–89*
Periods with Lowest Returns								
1-year	Return	-38	-57	-40	-91	-86	-35	-41
periods	*Period*	*1931*	*1974*	*1945*	*1948*	*1946*	*1931*	*1946*
2-year	Return	-53	-71	-54	-90	-95	-47	-52
periods	*Period*	*1930–31*	*1973–74*	*1944–45*	*1947–48*	*1945–46*	*1973–74*	*1946–47*
5-year	Return	-45	-63	-78	-93	-98	-50	-56
periods	*Period*	*1916–20*	*1970–74*	*1943–47*	*1944–48*	*1943–47*	*1916–20*	*1944–48*
Longest Runs of Negative Real Returns								
Longest	Return	-7	-4	-8	-8	-1	-9	-11
runs over	*Period*	*1905–20*	*1900–21*	*1900–52*	*1900–54*	*1900–50*	*1901–20*	*1928–50*
106 years	Number of Years	16	22	53	55	51	20	23

aggregate negative for equities, there were relative winners and losers, corresponding to each country's fortunes in war. Thus in World War I, German equities performed the worst (–66%), while Japanese stocks fared the best (+66%), as Japan was a net gainer from the war. In World War II and its aftermath,[8] Japanese and German equities were decimated (–96% and –88% respectively), while both U.S. and U.K. equities enjoyed small positive real returns.

Table 2 shows that the world wars were less damaging to world equities than the peacetime bear markets. From 1929–31, during the Wall Street Crash and ensuing Great Depression, the world index fell by 54% in real, U.S. dollar terms, compared with 20% during World War I and 13% in World War II. For the United States, Germany, and the world index this was the most savage of the three great bear markets, and from 1929–31 the losses in real terms were 60%, 59%, and 54%, respectively. From peak to trough, the falls were even greater. Table 2 records calendar year returns, but the U.S. equity market did not start falling until September 1929, reaching its nadir in June 1932, 79% (in real terms) below its 1929 peak.

British and Japanese investors, in contrast, suffered greater losses in 1973–74 than during the 1930s. This was the time of the first OPEC oil squeeze after the 1973 October War in the Middle East, which drove the world into deep recession. Over 1973–74, the real returns on U.K., U.S., Japanese, and world equities were –71%, –52%, –49%, and –47%, respectively. The last row of the top panel of Table 2 shows that the world equity index fell by almost as much (44% in real terms) in the bear market of 2000–02, which followed the late 1990s internet bubble. Table 2 shows the returns over calendar years, and from the start of 2000 until the trough of the bear market in March 2003, the real returns on U.S., U.K., Japanese, and German equities were even lower at –47%, –44%, –53%, and –65%, respectively.

The top panel of Table 2 also summarizes real returns over four "golden ages" for equity investors. The 1990s, which we highlighted in Table 1 as a recent period of exceptional performance, was the most muted of the four, with the world index showing a real return of 113%. While the 1990s was an especially strong period for the U.S. market (279% real return), the world index was held back by Japan.[9] The world index rose by appreciably more during the 1980s (255% in real terms) and the two post-world war recovery periods (209% in the decade after World War I and 517% from 1949–59). During the latter period, a number of equity markets enjoyed quite staggering returns. For example, Table 2 shows that during these nascent years of the German and Japanese "economic miracles", their equity markets rose in real terms by 4094% (i.e., 40.4% p.a.) and 1565% (29.1% p.a.), respectively.

---

[8] To measure the full impact of World War II on German and Japanese equity returns, it is necessary to extend the period through to 1948 to include the aftermath of the war. This is because, as noted above, stock prices in Germany were effectively fixed after January 1943, and the exchanges closed in 1944 with the Allied invasion, and did not reopen until July 1948, when prices could finally reflect the destruction from the war. Meanwhile, German inflation from 1943–48 was 55%. In Japan, the stock market closed in 1944, but over-the-counter trading resumed from 1946 onwards. In Japan, the sharp negative real returns recorded in 1945, 1946, and 1947 thus reflect the hyperinflation that raged from 1945 onward (inflation from 1945–48 was 5,588%), the resumption of trading at market-determined prices in 1946, and the break-up of the zaibatsu industrial cartels and the distribution of their shares to the workforce.

[9] Table 2 shows that Japan experienced a real return of –42% during the 1990s (equivalent to an annualized real return of –5.2% p.a. as shown in the third column of Table 1). At the start of the 1990s, the Japanese stock market was the largest in the world by market capitalization, with a 40.4% weighting in the world index, compared with 32.2% for the United States. Japan's poor performance, coupled with its high weighting in the world index, and even higher weighting (60%) in the world ex-U.S. naturally had a depressing effect on the returns on the world and world-ex U.S indexes (see Table 2 and column 2 of Table 1).

The second and third panels of Table 2 show the returns for, and dates of, the one-, two-, and five-year periods during which each country and the world indexes experienced their highest and lowest returns. The picture that emerges reinforces the discussion above: in nearly all cases, the best and worst periods are drawn from, and are subsets of, the episodes listed in the top panel. Note that the spreads between worst and best are wide. One-year real returns range from –35% to +70% (world), –38% to +57% (United States), –91% to +155% (Germany), and –86% to 121% (Japan). Five-year real returns extend from –50% to +174% (world), –45% to +233% (United States), –93% to +652% (Germany), and –98% to 576% (Japan).

Finally, the bottom panel of Table 2 reports the longest period over which each country (or world index) has experienced a cumulative negative real return. It shows that for the United States, the longest such period was the 16 years from 1905–20, when the cumulative return was –7%. This reconfirms Siegel's (2002) observation that U.S. investors have historically always enjoyed a positive real return as long as they have held shares for at least 20 years. However, Table 2 shows that investors in other countries have not been so fortunate, with Japan, France, and Germany suffering extended periods lasting over half a century during which cumulative equity returns remained negative in real terms. Dimson, Marsh, and Staunton (2004) report that three-quarters of the DMS countries experienced intervals of negative real stock market returns lasting for more than two decades.

The Long-Run Perspective

The statistics presented in Tables 1 and 2 and the discussion in the previous section serve to emphasize the volatility of stock markets, and the substantial variation in year-to-year and period-to-period returns. Clearly, because of this volatility, we need to examine intervals that are much longer than five years or a decade when estimating means or equity premiums. The fourth column of Table 1 (shown in boldface) illustrates the perspective that longer periods of history can bring by displaying real equity returns over the 106-year period 1900–2005. Clearly, these 106-year returns contrast favourably with the disappointing returns over 2000–2005 (second column), but they are much lower than the returns in the 1990s (third column).

The remaining columns of Table 1 present formal statistics on the distribution of annual real returns over 1900–2005, and again, they emphasize how volatile stock markets were over this period. The arithmetic means of the 106 one-year real returns are shown in the fifth column. These exceed the geometric means (fourth column) by approximately half the variance of the annual returns. The standard deviation column shows that the U.S., U.K., Swiss, and Danish equity markets all had volatilities of around 20%. While this represents an appreciable level of volatility, these countries are at the lower end of the risk spectrum, with only Australia and Canada having lower standard deviations. The highest volatility markets were Italy, Japan, and Germany, with volatilities close to, or above, 30%. These high levels of volatility imply that the arithmetic means are estimated with high standard errors (see column six), and we return to this issue below when we discuss the precision of equity premium estimates.

The skewness and excess kurtosis columns in Table 1 show that returns were positively skewed except in the United States, and in most countries, they were noticeably more fat-tailed than would be expected if they were normally distributed.[10] Finally, the serial correlation column shows that to a good approximation, returns are serially independent. The average serial correlation coefficient was 0.07, and only two out of 17 coefficients were significant at the 95% level—only slightly higher than the proportion that would be expected from chance.

The fourth column of Table 1 shows that the 106-year annualized real return on U.S. equities was 6.5%. The equivalent real return on non-U.S. equities—from the perspective of a U.S. investor, and as measured by the world index excluding the United States—was lower at 5.2%. This lends initial support to the concern about success bias from focusing solely on the United States. At the same time, the gap is not large, and it is also clear from Table 1 that the stock markets of several other countries performed even better than the United States. Table 1 shows real returns in local currency terms, however, rather than equity premiums, and we defer presenting comprehensive comparisons of the latter until Section 5 below.

However, to reinforce the importance of focusing on long-run data, we briefly preview the equity premium data for the U.S. market. The bars in **Figure 2** show the year-by-year historical U.S. equity premium calculated relative to the return on Treasury bills over 1900–2005.[11] The lowest premium was –45% in 1931, when equities earned –44% and Treasury bills 1%; the highest was 57% in 1933, when equities earned 57.6% and bills 0.3%. Over the entire 106-year interval, the mean annual excess return over treasury bills was 7.4%, while the standard deviation was 19.6%. On average, therefore, this confirms that U.S. investors received a positive, and large, reward for exposure to equity market risk.

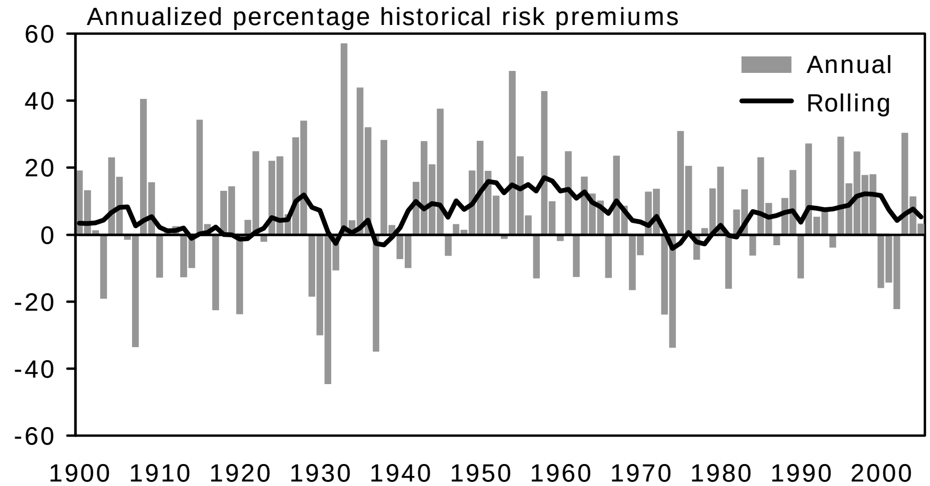
Because the range of year-to-year excess returns is very broad, it would be misleading to label these as "risk premiums." As noted above, investors cannot have expected, let alone required, a negative risk premium from investing in equities. Many low and all negative premiums must therefore reflect unpleasant surprises. Nor could investors have required premiums as high as the 57% achieved in 1933. Such numbers are quite implausible as a required reward for risk, and the high realizations must therefore reflect pleasant surprises. To avoid confusion, it is helpful to refer to a return in excess of the risk free rate, measured over a period in the past, simply as an excess return or as the "historical" equity premium (rather than equity premium). When looking to the future, it is helpful to refer to the "expected" or "prospective" equity premium.

---

[10] The average coefficients of skewness and kurtosis for the 17 countries were 0.76 and 2.60. This is consistent with our expectation that the distribution of annual stock returns would be lognormal, rather than normal, and hence positively skewed. But when we examine the distribution of log returns (i.e., the natural logarithm of one plus the annual return), we find average skewness and kurtosis of –0.48 and 3.25, i.e., the skewness switches from positive to negative, and the distributions appear even more leptokurtic. This finding is heavily influenced by the extreme negative returns for Germany in 1948 and Japan in 1946. As noted in section 3 above, German returns from 1943–48 and Japanese returns from 1945–46 must be treated with caution, as although the total return over these periods is correct, the values for individual years cannot be regarded as market-determined. The values recorded for Germany in 1948 and Japan in 1946 thus almost certainly include accumulated losses from previous years. Excluding Germany and Japan, the coefficients of skewness and kurtosis based on log returns were –0.20 and 1.40, which are much closer to the values we would expect if annual returns were lognormally distributed.

[11] For convenience, we estimate the equity premium from the arithmetic difference between the logarithmic return on equities and the logarithmic return on the riskless asset. Equivalently, we define $1{+}Equity\,Premium$ to be equal to $1{+}Equity\,Return$ divided by $1{+}Riskless\,Return$. Defined this way, the equity premium is a ratio and therefore has no units of measurement. It is identical if computed from nominal or real returns, or if computed from dollar or euro returns.

Figure 2: Annual and Rolling Ten-Year U.S. Premiums Relative to Bills, 1900–2005



The ten-year excess returns were sometimes negative, most recently in the 1970s and early 1980s. Figure 2 also reveals several cases of double-digit ten-year premiums. Clearly, a decade is too brief for good and bad luck to cancel out, or for drawing inferences about investor expectations. Indeed, even with over a century of data, market fluctuations have an impact. Taking the United Kingdom as an illustration, the arithmetic mean annual excess return from 1900–49 was only 3.1%, compared to 8.8% from 1950–2005. As over a single year, all we are reporting is the excess return that was realized over a period in the past.

To quantify the degree of precision in our estimates, we can compute standard errors. Assuming that each year's excess return is serially independent,[12] the standard error of the mean historical equity premium estimate is approximately $\sigma/\sqrt{T}$, where σ is the standard deviation of the annual excess returns, and T is the period length in years. Since we have seen that σ was close to 20% for the U.S. market, this implies that the standard error of the mean historical equity premium estimated over ten years is 6.3%, while the standard error using 106 years of data remains quite high at approximately 2%. Since we saw in Table 1 above that most countries had a standard deviation that exceeded that of the U.S. market, the standard error of the mean equity premium is typically larger in non-American markets.

When estimating the historical equity premium, therefore, the case for using long-run data is clear. Stock returns are so volatile that it is hard to measure the mean historical premium with precision. Without long-run data, the task is impossible, and even with over a century of data, the standard error remains high—even if we assume that the underlying series is stationary.

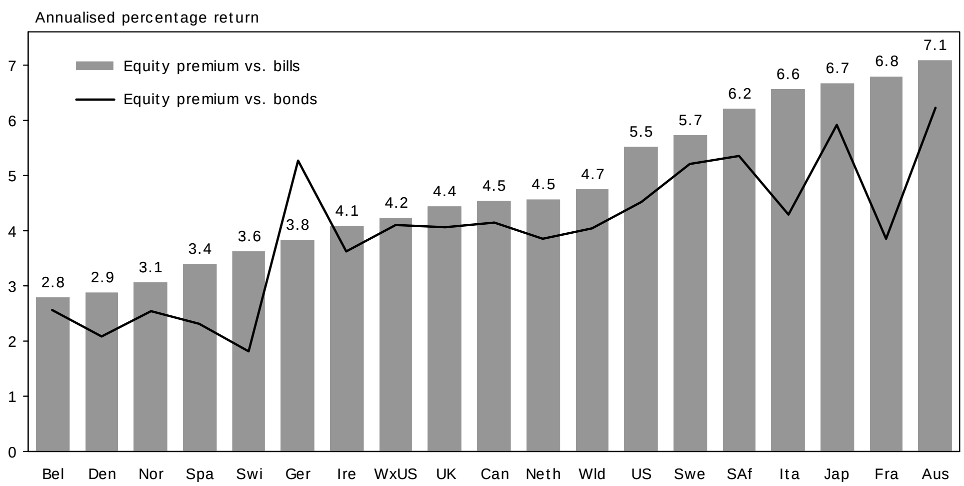
---

[12] We saw in Table 1 above that this was a good approximation for real returns, and the same holds true for excess returns. For the United States, the serial correlation of excess returns over 1900–2005 was 0.00, while the average across all 17 countries was 0.05. For excess returns defined relative to bonds rather than bills, the average serial correlation was 0.04.

5. NEW GLOBAL EVIDENCE ON THE EQUITY PREMIUM

Figure 3 shows the annualized (geometric mean) historical equity premiums over the 106-year period from 1900–2005 for each of the 17 countries in the DMS database, as well as the world index and the world excluding the United States. Countries are ranked by the equity premium relative to bills (or the nearest equivalent short-term instrument), displayed as bars. The line-plot shows each country's equity premium relative to bonds (long-term government bonds). Since the world indexes are computed here from the perspective of a U.S. (dollar) investor, the world equity premiums relative to bills are calculated with reference to the U.S. risk-free (Treasury bill) rate. The world equity premiums relative to bonds are calculated relative to the world bond indexes.

Figure 3: Worldwide Annualized Equity Premiums 1900–2005*



* Germany omits 1922-23

Figure 3 shows that equities outperformed both bills and bonds in all 17 countries over this period, and that, in general, the equity premium was large. The chart lends support to the concern about generalizing from the U.S. experience by showing that the U.S. equity premium relative to bills was 5.5% compared with 4.2% for the rest of the world. But while noteworthy, this difference is not that large, and Figure 3 shows that several countries had larger premiums than the United States. For the world index (with its large U.S. weighting), the premium relative to bills was 4.7%. The U.K. equity premium was a little below the world average at 4.4%.

Relative to long bonds, the story for the 17 countries is similar, although on average, the premiums were around 0.8% lower, reflecting the average term premium, i.e., the annualized amount by which bond returns exceeded bill returns. The annualized U.S. equity premium relative to bonds was 4.5% compared with 4.1% for the world ex-U.S. Across all 17 countries, the equity premium relative to bonds averaged 4.0%, and for the world index it was also 4.0%.[13] Thus,

[13] Over the entire period, the annualized world equity risk premium relative to bills was 4.74%, compared with 5.51% for the United States. Part of this difference, however, reflects the strength of the dollar. The world risk premium is computed here from the world equity index expressed in dollars, in order to reflect the perspective of a U.S.-based global investor. Since the currencies of most other countries depreciated against the dollar over the twentieth century, this lowers our estimate of the world equity risk premium relative to the (weighted) average of the local-currency-based estimates for individual countries.

while U.S. and U.K. equities have performed well, both countries are toward the middle of the distribution of worldwide equity premiums, and even the United States is not hugely out of line compared to other markets.

The Equity Premium Around the World

Table 3 provides more detail on the historical equity premiums. The left half of the table shows premiums relative to bills, while the right half shows premiums relative to government bonds. In each half of the table we show the annualized, or geometric mean, equity premium over the entire 106 years (i.e., the data plotted in Figure 3); the arithmetic mean of the 106 one-year premiums; the standard error of the arithmetic mean; and the standard deviation of the 106 one-year premiums. The geometric mean is, of course, always less than the arithmetic mean, the difference being approximately one-half of the variance of the historical equity premium.

Table 3 shows that the *arithmetic* mean annual equity premium relative to bills for the United States was 7.4% compared with 5.9% for the world excluding the United States. This difference of 1.5% again lends support to the notion that it is dangerous to extrapolate from the U.S. experience because of *ex post* success bias. But again we should note that Table 3 shows that the United States was by no means the country with the largest arithmetic mean premium. Indeed, on a strict ranking of arithmetic mean premiums, it was eighth largest out of 17 countries.

Table 3: Annualized Equity Premiums for 17 Countries, 1900–2005

% p.a.	**Historical Equity Premium Relative to Bills**				**Historical Equity Premium Relative to Bonds**			
Country	**Geometric Mean**	**Arithmetic Mean**	**Standard Error**	**Standard Deviation**	**Geometric Mean**	**Arithmetic Mean**	**Standard Error**	**Standard Deviation**
Australia	7.08	8.49	1.65	17.00	6.22	7.81	1.83	18.80
Belgium	2.80	4.99	2.24	23.06	2.57	4.37	1.95	20.10
Canada	4.54	5.88	1.62	16.71	4.15	5.67	1.74	17.95
Denmark	2.87	4.51	1.93	19.85	2.07	3.27	1.57	16.18
France	6.79	9.27	2.35	24.19	3.86	6.03	2.16	22.29
Germany*	3.83	9.07	3.28	33.49	5.28	8.35	2.69	27.41
Ireland	4.09	5.98	1.97	20.33	3.62	5.18	1.78	18.37
Italy	6.55	10.46	3.12	32.09	4.30	7.68	2.89	29.73
Japan	6.67	9.84	2.70	27.82	5.91	9.98	3.21	33.06
Netherlands	4.55	6.61	2.17	22.36	3.86	5.95	2.10	21.63
Norway	3.07	5.70	2.52	25.90	2.55	5.26	2.66	27.43
South Africa	6.20	8.25	2.15	22.09	5.35	7.03	1.88	19.32
Spain	3.40	5.46	2.08	21.45	2.32	4.21	1.96	20.20
Sweden	5.73	7.98	2.15	22.09	5.21	7.51	2.17	22.34
Switzerland	3.63	5.29	1.82	18.79	1.80	3.28	1.70	17.52
U.K.	4.43	6.14	1.93	19.84	4.06	5.29	1.61	16.60
U.S.	5.51	7.41	1.91	19.64	4.52	6.49	1.96	20.16
Average	**4.81**	**7.14**	**2.21**	**22.75**	**3.98**	**6.08**	**2.11**	**21.71**
World-ex U.S.	**4.23**	**5.93**	**1.88**	**19.33**	**4.10**	**5.18**	**1.48**	**15.19**
World	**4.74**	**6.07**	**1.62**	**16.65**	**4.04**	**5.15**	**1.45**	**14.96**

* Germany omits 1922–23

Care is needed, however, in comparing and interpreting long-run arithmetic mean equity premiums. For example, Table 3 shows that, relative to bills, Italy had the highest arithmetic equity premium at 10.5%, followed by Japan at 9.8%, France at 9.3%, and Germany at 9.1%. Yet these four countries had below average equity returns (see Table 1). Table 3 shows that part of the explanation lies in the high historical volatilities in these four markets, 32%, 28%, 24% and 33%, respectively. As we saw above, much of this volatility arose during the first half of the twentieth century, during, or in the aftermath of, the World Wars. In all four cases, therefore, the long-run equity premium earned by investors (the geometric mean) was well below the arithmetic mean. But this is only part of the story, since Table 3 shows that these countries still had above-average geometric equity premiums, despite their below-average equity market returns. (Italy, Japan, and France had above average premiums relative to bills, while Italy, Japan, and Germany had above average premiums relative to bonds). The explanation, of course, lies in the very poor historical bill and/or bond returns in these four countries, and we return below to the issue of poor equity returns coinciding with poor bill and bond returns.

Table 3 shows that both the U.S. and U.K. equity premiums relative to bills had similar standard deviations of close to 20% per annum, and that only four other countries had standard deviations that were as low, or lower than this. As noted above, the relatively high standard deviations for the equity premiums for the 17 countries, ranging from 17–33%, indicate that, even with 106 years of data, the potential inaccuracy in historical equity premiums is still fairly high. Table 3 shows that the standard error of the equity premium relative to bills is 1.9% for the United States, and the range runs from 1.6% (Canada) to 3.3% (Germany).

A Smaller Risk Premium

By focusing on the world, rather than the United States, and by extending the time span to 1900–2005, the equity premium puzzle has become quantitatively smaller. We saw in Section 2 that, before our new database became available in 2000, the most widely cited number for the U.S. arithmetic mean equity premium relative to bills was the Ibbotson (2000) estimate for 1926–99 of 9.2%. Table 3 shows that by extending the time period backwards to include 1900–25 and forwards to embrace 2000–05, while switching to more comprehensive index series, the arithmetic mean equity premium shrinks to 7.4%. Table 3 also shows that the equivalent world equity premium over this same period was 6.1%.

But while the puzzle has become smaller than it once was, 6.1% remains a large number. Indeed, Mehra and Prescott's original article documented a premium of 6.2%, albeit for a different time period. As we noted in the introduction to this paper, the equity premium, and hence the equity premium puzzle, continued to grow larger in the years after their paper was written. By extending the estimation period, and expanding our horizons to embrace the world, we have simply succeeded in reducing the puzzle back down to the magnitude documented in Mehra-Prescott's original paper. If 6.2% was a puzzle, it follows that 6.1% is only a very slightly smaller puzzle.

In terms of the empirical evidence, if we are to further shrink our estimate of the expected premium, two further possibilities remain. The first is that our world index is still upward biased because of survivorship bias in terms of the countries included. The second possibility relates to "good luck" and/or a systematic repricing of equities and their riskiness to investors over the last century. As we have seen, however, although the U.S. equity market has performed well, it was not a massive outlier. The challenge for the good luck/repricing hypothesis is thus to explain not just why the United States had "100 years of good luck", but why the rest of the world was almost as fortunate. In the next subsection, we assess the possible impact of survivorship bias. Section 6 then addresses the issues of good luck and repricing.

Survivorship of Markets

Several researchers, most notably Brown, Goetzmann, and Ross (1995) and Jorion and Goetzmann (1999), have suggested that survivorship bias may have led to overestimates of the historical equity premium. Li and Xu (2002) argue on theoretical grounds that this is unlikely to explain the equity premium puzzle, since, for survival models to succeed, the *ex ante* probability of long-term market survival has to be extremely small, which they claim contradicts the history of the world's financial markets. In this section, we look at the empirical evidence on returns and survivorship, and reach the same conclusion as Li and Xu, namely that concerns over survivorship are overstated, especially with respect to true survivorship bias, namely, the impact of markets that failed to survive.

In practice, however, the term "survivorship bias" is often used to also embrace *ex post* success bias as well as true survivorship bias. By comparing U.S. history with that of 16 other countries, we have already addressed the issue of success bias. While a legitimate concern, we are still left with a high historical 17-country world equity premium. Mehra (2003) has also noted that, with respect to its impact on the equity premium, success bias is partly mitigated by the tendency of successful markets to enjoy higher bond and bill returns, as well as higher equity returns; similarly, unsuccessful markets have tended to have lower real returns for both government securities and equities. In other words, there has been a positive correlation between real equity and real bill (or bond) returns.[14] Among markets with high *ex post* equity premiums there are naturally countries with excellent equity performance (like Australia); but there are also countries whose below-average equity returns nevertheless exceeded their disastrous bond returns (like Germany or Japan). Consequently, the cross-sectional dispersion of equity premiums is narrower than the cross-sectional dispersion of equity returns.

Our equity premiums are, of course, measured relative to bills and bonds. In a number of countries, these yielded markedly negative real returns, often as a result of periods of very high or hyperinflation. Since these "risk-free" returns likely fell below investor expectations, the corresponding equity premiums for these countries are arguably overstated. Even this is not clear, however, as equity returns would presumably have been higher if economic conditions had not given rise to markedly negative real fixed-income returns. Depressed conditions were a particular feature of the first half of the twentieth century, a period in which hyperinflations were relatively prevalent.[15] Had economic conditions been better, it is possible that the equity premium could have been larger. Similarly, it could be argued that in the more successful economies, the *ex post* bill and bond returns may, over the long run, have exceeded investors' expectations.

---

[14] Over the entire 106-year period, the cross-sectional correlation between the 17 real equity and 17 real bill (bond) returns was 0.63 (0.66). Measured over 106 individual years, the time-series correlations between real equity and real bill returns ranged from 0.01 in The Netherlands to 0.44 in Japan, with a 17-country mean correlation of 0.22, while the time-series correlations between real equity and real bond returns ranged from 0.11 in The Netherlands to 0.55 in the United Kingdom, with a 17-country mean correlation of 0.37.

[15] In our sample of countries over 1900–1949, the cross-sectional correlation between real equity and real bill (bond) returns was 0.68 (0.80). The time-series correlations between annual real equity and real bill (bond) returns had a 17-country mean of 0.31 (0.42).

We concluded above, therefore, that provided a very long run approach is taken, inferences from the United States do not appear to have given rise to very large overestimates of the historical world equity premium. It is still possible, however, that our world index overstates worldwide historical equity returns by omitting countries that failed to survive. The most frequently cited cases are those of Russia and China, whose equity markets experienced a compound rate of return of –100%.[16] However, there are other stock markets, apart from Russia and China, which we have so far been unable to include in our sample due to data unavailability.[17]

At noted earlier, at the start-date of our database in 1900, stock exchanges already existed in at least 33 of today's nations. Our database includes 17 of these, and we would ideally like to assess their importance in terms of market capitalization relative to the countries for which we have no data. Unfortunately, the required data are not available. Such aggregate data were neither recorded nor even thought of in 1900.[18] Rajan and Zingales (2003), however, do report a set of market capitalization to GDP ratios for 1913. By combining these with Maddison (1995) GDP data, coupled with some informed guesses for countries not covered by Rajan and Zingales, we can calculate approximate equity market capitalizations at that date.

Based on these estimates, it is clear that the 17 DMS database countries dominated the early twentieth century world equity market. The largest omitted market is Russia, which we estimate in those days represented just under 5% of total world capitalization. Next is Austria-Hungary, which then incorporated Austria, Hungary, the Czech Republic, Slovakia, Slovenia, Croatia, Bosnia, and parts of modern-day Ukraine, Poland, and even Italy (Trieste), and which accounted for some 2% of world capitalization. Data described in Goetzmann, Ukhov, and Zhu (2006) suggest that the Chinese equity market accounted for 0.4% of world equity market capitalization in 1900. In addition, there was a group of Latin American markets, including Argentina, Brazil, Mexico, and Chile that in total made up around 1½% of overall capitalization; and a number of small markets that total less than 1%.[19] In addition to Russia and China, several other exchanges from 1900 did not survive World War II and ended in disaster, notably those in Czechoslovakia (now the Czech Republic and Slovakia), Hungary, and Poland (though these three countries were not independent states in 1900, being part of the Russian and the Austria-Hungary empires). We believe that the DMS database accounted for 90% of world equity capitalization at the start of the twentieth century, and that omitted countries represented just 10%.

---

[16] It could be argued that the nationalization of corporations in Russia after the revolution of 1917 and in China after the communist victory in 1949 represented a redistribution of wealth, rather than a total loss. But this argument would not have been terribly persuasive to investors in Russian and Chinese equities at the time. It is possible, however, that some small proportion of equity value was salvaged in Russian and Chinese companies with large overseas assets, e.g., in Chinese stocks with major assets in Hong Kong and Formosa (now Taiwan).

[17] We are endeavouring to assemble total return index series over 1900-2005 for countries such as New Zealand, Finland, and Austria; and we believe that, in principle, series for Argentina, India, Hong Kong, and other markets might also be compiled.

[18] The few snippets of historical data that exist, e.g., Conant (1908) are expressed in terms of the nominal value of the shares outstanding rather than the total market value of the shares. Furthermore, figures are often given only for the total nominal value of all securities, rather than that of equities. For the U.S., U.K., and two other countries we have meticulously constructed market capitalization data from archival sources relating to individual stocks. But for many of the other markets, it is possible that even the disaggregated archive source data may not have survived from the end of the nineteenth century to the present time.

[19] The Latin American stock markets suffered several episodes of political and economic instability and hyperinflation; today, they account for some 1.15% of world market capitalization, which is roughly three-quarters of their weighting in 1913. The other markets, that in 1913 totalled less than 1% of world market capitalization, today account for some 2.3% of the world market; this group includes countries such as Egypt, Finland, Greece, Hong Kong (China), India, New Zealand, and Sri Lanka.

Survivorship Bias is Negligible

Our estimates of the equity premium are based on 17 surviving markets and, as noted earlier, ignore at least 16 non-surviving markets. To quantify the global impact of omitted markets, it is unnecessary to focus on individual markets as in Li and Xu (2002). We assume the annualized historical equity return for markets that survived for T years was $R_{survivors}$ and that for markets which are missing from the DMS database, it was $R_{omitted}$. Assume a proportion S of the worldwide equity market survived the entire period. Then the cumulative worldwide equity premium $ERP_{worldwide}$ is given by:

$$(1 + ERP_{worldwide})^T = [S\,(1 + R_{survivors})^T + (1\text{-}S)\,(1 + R_{omitted})^T] \,/\, [(1 + R_{riskfree})^T] \qquad \textbf{[1]}$$

where $R_{riskfree}$ is the riskfree interest rate for the reference country. An extreme assumption would be that all omitted markets became valueless, namely $R_{omitted} = -1$; and that this outcome occurred, for every omitted country in a single disastrous year, rather than building up gradually. The worldwide equity premium, incorporating omitted as well as surviving markets, would therefore be given by:

$$(1 + ERP_{worldwide}) = S^{1/T}\,(1 + R_{survivors}) \,/\, (1 + R_{riskfree}) = S^{1/T}\,(1 + ERP_{survivors}) \qquad \textbf{[2]}$$

where $ERP_{survivors}$ is the historical equity premium for markets that survived. In our case, we estimate the proportion of the world equity market capitalization that survived was at least S=0.9 and our time horizon is T=106 years. To account for the omission of markets that existed in 1900 but did not survive, we must therefore adjust the *ex post* equity premium of the 17-country world index using a factor of $S^{1/T} = 0.9^{1/106} = 0.999$. The survivorship bias in the estimated equity premium is therefore the following:

$$ERP_{survivors} - ERP_{worldwide} = (1 - S^{1/T})(1 + ERP_{survivors}) = (1 - 0.999)(1 + ERP_{survivors}) \approx 0.001 \qquad \textbf{[3]}$$

where the final approximation reflects the fact that $ERP_{survivors}$ is an order of magnitude below 1. We see that, at most, survivorship bias could give rise to an overstatement of the geometric mean risk premium on the world equity index by about one-tenth of a percentage point. If disappearance were a slower process, the index weighting of countries destined to disappear would have declined gradually and the impact of survivorship bias would have been even smaller. Similarly, if omitted markets did not all become valueless, the magnitude of survivorship bias would have been smaller still.

While there is room for debate about the precise impact of the bias arising because some, but not all, equity markets experienced a total loss of value, the net impact on the worldwide geometric mean equity premium is no more than 0.1%. The impact on the arithmetic mean is similar.[20] At worst, an adjustment for market survivorship appears to reduce the arithmetic mean world equity premium relative to bills from around 6.1% (see Table 3 above) to approximately 6.0%. Thus the equity premium puzzle has once again become smaller, but only slightly so.

[20] It is duplicative to derive this formally. The intuition involves disappearance of 10% of the value of the market over a century, which represents a loss of value averaging 0.1% per year.

6. DECOMPOSING THE HISTORICAL EQUITY PREMIUM

The conventional view of the historical equity premium is that, at the start of each period, investors make an unbiased, albeit inaccurate, appraisal of the end-of-period value of the stock market. Consequently, the *ex post* premium, averaged over a sufficiently long interval, is expected to be a relatively accurate estimate of investors' expectations. A key question is whether the historical premium may nevertheless be materially biased as a proxy for expectations because the past was in some sense unrepresentative. For instance, investors may have benefited from a century of exceptional earnings, or stock prices may have enjoyed a major, but non-sustainable, expansion in their valuation ratios. Our argument, which has some roots in Mehra and Prescott (1988), is that the historical equity premium may have beaten expectations not because of survivorship, but because of unanticipated success within the equity market. This analysis therefore draws on, and complements, Fama and French (2002), Ibbotson and Chen (2003), and Arnott and Bernstein (2003).

Unanticipated Success

To examine whether history may have witnessed exceptional earnings and/or expanding valuation ratios, consider how the stock market's past performance could, over multiple decades, be below or above expectations. The twentieth century opened with much promise, and only a pessimist would have believed that the next 50 years would involve widespread civil and international wars, the 1929 Crash, the great depression, episodes of hyperinflation, the spread of communism, conflict in Korea, and the Cold War. During 1900–1949 the annualized real return on the world equity index was 3.5%, while for the world excluding the U.S. it was just 1.5%. By 1950, only the most rampant optimist would have dreamt that over the following half-century, the annualized real return on world equities would be 9.0%. Yet the second half of the twentieth century was a period when many events turned out better than expected. There was no third world war, the Cuban missile crisis was defused, the Berlin Wall fell, the Cold War ended, productivity and efficiency accelerated, technology progressed, and governance became stockholder driven. As noted by Fama and French (2002), among others, the 9.0% annualized real return on world equities from 1950 to 1999 probably exceeded expectations.

In many countries valuation ratios expanded, reflecting—at least in part—reduced investment risk. Over the course of the twentieth century, the price/dividend ratio rose in all the DMS countries. Davis et al (2000) and Siegel (2002) report that for the U.S. over the period since the 1920s, the aggregate stock market price/earnings and price/book ratios also rose, and Dimson, Nagel and Quigley (2003) make similar observations for the U.K. In 1900 investors typically held a limited number of domestic securities from a few industries (Newlands (1997)). As the century evolved, new industries appeared, economic and political risk declined, closed- and open-ended funds appeared, liquidity and risk management improved, institutions invested globally, and finally, wealthier investors probably became more risk tolerant. Yet even if their risk tolerance were unchanged, as equity risk became more diversifiable, the required risk premium is likely to have fallen. These trends must have driven stock prices higher, and it would be perverse to interpret higher valuation ratios as evidence of an *increased* risk premium. Furthermore, insofar as stock prices rose because of disappearing barriers to diversification, this phenomenon is non-repeatable and should not be extrapolated into the future.

To unravel whether twentieth-century equity premiums were on balance influenced by exceptional earnings and expanding valuation ratios, we decompose long-term premiums into several elements. We use the fact that the historical equity premium is equal to the sum of the growth rate of real dividends, expansion in the price/dividend ratio, the mean dividend yield, and the change in the real exchange rate, *less* the risk-free real interest rate. As shown in Appendix 1, provided the summations and subtractions are geometric, this relationship is an identity.[21]

Decomposition of the Equity Premium

Table 4 reports these five components of the equity premium for each country. The first two columns show the growth rate of real dividends and the expansion in the price/dividend ratio. There is a widespread belief, largely based on the long-term record of the U.S. (Siegel (2002)), that nominal dividends can be expected to grow at a rate that exceeds inflation. In fact, only three countries have recorded real dividend growth since 1900 of more than 1% per year, and the average growth rate is –0.1%, i.e., the typical country has not benefited from dividends (or, in all likelihood, earnings) growing faster than inflation. Equally, there is the belief that superior stock market performance may be attributed to the expansion of valuation ratios. While there is some truth in this, it should not be overstated. Over the last 106 years, the price/dividend ratio of the average country grew by just 0.6% per year. Given the improved opportunities for stock market diversification, 0.6% seems a modest contribution to the historical equity premium.

Each country's real (local currency) capital gain is attributable to the joint impact of real dividend growth and expansion in the price/dividend ratio. Although the real capital gain is not reported explicitly in Table 4, note that only two countries achieved a real, local-currency capital gain of at least 2% per year: the U.S. (2.1%) and Sweden (3.6%). We should be cautious about extrapolating from these relatively large rates of capital appreciation to other markets around the world.

The middle column of Table 4 is the geometric mean dividend yield over the 106-year sample period. Averaged across all 17 countries, the mean dividend yield has been 4.5%, though it has been as large as 6.0% (in South Africa) and as low as 3.5% (in Switzerland). Interestingly, the countries whose mean dividend yield is closest to the cross-sectional average are Canada (4.5%) and the U.S. (4.4%). Drawing on Grullon and Michaely (2002) and Mauboussin (2006) to adjust for the impact of repurchases,[22] which are more important in the U.S. than elsewhere, that country's (adjusted) historical dividend yield rises to approximately 4.7%, which is just above the (unadjusted) 17-country average of 4.5%.

---

[21] Let G_{dt} be the growth rate of real dividends; G_{PDt} be the rate at which the price/dividend ratio has expanded; $Y_t = D_t / P_t$ be the dividend yield, the ratio of aggregate dividends paid during period t divided by the aggregate stock price at the end of period t; X_t be the change in the real exchange rate; and R_{ft} be the risk-free real interest rate. The geometric mean from period 1 through period t, denoted by boldface italic, is calculated like this for all variables: $(1 + \boldsymbol{Y}_t) = [(1 + Y_1)(1 + Y_2)\ldots(1 + Y_t)]^{1/t}$. Appendix 1 shows that the equity risk premium is given by: $(1 + \boldsymbol{ERP}_t) = (1 + \boldsymbol{G_{dt}})(1 + \boldsymbol{G_{PDt}})(1 + \boldsymbol{Y}_t)(1 + \boldsymbol{X}_t) / (1 + \boldsymbol{R_{ft}})$ where boldface italic indicates a t-period geometric mean.

[22] Since the 1980s, U.S. yields have been low relative to the past partly because, under prior tax rules, companies could return capital to shareholders more effectively on an after-tax basis by means of stock repurchases. From 1972–2000, Grullon and Michaely (2002) estimate that annual repurchases averaged 38.0% of cash dividends (57.5% from 1984–2000), while over 1977–2005, Mauboussin (2006) estimates the average to be 64.8%. Adding repurchases to the yield, the "adjusted dividend yield" for the U.S. rises from its raw historical average of 4.4% to 4.7%, whether we use the data from Grullon and Michaely (2002) or Mauboussin (2006). The impact of a similar adjustment to other countries' dividend yield is smaller and often zero (see Rau and Vermaelen (2002)).

Table 4: Decomposition of the Historical Equity Premium for 17 Countries, 1900–2005

% p.a. **Country**	 **Real dividend growth rate**	***plus**** **Expansion in the P/D ratio**	***plus*** **Geometric mean dividend yield**	***plus*** **Change in real exchange rate**	***minus*** **U.S. real interest rate**	***equals*** **Equity premium for U.S. investors**
Australia	1.30	0.46	5.83	-0.24	0.96	6.42
Belgium	-1.57	0.08	3.95	0.62	0.96	2.05
Canada	0.72	0.98	4.46	-0.04	0.96	5.18
Denmark	-0.87	1.43	4.68	0.47	0.96	4.74
France	-0.74	0.42	3.93	-0.14	0.96	2.47
Germany	-1.54	0.97	3.69	0.23	0.96	2.35
Ireland	-0.25	0.38	4.66	0.25	0.96	4.05
Italy	-1.46	-0.08	4.05	0.10	0.96	1.58
Japan	-2.39	1.59	5.39	0.32	0.96	3.85
Netherlands	-0.16	0.41	5.00	0.27	0.96	4.54
Norway	-0.25	0.50	4.02	0.25	0.96	3.54
South Africa	0.91	0.31	5.95	-0.80	0.96	5.38
Spain	-0.62	0.24	4.13	0.00	0.96	2.75
Sweden	2.88	0.67	4.09	-0.05	0.96	6.72
Switzerland	0.32	0.60	3.52	0.72	0.96	4.22
U.K.	0.61	0.18	4.68	-0.03	0.96	4.46
U.S.	1.32	0.75	4.36	0.00	0.96	5.51
Average	**-0.10**	**0.58**	**4.49**	**0.11**	**0.96**	**4.11**
Std deviation	**1.32**	**0.45**	**0.71**	**0.35**	**0.00**	**1.51**
World (USD)	0.77	0.68	4.23	0.00	0.96	4.74

* Note: Premiums are relative to bill returns. All summations and subtractions are geometric

To examine the equity premium from the perspective of a global investor located in a specific home country, such as the U.S., we convert from real, local-currency returns to real, common-currency returns. Taylor (2002) demonstrates that, over the very long term, exchange rate changes reflect purchasing power changes. It is unsurprising, then, to see that the annualized change in our 17 countries' real exchange rate averages only 0.1% per year, and that every country's real exchange rate change was within the range ±1%. Note that, for the average country, the capital gain in real U.S. dollars (the sum of the second, third and fifth columns) was just 0.6% per year (not reported in Table 4). Measured in real U.S. dollars, only two countries achieved a capital gain that exceeded 2% per year. Nine countries achieved a real U.S. dollar capital gain that was between zero and +2%; and six achieved between zero and –2%.

The annualized real, local-currency returns were reported for all countries in Table 1; across all 17 countries, the average 106-year return is 5.0%. The real, USD-denominated returns (the sum of the second to the fifth columns in Table 4) average 5.1%. Deducting the U.S. risk-free interest rate of 0.96% in real terms, the equity premium for a U.S. investor buying stocks in each of the 17 markets is as listed on the right of Table 4: on average the premium is 4.1%.

The *ex post* equity premiums on the right of Table 4 vary cross-sectionally for two reasons: the expected reward for risk, and the impact of chance. In 1900 the expected premium for higher risk markets may have merited a high reward that was subsequently realised; if Australia, Canada, South Africa and Sweden were such economies, they achieved relatively large *ex post* premiums of over 5%. The expected premium for safer markets may have been low; if these markets are typified by Belgium, France, Germany, Italy and Spain, their *ex post* premiums were below 3%. However, this rationalization is not a credible explanation for historical performance. It is more likely that, in 1900, investors underestimated the probability of wars in Europe, not to mention the ultimate value of resource-rich economies like the U.S. and Canada. National returns thus probably had more to do with noise than with the expected premium in 1900, and averaging mitigates the impact of noise. In projecting the equity premium into the future, we therefore focus on the equally weighted worldwide average of 4.1% and on the market-capitalization weighted world index. The world index is shown in the bottom-right corner of Table 4; from the point of view of a U.S. based investor, the world equity premium was 4.7%.[23]

From the Past to the Future

Over the long run, real returns accrued largely from dividend payments, but Dimson, Marsh and Staunton (2000, 2002), Arnott and Ryan (2001), and Ritter (2005) highlight the time-series and cross-sectional variation of global equity premiums. Given the large standard errors of historical estimates, and the likelihood that risks and equity premiums are nonstationary, one cannot determine a precise, forward-looking expected premium. However, by considering separately each component of the historical equity premium, we can develop a framework for making inferences. We start by discussing the real dividend growth rate, followed by expansion in the price/dividend ratio, and then the average dividend yield. We also consider changes in the real exchange rate.

The second column of Table 4 indicates that, over the last 106 years, real dividends in the average country fell by 0.1% per year; in the world index, they rose by +0.8%; and in the U.S., they rose by +1.3%. Siegel (2005) and Siegel and Schwartz (2006), among others, observe that these long-term dividend growth rates were not achieved by a cohort of common stocks. The growth is that of a portfolio whose composition evolved gradually; today it contains almost no stocks from 1900, and largely comprises companies that gained a listing subsequently.[24] In large part, the long-term increase in index dividends reflects companies that not only gained a listing after 1900, but ceased to exist quite some years ago.[25] So what real dividend growth can we anticipate for the future? The worldwide growth rate was 0.8% per year; relative pessimists might project real dividend growth that is zero or less (Arnott and Bernstein (2002)), while relative optimists might forecast indefinite real growth in excess of 1% (Ibbotson and Chen (2003)).

---

[23] We also computed the premium from the viewpoint of investors in the other 16 countries (for example, with a Japanese investor's premium based on every market's local-currency return converted into yen); the 17-country average equity premium varied between 2.3% for Denmark and 9.2% for Italy, with an average across all 17 reference currencies of 4.8%. Similarly, we computed the world premium from the viewpoint of investors in the other 16 countries (again converting every market's return into yen, and so on); the world equity premium varied between 2.9% for Denmark and 9.9% for Italy, with an average across all 17 reference currencies of 5.4%. This wide range of values is attributable mostly to differences in the annualized real risk-free rate between countries, rather than to exchange rate differences.

[24] To illustrate how much the listed equity market has evolved, Dimson, Marsh and Staunton (2002) report that almost two-thirds of the value of the U.S. market and half the value of the U.K. market was represented by railroad stocks at the end of 1899.

[25] There can also be a spurious jump in measured dividends when indexes are chain-linked. As a dividend series switches from narrower to broader composition, or from pre-tax to net-of-tax dividend payments, this can give rise to a step in income that impacts dividend growth estimates and (in the opposite direction) changes in the price/dividend ratio. We experimented with making adjustments for this for the U.S. and U.K. but the impact on estimated long-term dividend growth from splicing index series was small, and we abandoned this idea.

The third column of Table 4 reports that, over the last 106 years, the price/dividend ratio in the average country expanded by +0.6% per year; in the world and U.S. indexes it expanded by +0.7% and +0.8% respectively. As discussed earlier, this expansion reflected, at least in part, the enhanced opportunity to reduce portfolio risk as institutions increased the scope for diversification both domestically and internationally. If investors' risk tolerances are today similar to the past, we have already argued that the required risk premium is likely to have fallen and valuation ratios to have risen. There is no reason to expect the required risk premium to fall further over the long haul, so persistent multiple expansion seems unlikely. Without further expansion in the price/dividend ratio, this source of historical performance cannot contribute to forward-looking equity premiums.

The fourth column of Table 4 shows that, over the last 106 years, the geometric mean dividend yield in the U.S. was 4.4%, compared with 4.5% for the average country and 4.2% for the world index. Contemporary dividend yields (i.e., yields at end-2005, at the conclusion of the 106-year period) are lower than the historical average, even when buybacks are incorporated (see footnote 22 above). Whether adjusted for stock repurchases or not, projected levels for the long-term, geometric mean dividend yield are unlikely to be as large as the worldwide historical average of 4.2%. To the extent that the current (end-2005) level of dividends is indicative, the mean yield is likely to be lower in the future by at least ½–1%.

Over the long term, nominal exchange rates tend to follow fluctuations in relative purchasing power. The consensus forecast for changes over the long term in the real (inflation adjusted) exchange rate is zero. While the fifth column of Table 4 indicates that, historically, Americans gained (and others lost) from the rising real value of the U.S. dollar, this pattern cannot be extrapolated. We may assume that, over the long term, the real exchange rate change is expected to average zero.

The historical equity premium comprises the sum of the factors discussed in the preceding paragraphs, minus the real interest rate (see the penultimate column of Table 4). The final column of Table 4 reports the historical equity premiums for our 17 countries; they have an average of a 4.1% premium, with a cross-sectional standard deviation of 1.5%. While forward-looking estimates cannot be precise, a long-term projection of the annualized equity premium might, at the very least, involve making an adjustment to the historical record for components of performance that cannot be regarded as persistent. First, the expected change in the real exchange rate may be assumed to be zero, which implies an upward bias of 0.1% in the cross-sectional average of the country equity premiums. Second, the historical expansion in the price/dividend ratio cannot be extrapolated and might be assumed to be zero, which implies an upward bias of 0.6% in the cross-sectional average. These two adjustments, alone, attenuate the average country equity premium from 4.1% to 3.4%. When the same adjustments are made to the world index, the world equity premium shrinks from 4.7% to 4.0%. We noted above that if current dividend levels are a guide to the future, then the prospective mean dividend yield on the world index is likely to be lower than the historical average by at least ½–1%. This suggests a current equity premium of approximately 3–3½%.

Goyal and Welch (2006) conclude that for forecasting the equity risk premium one cannot do better than to project the historical average equity premium into the future, and Mehra (2003) contends that "over the long term, the equity premium is likely to be similar to what it has been in the past." However, as Campbell and Thompson (2005) point out, this cannot be the full story. History suggests that some part of the historical premium represents equity investors' good luck, and Fama and French (2002) say in relation to the period 1951–2000 that their "main message is that the unconditional expected equity premium…is probably far below the realized premium."

Jorion and Goetzmann (1999) justified estimating equity premiums from capital-appreciation indexes, stating "to the extent that cross-sectional variations in [dividend return minus real interest rate] are small, this allows comparisons of equity premiums across countries." They compared six markets with and without dividends, with similar conclusions, albeit over a sample period differing from the 1900-2005 interval used here. However, there is a cross-country standard deviation in dividend yields of 0.7% (see Table 4). If one computes the sum for each country of dividend yield plus dividend growth, the cross-sectional standard deviation is 1.6%. Our estimates of the equity premium avoid the inaccuracies that arise from the Jorion-Goetzmann approximation.

The debate on the size of the equity premium is sometimes conducted in terms of the arithmetic mean. For a stationary series the arithmetic mean is straightforward to interpret, but as Lettau and Nieuwerburgh (2006) highlight, the underlying parameters are unstable. This makes arithmetic means harder to interpret, which is why we undertake our decompositions using annualized returns.[26] For those who focus on the arithmetic mean equity premium, for the world index the latter is 1.3% larger than the geometric mean (see Table 3), and our forward-looking estimate of the arithmetic mean premium for the world index would be approximately 4½–5%.

Twentieth-century financial history was a game of two halves. In the first half, markets were harsh on equity investors; but in the second half they were benevolent.[27] As we show in Dimson, Marsh and Staunton (2002), early in the century dividend yields were mostly high relative to interest rates, whereas more recently yields have generally been lower. Looking at the 1900-2005 period as a whole, the world equity market experienced dividend growth and price/dividend multiple expansion that contributed 0.8% and 0.7% per year respectively to long-run real returns and hence to the *ex post* equity premium. The remainder was contributed by the annualized dividend yield of 4.2% (for the world index) and a real exchange rate adjustment. This suggests that the equity premium *expected* by investors was lower than the realized premium. The fact that *ex post* equity premiums were enhanced by this rate of dividend growth and multiple expansion is the "triumph" experienced by twentieth-century stock market investors.

---

[26] For example, consider a hypothetical index that provides a zero equity premium over a two-period interval. Assume that, within this interval, it suffers from transient volatility; for instance, the single-period returns might be +900% and –90%. Unless there is reason to suppose that volatility will persist at its historical level, the expected equity premium will be lower than the high arithmetic mean of +405% per period. In contrast with formerly turbulent countries like Germany, Italy and Japan, the U.S. and world indexes did not experience volatility on this scale—at least, not during the twentieth century.

[27] Averaged across all 17 countries, the real, local-currency annualised equity returns were 2.7% in the first half of the twentieth century, versus 7.1% over the following 55 years. Note, however, that adverse stock market conditions also tended to impact the real returns from bonds and bills (see section 5).

7. CONCLUSION

We have presented new evidence on the historical equity premium for 17 countries over 106 years. Our estimates, including those for the U.S. and U.K., are lower than frequently quoted historical averages. The differences arise from bias in previous index construction for the U.K. and, for both countries, our use of a longer time frame that incorporates the earlier part of the twentieth century as well as the opening years of the new millennium. Prior views have been heavily influenced by the U.S. experience, yet we find that the U.S. equity premium is somewhat higher than the average for the other 16 countries.

The historical equity premium, presented here as an annualized estimate (i.e., as a geometric mean), is equal to investors' *ex ante* expectations plus the impact of luck. In particular, expanding multiples have underpinned past returns. In part, this reflects a general decline in the risk faced by investors as the scope for diversification has increased, and stocks have become more highly valued. In addition, past returns have also been enhanced during the second half of the twentieth century by business conditions that improved on many dimensions.

We cannot know today's consensus expectation for the equity premium. However, after adjusting for non-repeatable factors that favoured equities in the past, we infer that investors expect an equity premium (relative to bills) of around 3–3½% on a geometric mean basis and, by implication, an arithmetic mean premium for the world index of approximately 4½–5%. These estimates are lower than the historical premiums quoted in most textbooks or cited in surveys of finance academics. From a long-term historical and global perspective, the equity premium is smaller than was once thought. The equity premium survives as a puzzle, however, and we have no doubt that it will continue to intrigue finance scholars for the foreseeable future.

APPENDIX 1: DECOMPOSITION OF THE EQUITY PREMIUM

This appendix explains how we decompose the historical equity premium into five elements. These are, firstly, the average dividend yield over the sample period; next, the impact of real dividend growth, expansion of the price/dividend ratio, and the change in the real exchange rate; and finally, the risk-free interest rate that is used to compute the equity premium. Without loss of generality, the decomposition is in real (inflation adjusted) terms.

Capital Appreciation and Income

We assume the dividend payment on the equity index portfolio is received at the end of period t and is equal to D_t, that the price at the end of period t-1 is P_{t-1}, and that inflation over period t runs at the rate I_t.

Real dividends are $d_t = D_t / (1 + \boldsymbol{I}_t)^{t}$, where the denominator measures the inflation rate from period 1 to period t, namely $(1 + \boldsymbol{I}_t)^{t} = (1 + I_1)(1 + I_2)\ldots(1 + I_t)$. The price/dividend ratio is $PD_t = P_t / D_t$. The real capital gain over period t is given by:

$$
\begin{aligned}
1+\text{Real gain}_t &= (P_t / P_{t-1}) / (1 + I_t) \\
&\equiv [(D_t / D_{t-1}) / (1 + I_t)]\ (PD_t / PD_{t-1}) \\
&= (d_t / d_{t-1})(PD_t / PD_{t-1}) \\
&= (1 + G_{dt})(1 + G_{PDt}) \qquad \textbf{[A1]}
\end{aligned}
$$

where the growth rate of real dividends is $G_{dt} = d_t / d_{t-1} - 1$, and the rate at which the price/dividend ratio has expanded is $G_{PDt} = PD_t / PD_{t-1} - 1$.

As a proportion of the initial investment, real dividend income during period t is:

$$
\begin{aligned}
\text{Real income}_t &= (D_t / P_{t-1}) / (1 + I_t) \\
&\equiv (D_t / P_t)(P_t / P_{t-1}) / (1 + I_t) \\
&= Y_t (P_t / P_{t-1}) / (1 + I_t) \qquad \textbf{[A2]}
\end{aligned}
$$

where $Y_t = D_t / P_t$ is the dividend yield, defined as the ratio of aggregate dividends paid over period t divided by the aggregate stock price at the end of period t. Note that the terms to the right of Y_t measure (one plus) the real capital gain over period t, as defined above.

Total Returns

The real return is equal to the arithmetic sum of [1] real capital gain and [2] real income, namely:

$$
\begin{aligned}
1+\text{Real return}_t &\equiv [D_t / P_{t-1} + (P_t / P_{t-1})] / (1 + I_t) \\
&= (1 + G_{dt})(1 + G_{PDt})(1 + Y_t)
\end{aligned}
$$

So far we have decomposed returns denominated in a single currency. If the assets are purchased in unhedged foreign currency, we assume that each period's return is converted from foreign currency into home currency. The real return is then:

$$1+ \text{Real return}_t = (1 + G_{dt}) (1 + G_{PDt}) (1 + Y_t) (1 + X_t) \quad \textbf{[A3]}$$

where X_t is the increase in the inflation-adjusted value of the home currency relative to the foreign currency, namely the change in the real exchange rate.[28]

The Equity Premium

Finally, we define the equity premium as the geometric difference between the real return defined in [3] and the risk-free real interest rate, R_{ft}. Hence the historical equity premium is:

$$1 + ERP_t = (1+ \text{Real return}_t) / (1 + R_{ft})$$

$$= (1 + G_{dt}) (1 + G_{PDt}) (1 + Y_t) (1 + X_t) / (1 + R_{ft}) \quad \textbf{[A4]}$$

The historical equity premium is therefore equal to the sum of the real dividend growth rate, expansion in the price/dividend ratio, the dividend yield, and the change in the real exchange rate; less the risk-free real interest rate. All additions and subtractions are geometric.

Consequently, the geometric mean equity premium from period 1 through period t may be decomposed as follows:

$$1 + \boldsymbol{ERP}_t = (1 + \boldsymbol{G_{dt}}) (1 + \boldsymbol{G_{PDt}}) (1 + \boldsymbol{Y}_t) (1 + \boldsymbol{X}_t) / (1 + \boldsymbol{R_{ft}}) \quad \textbf{[A5]}$$

where each term on the right hand side of [5] is the geometric mean of t single-period components. That is, $(1 + \boldsymbol{Y}_t)^t = (1 + Y_1) (1 + Y_2)\ldots(1 + Y_t)$, and so on.

To sum up, the annualized historical equity premium may be decomposed geometrically into five elements. These are as follows: firstly, the mean growth rate in real dividends; secondly, the mean rate of expansion in the price/dividend multiple; thirdly, the mean dividend yield; fourthly, the mean change in the real exchange rate; and finally, the mean risk-free real interest rate.

Finally, note that the reference country for the real exchange rate and the real interest rate must correspond. For example, the exchange rate may be relative to the U.S. dollar; and if so, the real interest rate should be the rate on the U.S. risk-free asset.

[28] Obviously, when the investment is in domestic securities, the change in the real exchange rate is $X_t = 0$.

APPENDIX 2: DATA SOURCES FOR THE DMS DATABASE

Section 3 outlined the general methodology and guiding principles underlying the construction of the DMS database (see also Dimson, Marsh, and Staunton (2002, 2006a, and 2006b)). This appendix describes the data sources used for each country.

Australian equities are described in Officer's chapter in Ball, Brown, Finn, and Officer (1989). Ball and Bowers (1986) provide a complementary, though brief, historical analysis. We are grateful to Bob Officer for making his database available to us. Officer compiled equity returns from a variety of indexes. The early period made use of data from Lamberton's (1958) classic study. This is linked over the period 1958–74 to an accumulation index of fifty shares from the Australian Graduate School of Management (AGSM) and over 1975–79 to the AGSM value-weighted accumulation index. Subsequently, we use the Australia All-Ordinary index. Bond returns are based on the yields on New South Wales government securities from the start of the century until 1914. For the period 1915–49 the yields were on Commonwealth Government Securities of at least five years maturity. During 1950–86 the basis is ten-year Commonwealth Government Bonds. From 1986 we use the JP Morgan Australian government bond index with maturity of over seven years. For 1900–28 the short-term rate of interest is taken as the three-month time deposit rate. From 1929 onward we use the Treasury bill rate. Inflation is based on the retail price index (1900–48) and consumer price index (1949 onward). The switch in 1966 from Australian pounds to Australian dollars has been incorporated in the Exchange Rate index history.

Belgium is being researched by Annaert, Buelens, de Ceuster, Cuyvers, Devos, Gemis, Houtman-deSmedt, and Paredaens (1998). We are grateful for access to their interim results for 1900–28, which are subject to correction. From 1929 we use the National Bank of Belgium's 80-share index. The market was closed from August 1944 to May 1945, and we take the closing level for 1944 as the year-end value. For 1965–79 we use the Banque Bruxelles Lambert 30 share index and from 1980 the Brussels Stock Exchange All Share Index. Up to 1956, bond returns are based on estimated prices for 4% government bonds. During the 1944–45 closure, we take the last available value from 1944 as the year-end level. Over 1957–67 the index is for bonds with a five to twenty year maturity, for 1968–85 for bonds with maturity over five years. Subsequent years use the JP Morgan Belgian government bond index with maturity of over five years. Short-term interest rates are represented over the period 1900–26 by the central bank discount rate, followed during 1927–56 by the commercial bill rate. From 1957 onward, we use the return on Treasury bills. Inflation is estimated for 1900–13 using the consumer price index, and for 1914 we take the French inflation rate. Over 1915–20 and 1941–46 we interpolate the Belgian consumer price index from Mitchell (1998). From 1921 inflation is measured using the Institut National de Statistique's consumer price index.

Canadian stocks, bonds, bills, and inflation since 1924 are presented in Panjer and Tan (2002), with supplementary data kindly compiled for us by Lorne Switzer. For 1900–14 the annual index returns are based on Switzer's equally weighted (2000) Montreal index, adjusted for dividends. The equity series for 1915–46 is taken from Urquhart and Buckley (1965). Houston (1900–14) provides dividends for 1900 and hence the Canadian yield premium relative to the 1900 S&P, and Panjer and Tan (2002) estimate the Canadian yield relative to the 1924 S&P. To compute yearly total returns over 1900–23, we interpolate the Canadian yield premium relative to the S&P. For the period 1947–56 returns are for the TSE corporates, and from 1957 the TSE 300 total return index. The bond index for 1900–23 is based on a 4% bond from Global Financial Data (GFD). For 1924–36 we use the Government of Canada long bond index from Panjer and Tan (2002). Starting in 1936 the index is the Cansim index of bonds with maturity of over ten years, switching in 2002 to the JP Morgan Canadian government bond index with maturity of over ten years. For 1900–33 the short-term rate is represented by U.S. Treasury bills or equivalent. From 1934 onward the short-term rate is based on Canadian Treasury bills. Inflation is measured using the Canadian wholesale price index for 1900–10. For 1911–23 we switch to the Canadian consumer price index, and thereafter consumer price inflation is taken from Cansim.

Danish stock market data has involved working with Claus Parum to extend his research back to 1900. We have also referred to the papers by Steen Nielsen and Ole Risager (1999, 2000) and Allan Timmermann (1992). Over the period 1900–14 we use Parum's (2002) equally weighted index of equity returns, which covers some forty to fifty constituents each year. Thereafter, all the studies cited above are based on equity price indexes from Statistics Denmark, though we incorporate Parum's adjustments for capital changes that are not incorporated into the published index numbers. For 1915–2001 we use the data compiled in Parum (1999a,b and 2002) switching from 2002 to the Copenhagen KAX Index. Danish bond returns are estimated from yields on government bonds until 1924. For 1925–2001 our data is from Parum (1999a,b and 2002) who uses the return on mortgage bonds, a large and liquid asset class throughout the period, in contrast to more thinly traded government bonds, as described in

Christiansen and Lystbaek (1994). From 2002 we use the JP Morgan Danish government bond index with maturity of over seven years. Short-term interest rates are represented by the central bank discount rate until 1975, and thereafter by the return on Treasury bills.

France is documented by Laforest (1958) then Laforest and Sallee (1977), for the first half of the twentieth century, followed by Gallais-Hamonno and Arbulu (1995) for the period commencing in 1950. The common basis for equity returns in all the primary studies is the index series compiled by the Institut National de la Statistique et des Etudes Economiques (INSEE). The INSEE equity index is a weighted average of price relatives with about three hundred constituents. Over the period from 1914-18 we interpolate, assuming constant real returns. We use the SBF-250 from 1991 onward. The bond series for France, also compiled by INSEE, is based on consol yields. Over the period from 1914-18 we interpolate, assuming constant nominal returns. We switch in 1950 to the Gallais-Hamonno and Arbulu (1995) series, which is the INSEE General Bonds Index, with coupons reinvested monthly as received. From 1993 we use the JP Morgan French government bond index with maturity of over ten years. The short-term interest rate for France is based on the central bank discount rate until 1930. The rate is measured by the return on Treasury bills starting in 1931. To measure consumer price inflation, we use the consumption price index that is compiled by the Institut National de la Statistique et des Etudes Economiques, taken from Laforest (1958), Gallais-Hamonno and Arbulu (1995) and directly since 1981.

German data was provided by George Bittlingmayer (1998) and Richard Stehle (1997); also see Stehle, Wulff, and Richter (1999),and also Gregor Gielen (1994) and Ulrich Ronge (2002). We use Ronge's reconstruction of the DAX 30 share index to provide nominal equity returns for 1900-53. For August 1914–October 1918 Ronge uses the Gielen over-the-counter index. For 1954–94 we use the Stehle (1997) comprehensive index, switching in 1995 to the CDAX as given in Stehle/Hartmond-Reihe. For 1900–23, German bond returns are based on the price of 3% perpetuals, which essentially lost all value during the 1922–23 hyperinflation. For 1924–35 the bond index is based on mortgage bonds, and for 1936–51 it is based on 4.5% conversion (to 1943), 4.5% western zone (1946–47) and 5% tax-free (from 1948) bonds. We use the REX performance index starting in 1968, switching in 1986 to the JP Morgan German government bond index with maturity of over seven years. The short-term rate of interest is represented by the discount rate on private bills through 1945. We assume rates of 2% during 1946–50, 3% for 1951–53, and use Treasury bills beginning in 1954. Inflation in Germany is from Gielen (1994), using consumer price level data from the Imperial Statistical Office (see Bittlingmayer (1998)). Inflation rates during 1922 and 1923 were inferred from exchange rates against the dollar. From 1993 we use the CPI from the Federal Statistical Office.

Ireland was first studied by Shane Whelan (1999), who used Irish Central Statistical Office (CSO) data from 1934, and British data before that. Thomas (1986) provides some additional early data, but only in graphical form. We therefore created a new, market capitalization-weighted index of Irish equity prices for 1900–33 from original archive stock price and dividend sources (and this index has now been adopted by Whelan (2002)). For 1934–83 we use the Irish CSO Price Index of Ordinary Stocks and Shares. Until 1987, we incorporate our estimates of U.K. dividend yields. From 1988 we use the Irish Stock Exchange Equity (ISEQ) total return index. The bond series for Ireland uses U.K. returns for 1900–78. For 1979–98, we use Whelan's (1999) return on a twenty-year representative Irish gilt, as estimated by Raida Stockbrokers, turning thereafter to the Datastream ten-year Irish government bond index. Short-term Irish interest rates again use U.K. Treasury bills for 1900-1969. From 1970 we use Irish Treasury bills. Up to the date of political independence from Britain, inflation is measured using Bowley's (1937) cost of living index for 1900–13 and the working-class cost of living index for 1914–22. For 1923–52 we use Meghen's (1970) Irish cost of living index, and from 1953, the Irish consumer price index.

Italian data was provided by Fabio Panetta and Roberto Violi (1999). The equity data for 1900–07 are from the Official List and supplementary sources, and this is extended through 1911 with data from Aleotti (1990). From 1912–77 the share price and dividend series are based on the Bank of Italy index, which covers at least three-quarters of the total market capitalization of the Italian equity market. Thereafter, the Bank of Italy's index is calculated from the bank's monthly share price database, which covers all listed shares. From 1999 onward, we use the Milan BCI performance index. The government bond returns over 1900–44 are from Bianchi (1979). For the period 1945–83, the index of total bond returns is based on a treasury bond index with a coverage of over half, and often over three-quarters, of the value of all treasury bonds in issue. Thereafter, the data are sourced from Panetta and Violi's (1999) study. From 1988, we use the JP Morgan Italian government bond index with maturity of over three years. The short-term bank deposit rate to 1940 is from Biscaini Cotula and Ciocca (1982). Panetta and Violi estimate the values for the period 1941–46, and for 1947–61 the figures are from the Bank of Italy's Bollettino Economico. After that, the source is the Bank of Italy's Bollettino Statistico.

Japanese data of good quality are available from the Hamao (1991) database, and from the study by Schwartz and Ziemba (1991). We are grateful to Kenji Wada for facilitating provision of pre–World War I equity data. For 1900–14 we use the Laspeyres price index for the Tokyo Stock Exchange (TSE), as published in Fujino and Akiyama (1977). Thereafter, share prices are represented by the Japan National Bank index for 1915–32; the Oriental Economist Index from 1933 until September 1948 (although trading was suspended in August 1945, and no index values were published again until May 1946 when black market trading resumed in Tokyo); the Fisher index from September 1948 until the market officially reopened in May 1949; and the Nikkei-225 from May 1949 to 1951. During 1952–70 we use the Japan Securities Research Institute total return index. From 1971 we use total returns from Hamao and Ibbotson (1989). Returns continue from 1995 with the TSE TOPIX index. The Japanese government bond index data is taken from Global Financial Data. Until 1957, the returns are estimated from yield data. No yield information is available for the end of 1947, and the yield for 1946 is used instead. The data for 1948–57 represent the yields on newly issued bonds. From 1957 through 1968, the bonds are those issued by Nippon Telephone and Telegraph. From 1971 we use the government bond index from Hamao and Ibbotson (1989), followed from 1995 by the JP Morgan Japanese government bond index with maturity of over ten years. The short-term riskless rate is available from 1900. It is based on call money rates to 1959, and on Treasury bills thereafter. Inflation is measured by the wholesale price index for 1900, the retail price index for 1901–46 and the consumer price index from 1947 onward.

The Netherlands is based on work by Eichholtz, Koedijk, and Otten (2000). The equity returns over 1900–18 are based on the Central Bureau of Statistics (CBS) general index of share prices, and historical yield data. For the period 1919–51 returns are based on the 50-stock, CBS weighted arithmetic index. The exchange was closed from August 1944 to April 1946, so the end-year index levels are represented by the intra-year values that are closest to the turn of the year. During 1952–80, returns are based on the CBS All Share index, with dividends estimated by the Dutch central bank. For 1981 onward we use the CBS total return index, which went live in 1989 with retrospective estimation of the impact of income reinvestment, changing to the Amsterdam AMS All Share index from 2004. During 1900–14, Dutch bond returns are represented by 2.5% and 3% consols. During 1915–73, the Eichholtz-Koedijk-Otten bond index is based on a series of 3.5% bonds. From 1974, the index is the JP Morgan Netherlands government bond index with maturity of over seven years. For the riskless rate, during 1900–40 we use the discount rate on three-month private bills. The rate is assumed unchanged when data were unavailable during August 1914 to December 1918, and from mid-May 1940 to the end of that year. From 1941 to date we use the rate on Dutch Treasury bills. Inflation is measured using the consumer price index. No data were available between August 1944 and June 1945, and the index was interpolated for end-1944.

Norway was introduced into the study through Thore Johnsen, Knut Kjær and Bernt Ødegaard who provided data and sources. Equity returns for 1900–17 are derived from an equally weighted index based on all stocks listed in Statistisk Arbok and supplemented with those shares listed in Kierulf's Handbook for which there was information on year-end prices and dividends. The index contained between 33–36 shares until the end of 1914, but this fell to 21 by the start of 1918. For the period 1918–72 we use an all-share index including industrial, banking and whaling/shipping shares calculated by Statistics Norway. From 1973 we use a comprehensive index compiled by Thore Johnsen, switching in 1981 to the Oslo Stock Exchange indexes. We first use the Industrial index, switching in 1983 to the General Index and then, from 1996, to the All Share index. During 1900–92 Norwegian bond returns are based on Global Financial Data's government bond yields. From 1993, the index is the Datastream government bond index with maturity of ten years. For the riskless rate, during 1900–71 we use the central bank discount rate, followed by money market rates until 1983. From 1984 to date we use the rate on Norwegian Treasury bills. Inflation is measured using the consumer price index published by Statistics Norway.

South African stocks, bonds, bills, and inflation since 1925 are presented in Firer and McLeod (1999) who, in turn, draw on earlier work going back to 1910 by Schumann and Scheurkogel (1948). These studies provide indexes for industrial and commercial companies in South Africa. However, mining and financial companies are of particular importance, especially early last century. We therefore create a market capitalization weighted index of mining and financial shares for 1900–59, based on London price quotations. We blend our mining and financial indexes with the Firer and McLeod industrial index, by starting with a weighting of 5% in the industrial index at the start of 1910, with weights increasing to 25% by the start of 1950. From 1960–78 we use the Rand Daily Mail Industrial Index and, from 1979, the Johannesburg Stock Exchange–Actuaries Equity Index. Up to 1924, bond returns are based on the yields for 4% government bonds. Subsequently we use the bond returns from Firer and McLeod, based first on market yields together with a notional twenty-year bond prior to 1980, followed by the JSE-Actuaries Fixed Interest Index (to 1985), the JSE-Actuaries All Bond Index (to 2000) and the BESA Government total return index from 2001 onward. Before 1925, short-term interest rates are represented by U.K. Treasury bills.

Subsequently, we use the bill returns from Firer and McLeod, based on three-month fixed deposits (1925–59), bankers' acceptances (1960–66), and thereafter negotiable certificates of deposits. Inflation is estimated prior to 1925 using the consumer price index and thereafter using the official price index from Central Statistical Services. The switch in 1961 from pounds to rand has been incorporated in the Exchange Rate index index history.

Spanish stock returns are presented in Gonzalez and Suarez (1994) for the period commencing in 1941. Valbuena (2000) provides a longer-term perspective. Valbuena's equity index for Spain over 1900–18 is from Bolsa de Madrid. For 1919–36 we use a total returns index from Valbuena (2000) that rectifies some problems in the Sandez and Benavides (2000) index. Trading was suspended during the Civil War from July 1936 to April 1939, and the Madrid exchange remained closed through February 1940. Over the closure we assume a zero change in nominal stock prices and zero dividends. During 1941–85 we use the Gonzalez and Suarez (1994) data, subsequently linking this to the Bolsa de Madrid total return index. The bond series for 1900–26 is based on the price of Spanish 4% traded in London through 1913 and in Madrid thereafter. For 1926–57 and 1979-87 it is based on Global Financial Data's (GFD) estimates for government bonds, with prices kept unaltered during the Civil War. A private bond index is used for 1958–78. From 1988 we use the JP Morgan Spanish government bond index series with maturity of over three years. The short-term interest rate over 1900–73 is the central bank discount rate. From 1974 we use the return on Treasury bills. Inflation during 1900–14 is measured using the wholesale price index from Mitchell (1998). For 1915–35 we use the consumer price index from Mitchell (1998); see also Vandellos (1936). During 1936–40 we revert to the wholesale price index from Mitchell. For 1941–85 we use the Spanish consumer price index from Gonzalez and Suarez (1994) and thereafter from the Instituto Nacional de Estadistica.

Sweden is studied in a series of papers by Per Frennberg and Bjorn Hansson's (1992a, 1992b, 2000) whose database on stocks, bonds, bills, and inflation covers the period 1919–99. The Swedish stock market data we use starts at the end of 1900, and we assume that stock prices did not move over 1900; thereafter we use the index values of the Swedish Riksbank. Over the period 1900–18, Swedish equity dividends are estimated from contemporaneous bond yields adjusted upwards by 1.33% (the mean yield premium over 1919–36). From the start of 1919, the Swedish equity series is based on the share price index published in the journal Affarsvarlden, plus the dividend income estimated by Frennberg and Hansson (1992b). The government bond series uses data for 1900–18 from The Economist. For 1919–49 the returns are for perpetuals, and after that the series measures the return on a portfolio of bonds with an average maturity of ten years. We use the JP Morgan Swedish government bond index with maturity of over five years from 2000. The short-term riskless rate of interest from 1900 is represented by the official discount rate of the Swedish Riksbank. Frennberg and Hansson (1992b) switch in 1980 to the return on short-term money market instruments, and from 1982 to Treasury bills. Inflation is represented by the Myrdal-Bouvin consumer price index before 1914, the cost of living index between 1914-54 and the Swedish consumer price index for 1955 onward.

Switzerland is investigated using the series spliced together by Daniel Wydler (1989, 2001) coupled with extra data kindly provided by Urs Walchli and Corina Steiner. We have created a new, equally-weighted index of Swiss equity prices for 1900-10. This used the series of annual prices and dividend yields collected from Neue Zurcher Zeitung, with an average of 66 year-end stock prices over the period. Over 1911–25 we use the index of 21 industrial shares from Statistiches Jahrbuch. The Swiss exchanges were closed during September 1914 to December 1915, so for end-1914 and end-1915 we use the index at the date closest to the year-end. For 1926–59 Ratzer (1983) estimates total returns. For 1960–83 Huber (1985) computes the returns from index levels and dividends on the SBC index. Over 1984–98 we use the Pictet return index, and then the Swiss All Share index. For Switzerland only, and solely for the period 1900–15, we estimate bond returns from the short rate. We use the latter as a proxy for the yield on seven-year bonds, and infer the annual returns for this series. For 1915–25 we use annual data from the Statistischen Bureau. The interval 1926–59 employs Ratzer's (1983) estimates based on redemption yields for new Swiss bond issues. The 1960–80 period is represented by Huber's (1985) bond index based on actual trading prices. From 1981 we use the Datastream ten-year Swiss government bond index. During 1900–55 short-term rates are represented by the central bank discount rate, and for 1956–79, by the return on three-month time deposits. From 1980 onward, we use the return on Treasury bills. Nominal returns are adjusted for inflation using movements in the Swiss consumer prices index.

The United Kingdom is analysed using index series described in Dimson and Marsh (2001) for the interval from 1955 to date, and in Dimson, Marsh, and Staunton (2002, 2006a) for the period 1900–1954. Because of biases and inaccuracies in prior index series, the last half-century is based on the fully representative record of equity prices maintained by London Business School and described in Dimson and Marsh (1983). The period up to the end of 1954 is based on an index of the returns from the 100 companies that, before each New Year, have the largest equity market capitalization. Share capital was checked against the annual Stock Exchange Official Yearbook up to 1955, to account for capital changes and corporate events. Before 1955, all cash flows are assumed to occur at the end of each year, including dividends, special dividends, returns of capital, and cash from acquisitions. Where companies are acquired for shares or merge, we base returns on the end-year share price of the acquirer or merged entity, taking account of the exchange ratio. Dividends were obtained from the Stock Exchange Ten-Year Record published by Mathiesons. The U.K. bond index was compiled from original British government bond data. For the 1900–54 period the returns are based on 2½% Consols, and for 1955–2000 the bond index measures the return on a portfolio comprising high-coupon government bonds with a mean maturity of twenty years. Throughout the century, Treasury bills are used to measure the short-term riskless rate of interest. Inflation is calculated using the retail price index and, before 1962, the index of retail prices.

The United States was first researched in the Ibbotson and Sinquefield (1976) article and subsequent Ibbotson Associates updates. The broadest index of U.S. stock market returns is in Wilson and Jones (2002), and we use the latter for this study. Earlier sources are described in Goetzmann, Ibbotson, and Peng (2001). Our series, however, commences with the Wilson-Jones index data over 1900–25. For 1926–61 we use the University of Chicago's Center for Research in Security Prices (CRSP) capitalization-weighted index of all New York Stock Exchange stocks. For 1962–70 we use the CRSP capitalization-weighted index of NYSE, American, and Nasdaq stocks. From 1971 onward we employ the Dow Jones Wilshire 5000 index. All indexes include reinvested dividends. The government bond series for 1900–18 is based on 4% government bonds. Over 1919–25 we use the Federal Reserve ten-to-fifteen year bond index. After that bond returns are based on Ibbotson Associates' long bond index. The bill index uses commercial bills during 1900–18. From 1919 onward, the series is based on U.S. Treasury bills. Inflation is based on the consumer price index.

The World is represented by an equity series that comprises a 17-country, common-currency (here taken as U.S. dollars) index. For each period, we take a market's local-currency return and convert it to U.S. dollars. We therefore have the return that would have been received by a U.S. citizen who bought foreign currency at the start of the period, invested it in the foreign market throughout the period, liquidated his or her position, and converted the proceeds back at the end of the period into U.S. dollars. We assume that at the beginning of each period our investor bought a portfolio of 16 such positions in each of the foreign markets in this study, plus domestic equities, weighting each country by its size. We use GDP weights with start-decade rebalancing before 1968 due to a lack of reliable data on capitalizations prior to that date. Thereafter, we use country capitalizations taken from Morgan Stanley Capital International (MSCI). The above procedure results in an index expressed in U.S. dollars. To convert this to real terms, we then adjust by the U.S. inflation rate. This gives rise to a global index return denominated in real terms, from the point of view of our notional U.S. investor. Our 17-country world bond market index is constructed in the same way. This is again weighted by country size, to avoid giving, say, Belgium the same weight as the United States. Equity capitalization weights are inappropriate here, so the bond index is GDP-weighted throughout. The short-term risk free rate is taken as the return on U.S. Treasury bills. The inflation rate is as for the United States.

REFERENCES

A. Aleotti, *Borsa e Industria 1861–1989: Cento Anni di Rapporti Difficili*. Milan: Edizioni Comunita (1990).

J. Annaert, F. Buelens, J. de Ceuster, L. Cuyvers, G. Devos, M. Gemis, H. Houtman-deSmedt, and J. Paredaens, “Ontwerp van een Databank m.b.t. het Archief van de Beurs van Brussel”, TEW Working paper 98:11, Antwerpen: RUCA, (1998).

R.D. Arnott and P. Bernstein, “What Risk Premium Is "Normal"?”, *Financial Analysts Journal*, Vol. 58, No. 2 (2002).

R.D. Arnott and W.J. Bernstein, “Earnings Growth: the 2% Dilution”, *Financial Analysts Journal*, Vol. 58, No. 5 (2003), pp. 47–55.

R.D. Arnott and R. Ryan, “The Death of the Risk Premium: Consequences of the 1990s”, *Journal of Portfolio Management*, Vol. 27, No. 3 (2001), pp. 61–74.

Barclays Capital, *Equity-Gilt Study* (1999).

R. Ball and J. Bowers, “Shares, Bonds, Treasury Notes, Property Trusts and Inflation: Historical Returns and Risks 1974–1985”, *Australian Journal of Management*, Vol. 11 (1986), pp. 117–137.

R. Ball, P. Brown, F. J. Finn, and R. R. Officer, *Share Markets and Portfolio Theory,* second edition. St Lucia, Queensland: University of Queensland Press (1989).

B. Bianchi, “Appendice Statistica: Il Rendimento del Consolidato dal 1862 al 1946”, in F. Vicarelli (Ed) *Capitale Industriale e Capitale Finanziaro: Il Caso Italiano.* Bologna: Il Mulino (1979).

A.M. Biscaini, F. Cotula and P.L. Ciocca, (1982), “Italian Financial Structures: Long-term Quantitative Aspects (1879–1970)”, in G. Federico (Ed) *The Economic Development of Italy Since 1870*. Aldershot: Elgar (1994).

G. Bittlingmayer, “Output, Stock Volatility and Political Uncertainty in a Natural Experiment: Germany 1880–1940”, *Journal of Finance*, Vol. 53 (1998), pp. 2243–2257.

Z. Bodie, “Longer Time Horizon ‘Does Not Reduce Risk’”, *Financial Times*, January 26, 2002.

Z. Bodie, A. Kane, and A.J. Marcus, *Investments,* fourth edition. NY: McGraw Hill (1999).

A.L. Bowley, *Wages and Income in the United Kingdom Since 1860.* Cambridge: Cambridge University Press (1937).

R.A. Brealey, R. Giammarino, E. Maynes, S.C. Myers, and A.J. Marcus, *Fundamentals of Corporate Finance* (first Canadian edition). NY: McGraw Hill (1996).

R.A. Brealey and S.C. Myers, *Principles of Corporate Finance* (sixth edition). NY: McGraw Hill (2000).

S. Brown, W. Goetzmann, and S. Ross, “Survival”, *Journal of Finance*, Vol. 50 (1995), pp. 853–73.

J.Y. Campbell and S.B. Thompson, “Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?”, NBER Working Paper 11468 (July 2005).

J. Christiansen and B. Lystbaek, “Afkast og Risiko pa Aktier og Obligationer 1915–1993”, *Finans Invest*, Vol. 3 (1994), pp. 10–13.

J.H. Cochrane, “Where is the Market Going? Uncertain Facts and Novel Theories”, *Economic Perspectives*, Vol. 21 (1997), pp. 3–37.

Credit Suisse First Boston, *The CSFB Equity-Gilt Study* (1999).

J.L. Davis, E.F. Fama, and K.R. French, “Characteristics, Covariances and Average Returns: 1929–19997”, *Journal of Finance*, Vol. 55, pp. 389–406.

E. Dimson and P. Marsh, “The Stability of U.K. Risk Measures and the Problem of Thin Trading”, *Journal of Finance*, Vol. 38 (1983), pp. 753–783.

E. Dimson and P. Marsh, “U.K. Financial Market Returns 1955–2000”, *Journal of Business*, Vol. 74 (2001), pp. 1–31.

E. Dimson, P. Marsh, and M. Staunton, *The Millennium Book: A Century of Investment Returns*. ABN AMRO/London Business School (2000).

E. Dimson, P. Marsh, and M. Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns*. New Jersey: Princeton University Press (2002).

E. Dimson, P. Marsh, and M. Staunton, “Irrational Optimism”, *Financial Analysts Journal*, Vol. 60, No. 1 (2004), pp. 15–25.

E. Dimson, P. Marsh, and M. Staunton, *Global Investment Returns Yearbook 2006*. ABN AMRO/London Business School (2006a).

E. Dimson, P. Marsh, and M. Staunton, *DMS Global Returns* data module. Chicago, IL: Ibbotson Associates (2006b).

E. Dimson, S. Nagel, and G Quigley, "Capturing the Value Premium in the United Kingdom, *Financial Analysts Journal*, Vol. 59, No. 6 (2003), pp. 35–45.

P. Eichholtz, K. Koedijk and R. Otten, "De Eeuw van Het Aandeel", *Economisch Statistische Berichten*, Vol. 85, No. 4238 (2000), pp. 24–27.

E. Fama and K. French, "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay", *Journal of Financial Economics*, Vol. 60 (2001), pp. 3–43.

E. Fama and K. French, "The Equity Premium", *Journal of Finance*, Vol. 57 (2002), pp. 637–659.

P. Fernandez, "Market Risk Premium: Required, Historical and Expected", Working paper, IESE Business School, Spain, October 2004.

C. Firer and H. McLeod, "Equities, Bonds, Cash and Inflation: Historical Performance in South Africa 1925–1998", *The Investment Analysts Journal*, Vol. 50 (1999), pp. 7–28.

P. Frennberg and B. Hansson, "Computation of a Monthly Index for Swedish Stock Returns 1919–1989", *Scandinavian Economic History Review*, Vol. 40 (1992a), pp. 3–7.

P. Frennberg and B. Hansson, "Swedish Stocks, Bonds, Bills and Inflation (1919–1990)", *Applied Financial Economics*, Vol. 2 (1992b), pp. 79–86.

P. Frennberg and B. Hansson, "Computation of a Monthly Index for Swedish Stock Returns 1919–1989: Update to end-1999", Unpublished note (2000).

S. Fujino and R. Akiyama, "Security Prices and Interest Rates: 1874–1975", The Documentation Center for Japanese Economic Statistics, Institute of Economic Research. Tokyo: Hitotsubashi University (1977).

G. Gallais-Hamonno and P. Arbulu, "La Rentabilité Reelle des Actifs Boursiers de 1950 à 1992", *Economie et Statistique*, Vol. 280 (1995), pp. 3–30.

G. Gielen, *Konnen Aktienkurse Noch Steigen? Langfristige Trendanalyse des Deutschen Aktienmarktes*. Wiesbaden: Gabler (1994).

W.N. Goetzmann and R.G. Ibbotson, "History and the Equity Risk Premium", in R. Mehra (Ed.), *Handbook of Investments: Equity Risk Premium* in the *Handbook of the Economics of Finance* series. Amsterdam: Elsevier (2006).

W.N. Goetzmann, R.G. Ibbotson and L. Peng, "A New Historical Database for the NYSE 1815 to 1925: Performance and Predictability", *Journal of Financial Markets*, Vol. 4, No. 1 (2001), pp. 1–32.

W.N. Goetzmann, A. Ukhov, and N. Zhu, "China and the World Financial Markets 1870–1930: Modern Lessons from Historical Globalization", *The Economic History Review*, forthcoming 2006.

A.S. Gonzalez and J.L. Suarez, "Historical Returns in the Spanish Equity Market", *Journal of Investing*, Vol. 3, No. 1 (1994), pp. 66–71.

J.R. Graham and C.R. Harvey, "The Equity Risk Premium in September 2005: Evidence from the Global CFO Outlook Survey", Working Paper, Duke University, September 2005.

G. Grullon and R. Michaely, "Dividends, Share Repurchases, and the Substitution Hypothesis", *Journal of Finance*, Vol. 57 (2002), pp. 1649–1684.

A. Goyal and I. Welch, "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, forthcoming 2006.

R.S. Grossman, "New Indices of British Equity Prices, 1870–1913", *Journal of Economic History*, Vol. 62 (2002), pp. 121–46.

Y. Hamao, "A Standard Database for the Analysis of Japanese Security Markets", *Journal of Business*, Vol. 64 (1991), pp. 87–101.

Y. Hamao and R.G. Ibbotson, *Stocks, Bonds and Inflation 1989 Yearbook*. Chicago, IL: Ibbotson Associates (1989 and annual supplements, 1990–).

W.R. Houston, *The Annual Financial Review (Canadian)*. Toronto: Houston's Standard Publications (1900–1914).

G. Huber, "Evidence sur la Performance Relative des Marchés Obligataire et des Actions en Suisse 1960–83", Working paper, Geneva: Pictet & Cie (1985).

S. Hwang and B.K. Song, "Irrational Exuberance in the Long History of the U.K. Stock Market", Working paper, Cass Business School, London (2004).

Ibbotson Associates, *Stocks, Bonds, Bills and Inflation Yearbook*. Chicago, IL: Ibbotson Associates (2000).

R. Ibbotson and P. Chen, "Long-Run Stock Returns: Participating in the Real Economy", *Financial Analysts Journal*, Vol. 59 (2003), pp. 88–98.

R.G. Ibbotson and R. Sinquefield, "Stocks, Bonds, Bills and Inflation: Year-by-year Historical Returns (1926–74)", *Journal of Business*, Vol. 49 (1976), pp. 11–43.

R. Jagannathan, E. McGrattan and A Scerbina, "The Declining U.S. Equity Premium", *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 24 (2000), pp. 3–19.

P. Jorion and W. Goetzmann, "Global Stock Markets in the Twentieth Century", *Journal of Finance*, Vol. 54 (1999), pp. 953–980.

N.R. Kocherlakota, "The Equity Premium: It's Still a Puzzle", *Journal of Economic Literature*, Vol. 34 (1996), pp. 42-71.

P. Laforest and P. Sallee, Le Pouvoir d'Achat des Actions, des Obligations et de l'Or de 1914 à 1976", *Economie et Statistique*, Vol. 86. (1977), pp. 61–67.

M. Lettau and S. van Nieuwerburgh, "Reconciling the Return Predictability Evidence: In-Sample Forecasts, Out-of-Sample Forecasts, and Parameter Instability", *Review of Financial Studies*, forthcoming 2006.

H. Li and Y. Xu, "Survival Bias and the Equity Premium Puzzle", *Journal of Finance*, Vol. 57 (2002) pp. 1981–1995.

D. McL. Lamberton, *Share Price Indices in Australia*. Sydney: Law Book Company (1958).

L. Pástor and R.F. Stambaugh, "The Equity Premium and Structural Breaks", *Journal of Finance*, Vol. 56 (2001), pp. 1207–39.

A. Maddison, *Monitoring the World Economy 1820–1992*. Paris: Organization for Economic Cooperation and Development (1995).

M.J. Mauboussin, *Clear Thinking About Share Repurchases*, Legg Mason Capital Management (2006).

P.J. Meghen, *Statistics in Ireland*. Dublin: Institute of Public Administration (1970).

R. Mehra, "The Equity Premium: Why is it a Puzzle?", *Financial Analysts Journal*, Vol. 59 (2003), pp. 54–69.

R. Mehra and E. Prescott, "The Equity Premium: A Puzzle", *Journal of Monetary Economics*, Vol. 15 (1985), pp. 145–161.

R. Mehra and E. Prescott, "The Equity Risk Premium: A Solution?", *Journal of Monetary Economics,* Vol. 22 (1988) pp. 133–136.

R. Mehra and E. Prescott, "The Equity Premium in Retrospect", in G.M. Constantinides, M. Harris, and R.M. Stulz (Eds.), *Handbook of the Economics of Finance: Volume 1B Financial Markets and Asset Pricing*. Amsterdam: Elsevier (2003).

R. Mehra and E. Prescott, "The Equity Premium: What have we learned in 20 years?", in R. Mehra (Ed.), *Handbook of Investments: Equity Risk Premium* in the *Handbook of the Economics of Finance* series. Amsterdam: Elsevier (2006).

B.R. Mitchell, *International Historical Statistics: Europe 1750–1993*. London: Macmillan Press (1998).

J. Newlands, *Put Not Your Trust in Money*. Association of Investment Trust Companies (1997).

S. Nielsen and O. Risager, "Macroeconomic Perspectives on Stock and Bond Investments in Denmark since the First World War", in T.M. Andersen, O. Risager, and S.H.E. Jensen (Eds), *Macroeconomic Perspectives on the Danish Economy* . London: Macmillan Press (1999).

S. Nielsen, and O. Risager, "Stock Returns and Bond Yields in Denmark 1922–99", Working paper, Department of Economics, Copenhagen Business School (2000).

F. Panetta and R. Violi, "Is there an Equity Premium Puzzle in Italy? A Look at Asset Returns, Consumption and Financial Structure Data over the Last Century", Temi di Discussione 353, Bank of Italy (1999).

H.H Panjer and K.S. Tan, *Report on Canadian Economic Statistics 1924–2001*. Ottawa: Canadian Institute of Actuaries (2002).

C. Parum, "Historisk Afkast af Aktier og Obligationer i Danmark", *Finans Invest*, Vol. 3 (1999a), pp. 4–13.

C. Parum, “Estimation af Realkreditobligationsafkast i Danmark i Perioden 1925–1998”, *Finans Invest*, Vol. 7 (1999b), pp. 12–15.

C. Parum, “An Equally Weighted Index of Copenhagen Stock Prices 1900–1914”, Unpublished note (2002).

R. Rajan and L. Zingales, “The Great Reversals: The Politics of Financial Development in the Twentieth Century”, *Journal of Financial Economics*, Vol. 69 (2003), pp. 5–50.

E. Rätzer, *Die Pensionskasse aus Ökonomischer Sicht*. Berne: Paul Haupt (1983).

R. Rau and T. Vermaelen, “Regulation, Taxes and Share Repurchases in the U.K.” *Journal of Business* Vol. 75 (2002), pp. 245–282.

J. Ritter, “Economic Growth and Equity Returns”, *Pacific-Basin Finance Journal*, Vol. 13, No. 5 (2005) pp. 489–503.

U. Ronge, *Die Langfristige Rendite Deutscher Standardaktien*. Frankfurt: Peter Lang (2002).

M. Sandez and F.G. Benavides, “La Bolsa de Madrid entre 1919 y 1936”, *Revista de la Bolsa de Madrid*, Vol. 87 (2000), pp. 32–39.

C.G.W. Schumann and A.E. Scheurkogel, “Industrial and Commercial Share Price Indices in South Africa”, Bureau for Economic Research, Faculty of Commerce, University of Stellenbosch (1948).

S.S. Schwartz and W.T. Ziemba, “The Japanese Stock Market 1949–1991”, in W.T. Ziemba, W. Bailey and Y. Hamao, (Eds) *Japanese Financial Market Research*. Amsterdam: Elsevier Science (1991).

G.W. Schwert, “Indexes of U.S. Stock Prices from 1802 to 1987”, *Journal of Business*, Vol. 63 (1990), pp. 399–426.

J.J. Siegel, *Stocks for the Long Run*, third edition. NY: McGraw Hill (2002).

J.J. Siegel, *The Future for Investors : Why the Tried and the True Triumph Over the Bold and the New*. NY: Crown Business (2005).

J.J. Siegel and J.D. Schwartz, "The Long-Term Returns on the Original S&P 500 Firms", *Financial Analysts Journal*, Vol. 62, No. 1 (2006), pp. 18-31.

R. Stehle, “Der Size-effekt am Deutschen Aktienmarkt”, *Zeitschrift für Bankrecht und Bankwirtschaft*, Vol. 9, No. 3 (1997), pp. 237–260.

R. Stehle, C. Wulff, and Y. Richter, “Die Rendite deutscher Blue-chip-Aktien in der Nachkriegszeit - Rückberechnung des DAX für die Jahre 1948 bis 1954”, Working Paper, Humboldt University, Berlin (1999).

L. Switzer, “An Equally Weighted Index of Montreal Stock Prices 1900–1914”, Unpublished note (2000).

A. Taylor, "A Century of Purchasing Power Parity", *Review of Economics and Statistics*, Vol. 84, No. 1 (2002), pp. 139–150.

W.A. Thomas, *The Stock Exchanges of Ireland*. NH: Francis Cairns Publications (1986).

A. Timmermann, “Udviklingen i de Danske Aktiekurser 1914–1990”, *Nationalokonomisk Tidsskrift*, Vol. 130 (1992), pp. 473–482.

M.C. Urquhart and K.A.H. Buckley, *Historical Statistics of Canada*. Toronto: Macmillan Company (1965).

S.F. Valbuena, “Sources and Methodology of Spanish Bolsa Returns in the 20th Century”, Working paper, Madrid: Universidad Complutense (2000).

J.A. Vandellos, “El Porvenir del Cambio de la Peseta”, Instituto de Investigaciones Economicas, Banca Mas Sarda (1936: reprinted 1974).

I. Welch, “Views of Financial Economists on the Equity Premium and Other Issues”, *Journal of Business*, Vol. 73 (2000), pp. 501–537.

I. Welch, “The Equity Premium Consensus Forecast Revisited”, Cowles Foundation Discussion Paper Number 1325, Yale School of Management, September 2001.

S. Whelan, *From Canals to Computers: The Friends First Guide to Long-Term Investment Returns in Ireland.* Dublin: Friends First Asset Management (1999).

J. Wilson and C. Jones, “An Analysis of the S&P 500 Index and Cowles’s Extensions: Price Indexes and Stock Returns, 1870–1999”, *Journal of Business*, Vol. 75 (2002), pp. 505–533.

D. Wydler, “Swiss Stocks, Bonds and Inflation 1926–1987”, *Journal of Portfolio Management*, Vol. 15, No. 2 (1989), pp. 27–32.

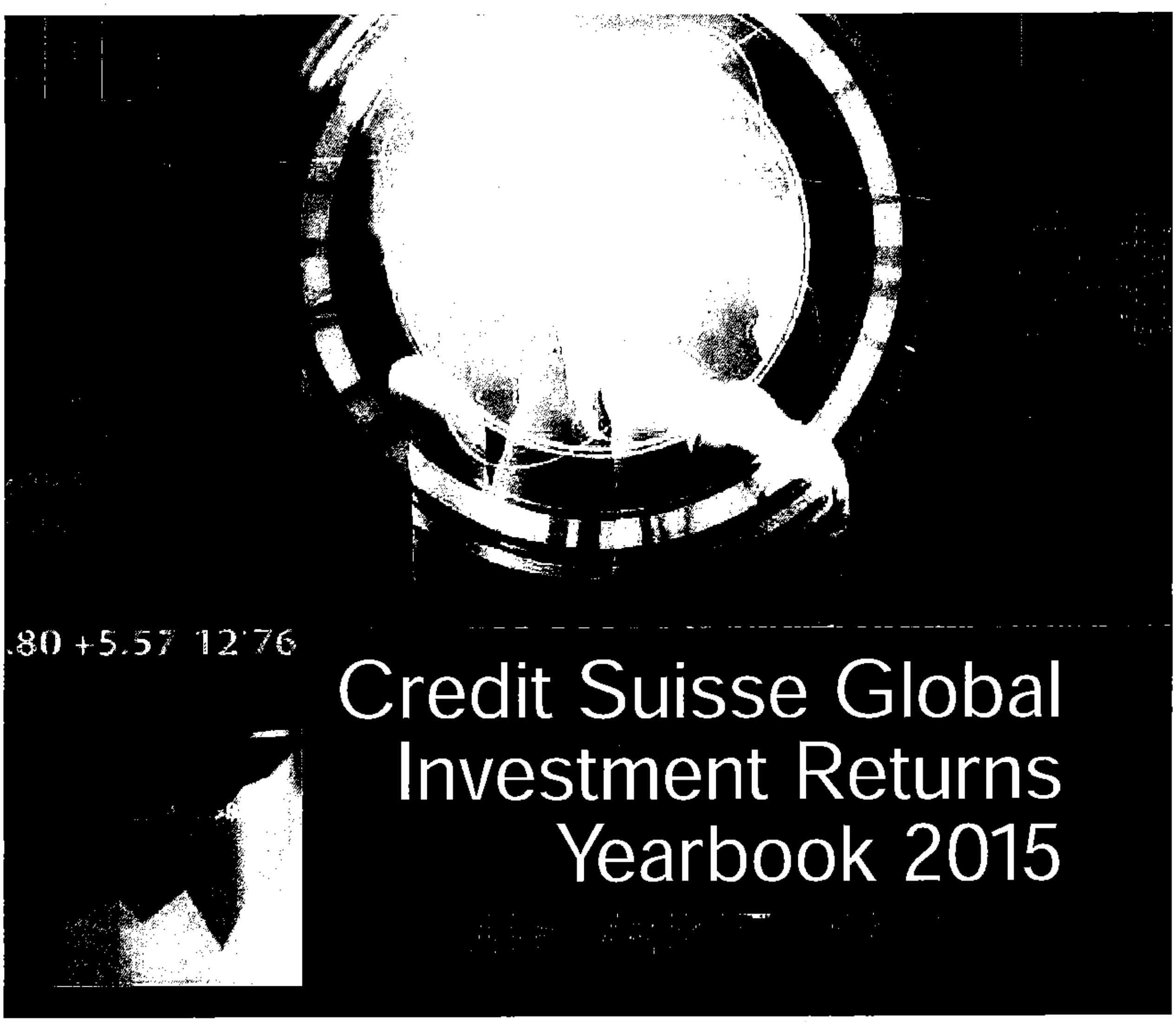
D. Wydler, *The Performance of Shares and Bonds in Switzerland: An Empirical Study Covering the Years Since 1925*. Geneva: Pictet & Cie (2001).

CREDIT SUISSE

February 2015

Research Institute

Thought leadership from Credit Suisse Research and the world's foremost experts



Credit Suisse Global Investment Returns Yearbook 2015

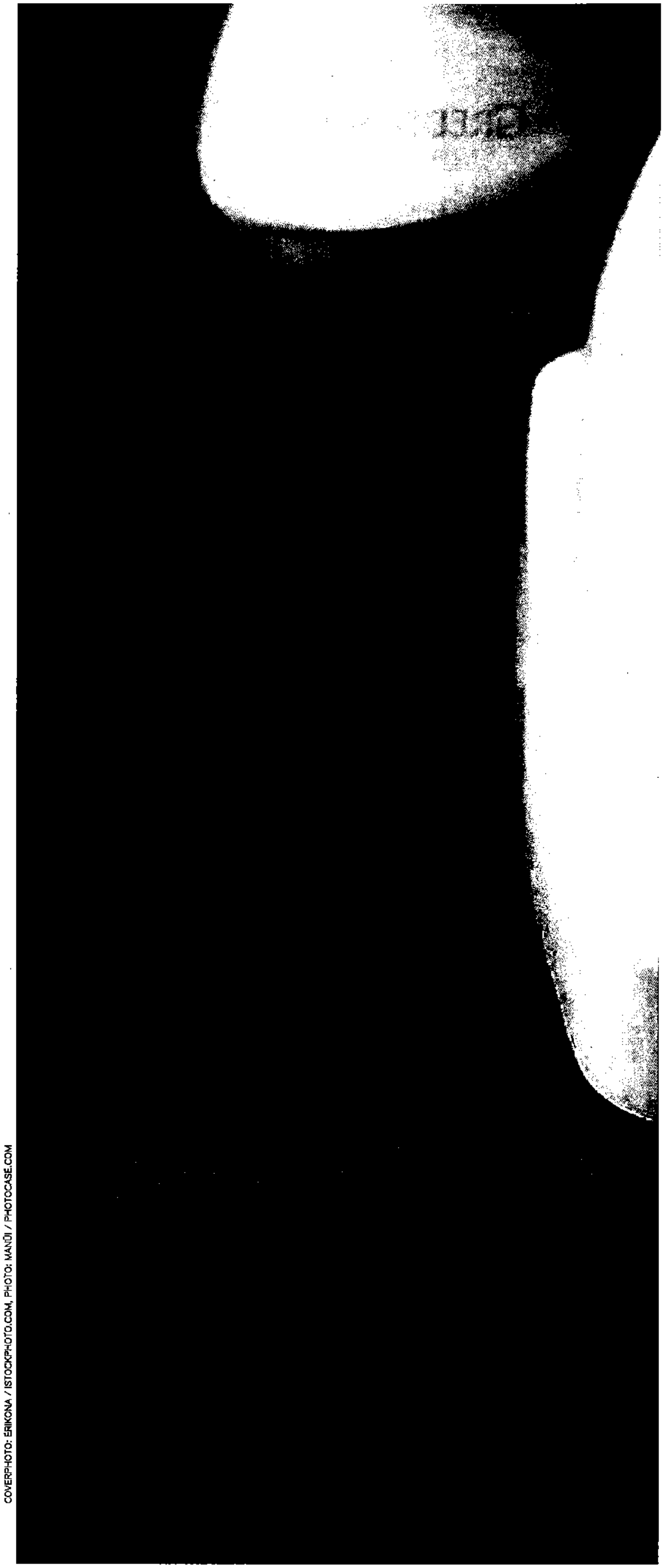
Contents

For more information on the findings of the Credit Suisse Global Investment Returns Yearbook 2015, please contact either the authors or:

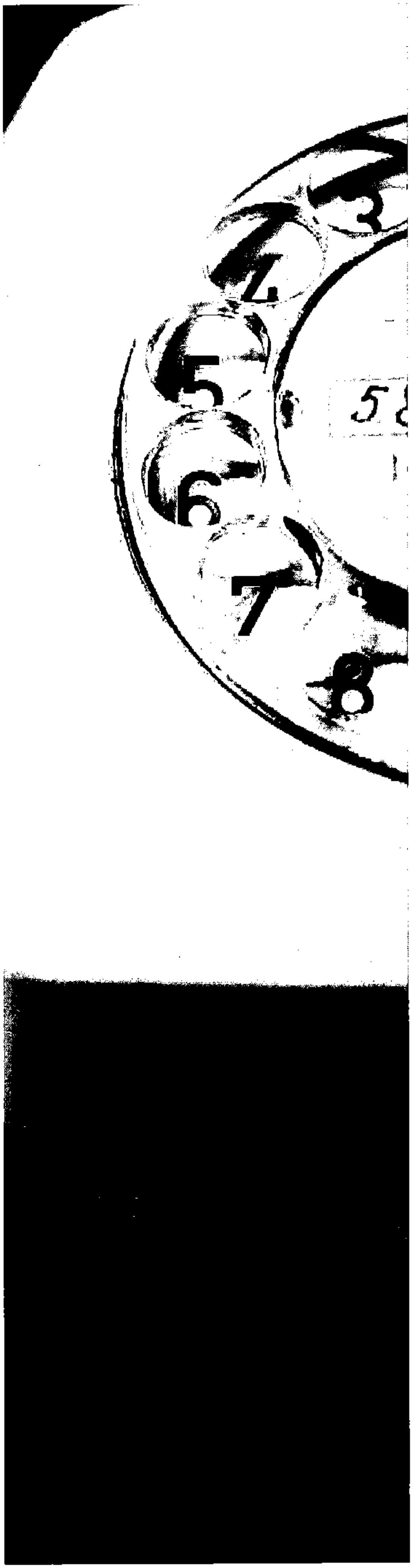
Michael O'Sullivan, Chief Investment Officer UK & EMEA, Credit Suisse Private Banking & Wealth Management, michael.o'sullivan@credit-suisse.com

Richard Kersley, Head of Global Securities Products and Themes, Credit Suisse Investment Banking, richard.kersley@credit-suisse.com

To contact the authors or to order printed copies of the Yearbook or of the accompanying Sourcebook, see page 66.



COVERPHOTO: ERIKONA / ISTOCKPHOTO.COM, PHOTO: MANŪI / PHOTOCASE.COM



Introduction

2015 has begun with a series of apparent contradictions and dramatic reversals. In the developed world, both equity and bond markets are at record highs. The price of oil has collapsed and the Swiss franc has jettisoned its link with the euro. Global economic growth is tepid and disinflation has caused many central banks to further cut interest rates or, in the recent case of the European Central Bank, to take extraordinary action in the shape of its quantitative easing program. Against this volatile backdrop, we launch the 2015 Credit Suisse Global Investment Returns Yearbook and hope that the wealth of stock, bond and inflation data in the Yearbook will help to frame market developments in the light of long-term asset price trends.

The 2015 Yearbook contains data spanning 115 years of history across 26 markets and the companion publication, the Credit Suisse Global Investment Returns Sourcebook 2015 extends the scale of this resource further with detailed tables, graphs, listings, sources and references for every country. In the first two chapters of the Yearbook, Elroy Dimson, Paul Marsh and Mike Staunton from the London Business School analyze this rich dataset in order to examine an established and new way of investing.

In the first chapter, they focus on the importance of industry weightings for long-term investors. Today, in the US and UK markets, only the banks and mining industries have weightings close to their 1900 levels. Indeed, in 1900, the railway industry made up 50% of the UK market and nearly two thirds of the US market. They examine the returns from new and old industries, as well as the implications for investors of structuring portfolios along industry lines by considering questions such as whether industry diversification is more important than country diversification and whether to overweight the old economy or the new? Interestingly, they find that returns can be higher from investing in old rather than new industries.

The second Yearbook chapter examines responsible investing – a topic we developed in a 2012 Credit Suisse Research Institute report "Investing for Impact." We believe that this is an important and growing area in the investment management field and this chapter measures several approaches to investing along social, environmental and ethical lines. It also provides evidence that corporate engagement can pay, whether the focus is on environmental and social issues or on corporate governance.

Finally, in Chapter 3, David Holland and Bryant Matthews of the CS HOLT team complement the historic data in the Yearbook with a market-implied approach. They study how the market-implied cost of capital mean reverts over time and the extent to which this is in any way predictable. They note that, at the country level, China and Switzerland currently have the lowest market-implied discount rates, while Russia, Italy and Argentina have the highest.

We are proud to be associated with the work of Elroy Dimson, Paul Marsh, and Mike Staunton, whose book Triumph of the Optimists (Princeton University Press, 2002) has had a major influence on investment analysis. The Yearbook is one of a series of publications from the Credit Suisse Research Institute, which links the internal resources of our extensive research teams with world class external research.

Giles Keating
Head of Research and Deputy Global CIO, Credit Suisse Private Banking and Wealth Management

Stefano Natella
Head of Global Equity Research, Credit Suisse Investment Banking


Q W E R T
A S D F G

Industries: Their rise and fall

This article focuses on the importance of industry weightings for long-term investors. We show how industries have risen and fallen as technology has advanced. Successive waves of new industries and companies have transformed the world, yet they have sometimes proved disappointing investments. We seek to explain how the decline of old industries, together with some investment disappointments from new ones, have somehow generated good overall returns. Finally, we examine some implications for investors. Is industry rotation worthwhile? Should investors pay attention to building portfolios that are well diversified across industries? Is industry diversification now more important than country diversification?

Elroy Dimson, Paul Marsh and Mike Staunton, London Business School

Understanding the factors that drive stock returns has long been the quest of professional investors. Greater knowledge has led to an increase in investing based on factor exposures, sometimes known as smart beta. This has moved far beyond the traditional emphasis on industry and country factors or even on factors such as size, value and momentum. Hsu (2014) reports that one quantitative investor is now using an 81-factor model.

Despite factor proliferation, industries remain one of the original and most important factors. They are a key organizing concept. Investment organizations continually review industrial classifications and, where necessary, recommend revisions. Companies often seek advantage by "window dressing" their industry affiliation. Investment research is mostly structured along industry lines.

When fund managers build, alter, or report on portfolios, they refer to industry weightings. Each year, there is a wide dispersion of returns across industries, so that getting these weightings right – or wrong – has consequences. Industry membership is the most common method for grouping stocks for portfolio risk management, relative valuation and peer-group valuation. And investors wrestle with whether to focus primarily on industries or countries in asset allocation, when taking active positions, and when seeking to diversify.

In research terms, however, industries are the Cinderella of factor investing. The two most comprehensive and influential books on factor investing, Antti Ilmanen's (2011) Expected Returns and Andrew Ang's (2014) Asset Management, have almost nothing to say about industries. This article contributes toward redressing this imbalance.

The great transformation

In 1900 – the start date of our global returns database – virtually no one had driven a car, made a phone call, used electric lighting, seen a movie or heard recorded music; no one had flown in an aircraft, listened to the radio, watched TV, used a computer, sent an email or used a smartphone. There were no x-rays, body scans, DNA tests or transplants, and no one had taken an antibiotic. Many would die young because of this.

Mankind has enjoyed a wave of transformative innovation dating from the Industrial Revolution, continuing through the golden age of invention of the late 19th century, and extending into today's information revolution. This has given rise to entire new industries – electricity and power generation, automobiles, aerospace, airlines, telecommunications, oil and gas, pharmaceuticals and biotechnology, computers, information technology, media and entertainment. Meanwhile, makers of horse-drawn carriages and wagons, canal boats, steam locomotives, candles, and matches have seen their industries decline. There have been profound changes in what is produced, how it is made, and in the way in which people live and work.

These changes can be seen in the shifting composition of the types of firms listed on world stock markets. Figure 1 shows the industrial composition of listed companies in the USA and UK. The top two pie charts show the position at start-1900, while the bottom two show start-2015.

Markets at the start of the 20th century were dominated by railroads. In the UK, railway companies accounted for almost half the value of the stock market, while in the USA they had a 63% weighting. Yet 115 years later, railroads have declined almost to the point of stock market extinction, representing less than 1% of the US market, and almost zero in the UK.

Of the US firms listed in 1900, more than 80% of their value was in industries that are today small or extinct; the UK figure is 65%. Besides railroads, other industries to have declined precipitously are textiles and iron, coal and steel. These industries still exist, but have moved to lower cost locations in the emerging world. Yet similarities between 1900 and 2015 are also apparent. The banking and insurance industries have continued to be important. Similarly, industries such as food, beverages (including alcohol), tobacco and utilities were present in 1900, just as they are today. And in the UK, quoted mining companies were important in 1900, just as they are in London today.

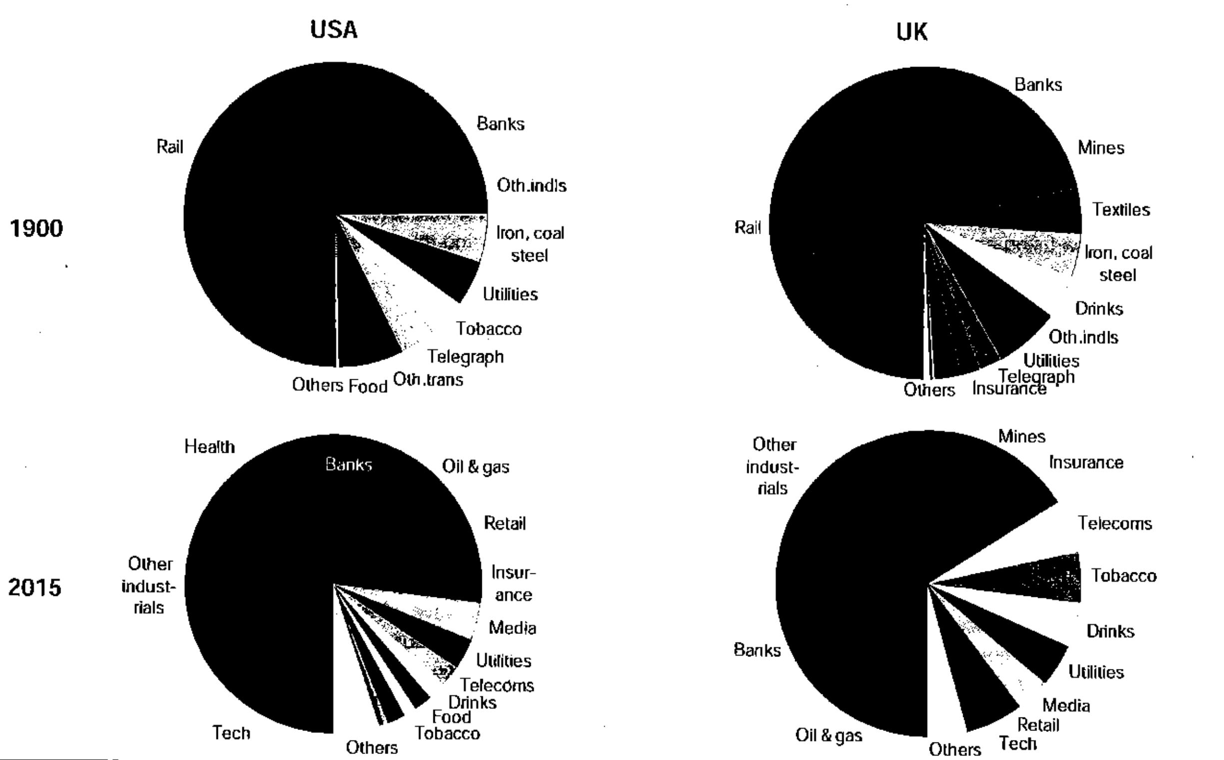
But even industries that initially seem similar have often altered radically. For example, compare telegraphy in 1900 with smartphones in 2014. Both were high tech at the time. Or contrast other transport in 1900 – shipping lines, trams, and docks – with their modern counterparts, airlines, buses and trucking. Similarly, within manufacturing and industrials, the 1900 list of companies includes the world's then largest candle maker and the world's largest manufacturer of matches.

Another statistic that stands out from Figure 1 is the high proportion of today's companies whose business is in industries that were small or non-existent in 1900 – 62% by value for the USA and 47% for the UK. The largest industries in 2015 are technology (notably in the USA), oil and gas, banking, healthcare, the catch-all group of other industrials, mining (for the UK), insurance, telecommunications and retail. Of these, oil and gas, technology, and health care (including pharmaceuticals and biotechnology) were almost totally absent in 1900. Telecoms and media, at least as we know them now, are also really new industries.

Figure 1

Industry weightings in the USA and UK, 1900 compared with 2015

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Triumph of the Optimists (for 1900: UK based on Top 100 companies, US on total market) and FTSE All World Indices (for 2015)



Our analysis relates purely to the quoted segment. Some industries existed throughout the period, but were not always listed. For example, there were many retailers in 1900, but apart from the major department stores, these were often small local outlets, rather than national retail chains like Walmart or Tesco. Similarly, in 1900, a higher proportion of manufacturing firms were then family owned and not stock-market listed.

In the UK and other countries, nationalization has caused entire industries – railroads, utilities, telecoms, steel, airlines, airports – to be de-listed and often later re-privatized. Our analysis includes the value of, for example, listed railroads, while omitting highways that remain largely in national or state ownership. Despite these caveats, the comparisons above mostly reflect the industrial evolution that has taken place over the last century, rather than just changes in ownership.

Long-run industry performance

It is instructive to look at long-run industry performance. For the USA, we use Ken French's industry data (Fama and French, 1997) for 1926–2014. There are 49 industries, 40 of which start in 1926. From 1900 to 1925, we use the 57 Cowles (1938) industries, 20 of which start in 1900. Our focus is on industries not sectors. These terms are often used interchangeably. However, we use "sector" to refer to a large segment of the economy, while an "industry" is a more detailed grouping of businesses. For example, the banking industry is part of the financial sector. Except where we state otherwise, this article is based on the more detailed industry groups.

Figure 2 shows the performance of the 15 US industries for which we have data back to 1900. The red line shows that a dollar invested in the US market at start-1900 would have grown, with dividends reinvested, to USD 38,255 by end-2014, representing an annualized return of 9.6%. The industries display a wide dispersion around this. A dollar invested in the worst performer, shipbuilding and shipping, would have grown to just USD 1,225, representing an annualized return of 6.4%. The best performer, tobacco, gave an annualized return of 14.6%, and a terminal value of USD 6.2 million, over 5,000 times as much as from shipbuilding and shipping.

This dispersion of long-run returns across industries is similar to the dispersion across countries (see pages 37–60 below and the companion Sourcebook). Just as some countries were "lucky" and others less fortunate, some industries prospered while others foundered. And just as we can infer little about future country returns from past returns, we can infer little about long-run future industry returns from their historical record. In fact, as Ilmanen (2011) concludes, countries and industries seem to be good examples of non-priced investment factors. If a factor is priced, investors can expect it to generate a long-run premium. For industries, however, the starting point is that they are likely to have similar expected returns, except to the extent that they are exposed to other factors. For example, an industry might, at a point in time, have a higher expected return because it has a higher beta, or is value-oriented with a low ratio of market price to fundamental value, or – as we explain in the accompanying article – is shunned by many investors.

But while industries may not be a "priced" factor, they remain important. The dispersion of industry returns is large, whether we look at the 115-year period, or at year-by-year returns. The average yearly cross-sectional dispersion across all US industries (not just those in Figure 2) averaged 22% over 1900–2014. Meanwhile, the average annual spread between the best- and worst-performing industries was 108%. Industries perform very differently from one another, even if it is hard to predict these differences in advance. Industries and industry weightings matter.

Figure 2

Long-run performance of industries in the USA

Source: Elroy Dimson, Paul Marsh and Mike Staunton; Cowles (1938), Ken French industry data; DMS USA index

Cumulative value of USD 1 invested in US industries at the start of 1900

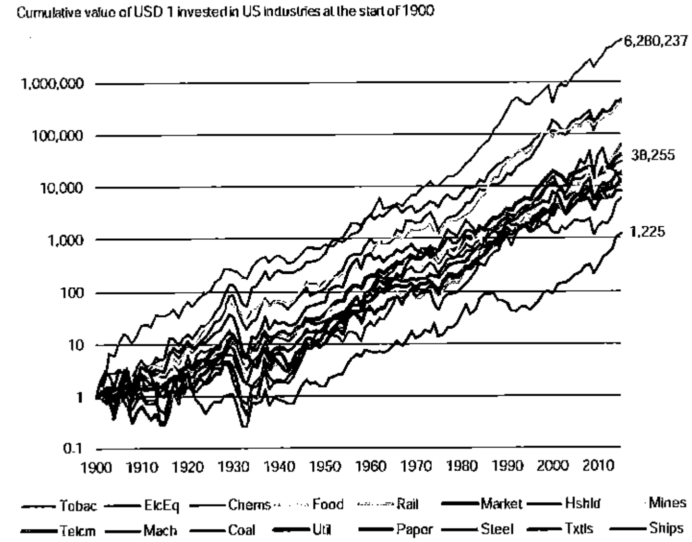


Figure 2 suffers from hindsight bias. By focusing on a full 115-year history, our sample contains only those industries that existed in 1900 and which survived. With hindsight, we know that many of the survivors declined in importance. Of the 15 industries in Figure 2, ten underperformed the market. Not surprisingly, these included coal, steel, textiles and shipbuilding. Since 1900, these industries declined in developed countries, but grew in importance in developing countries.

Figure 3

Long-run performance of industries in the UK

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Top 100 database; FTSE International; DMS UK index

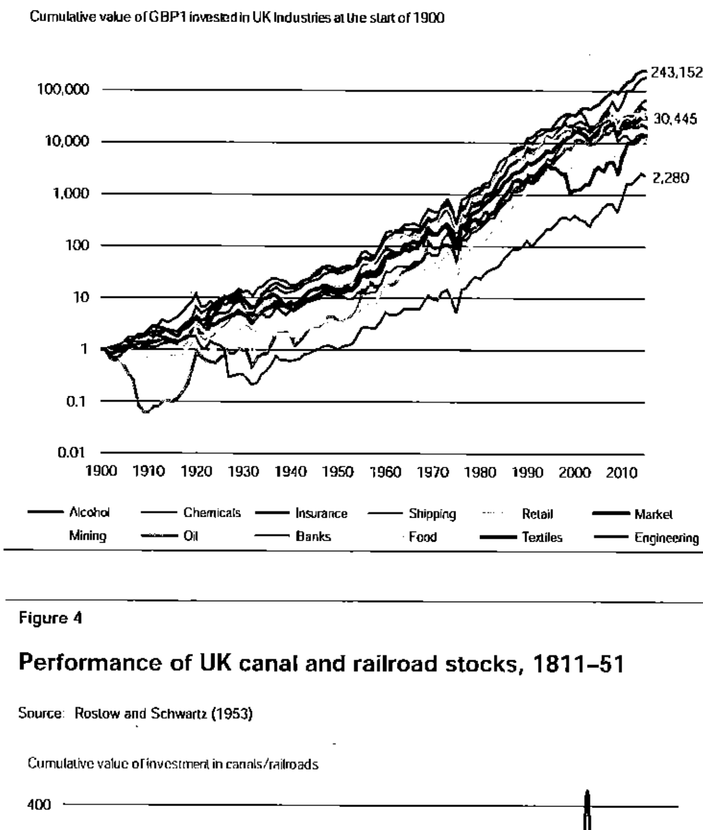
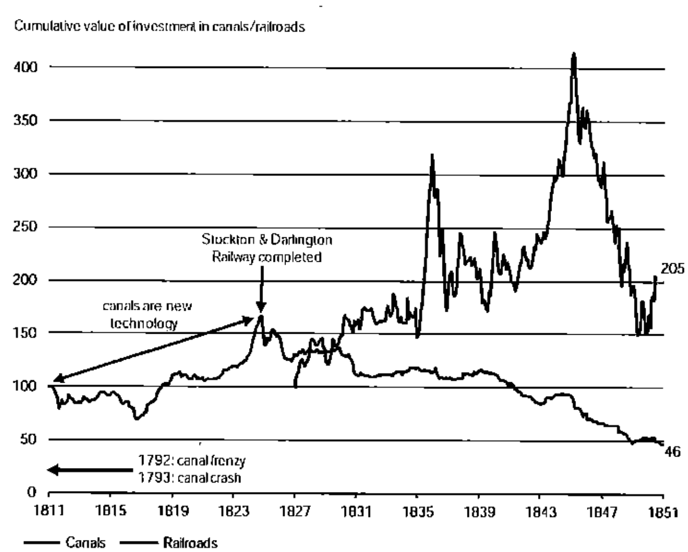


Figure 4

Performance of UK canal and railroad stocks, 1811–51

Source: Rostow and Schwartz (1953)



Since it excludes industries that emerged after 1900, Figure 2 provides only a partial perspective. But if we were to focus instead on the performance of industries that are substantial today, we would introduce success bias. Whether we start with industries as they existed in 1900, or with those that are important today, it is hard to avoid the intrusion of hindsight. It is implicit in most analysis of long-run industry performance.

To generate equivalent long-run industry histories for the UK, we use the FTSE International industry indices starting in 1962. There were originally 40 industries, and while there are still 40 today, they have changed over time. Pre-1962, we constructed our own industry indices based on the top 100 UK companies from 1900 to 1955 and the London Share Price Database thereafter.

Figure 3 displays the eleven UK industries for which we have a full 115-year history. Some, such as chemicals and textiles, are the same industries that we saw in Figure 2 for the USA. But several of the long-run US industry histories have no UK equivalent. This is because the UK's post-war nationalization programme took railroads, utilities, telecoms, steel, coal and shipbuilding into public ownership. Although they were later re-privatized, these industries lack continuous histories. However, Figure 3 does include three industries – banks, insurance and alcoholic beverages – for which there is no long-run US index. This is because financials were omitted from the Cowles data, and for alcohol, because of prohibition.

The red line in Figure 3 for the overall UK market shows that GBP 1 invested in 1900 would have grown to GBP 30,445 by end-2014, an annualized return of 9.4%. The remaining line plots in Figure 3 again show a wide dispersion of industry performance. The best industry was alcohol, while the worst was engineering. UK insurance companies performed strongly, while banks underperformed, due to the recent financial crisis. It is interesting to note that the best performer in the UK, alcohol, and the best US performer, tobacco, are both from "sin industries", an issue to which we return in the next article.

Rise and fall through disruptive technology

The Industrial Revolution began in the UK in the late 18th century fueled by inventions such as the spinning jenny and power loom, improvements in metallurgy and the harnessing of steam. But transportation for these new manufactured goods was inadequate. The solution was canals. To turnpike operators and the owners of fleets of wagons and horses, canals were a disruptive technology. Goods could be transported sixty times more efficiently in ton miles per day.

Nairn (2002) points out that between the late 18th century and 1824, more than 60 canal companies floated on the London Stock Exchange, raising the equivalent of USD 32 billion in today's money. 1792 saw canal frenzy, followed by a crash the next year. There are no indices to show the magnitude of these events, but Figure 4 shows an index of canal stock prices (the red line plot) from a later period, compiled by Rostow and Schwartz (1953). From 1816 to 1824 canal stock prices rose by 140%.

But in 1825, the Stockton and Darlington Railway was completed. Over the next quarter century, the stock prices of canals fell by over 70%. Although Rostow and Schwartz's index excludes dividends, it seems likely that, over the 40 years spanned by Figure 4, investors' total returns were low. Canals, the disruptive technology, had in turn been disrupted by railroads. Once rail freight became established, it was some 60 times more efficient than canals in ton miles per day.

Railroad mania took hold in Britain, peaking in 1846, when 272 new lines were approved. The blue line plot in Figure 4 shows the accompanying stock market frenzy. Railroad stock prices more than doubled in 1835, only to fall back again almost to their prior level. They more than doubled again by 1845, then fell two thirds by 1849. Many writers have described this as a bubble, but crucial infrastructure was built and, over the quarter century spanned by Figure 4, investors earned a reasonable return – an annualized capital gain of 3% plus dividends. However, it was a rocky ride.

In his book, Engines that Move Markets, Alasdair Nairn (2002) analyzes investment in successive new technologies, starting with canals and railroads, continuing through telegraph, electric light, crude oil, automobiles, wireless, radio and TV, computers, PCs, and ending with the internet. He finds that most new technologies were initially greeted with skepticism and derision, and faced a struggle for acceptance. He provides numerous amusing quotations, such as "What could be more palpably absurd than the prospect of locomotives travelling twice as fast as stagecoaches?" (Quarterly Review, 1825).

Once conquered, skepticism tends to be followed by over-enthusiasm, with new technologies often leading to stock market "bubbles," which Nairn defines as periods when investors seem to suspend rational valuation, which is typically followed by a calmer, more rational assessment. The firms that made money from the new technology over an extended period tended to have monopoly protection, effective barriers and a sustainable advantage.

Nairn concludes that stock market investors were not always the biggest beneficiaries of new technology. The latter tended to be the "insiders," i.e. the innovators, founders and providers of venture funding, along with consumers and society as a whole. In the 2014 Yearbook, we offer a similar explanation of how emerging-market growth may benefit local people more than stock-market investors.

New industries or old?

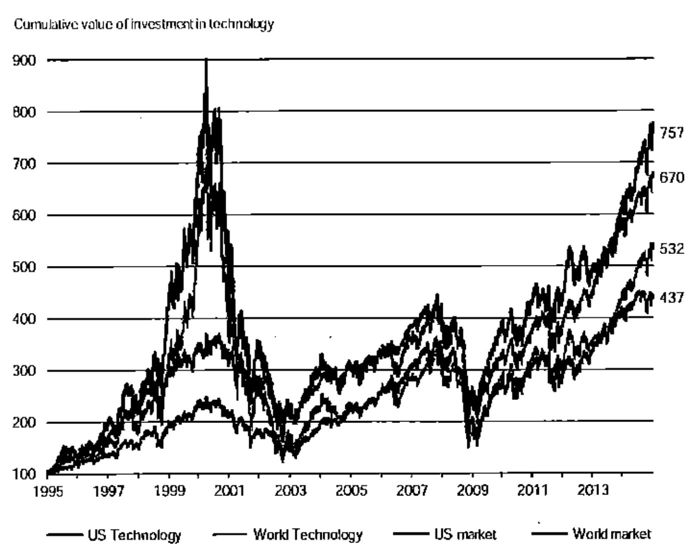
New industries can deliver disappointing returns if stock market prices are initially too optimistic about future growth. Declining industries can disappoint if investors fail to realize the speed and extent of their demise. But if this is the historical pattern, how have stock markets generated good long-run returns? Perhaps this has arisen from the "in-between" industries or "the tried and the tested" (Siegel (2005)).

The contrast between the new and the old is illustrated by the most celebrated "bubble" of recent times, the dot-com boom and crash. Figure 5 shows the total returns from the technology sector, comprising the hardware and software industries, over the 20 years since 1995, when the internet boom began. The dark blue line shows the FTSE US technology index, while the light blue line shows world technology. The gray and red lines show the overall US and world market returns. The spike centered on March 2000 is dramatic, representing a nine-fold rise over the previous five years. Tech stocks then fell by 82% over the next two-and-a-half years.

Figure 5

Performance of technology stocks: 1995 to date

Source: FTSE International All World Index series



But technology has not been a poor investment over this period. An investor in technology stocks over the last 20 years would have beaten the market, with an annualized return of 10.5% versus 9.9% for US stocks as a whole. Holders of the technology sector today would be losing money only if they had been unlucky enough to have bought between January and September 2000. Despite the bubble, the technology sector has, for most investors, generated good returns.

Old, declining industries can also provide good returns. We saw that railroads made up 63% of the US stock market in 1900, but less than 1% today, making them the ultimate declining industry. Figure 6 shows railroad returns from 1900 to date versus the US market. It also shows the returns on airlines and road transport companies (buses, trucks, and so on). The airline series starts in 1934 as there were no airlines in 1900, while the road transportation index begins in 1926, as there was no index before then. Both series start at the then-value of the rail index.

Figure 6 shows that from 1900 to date, railroads actually outperformed the market. But until the early 1970s, they trailed badly as their business model was disrupted by both air travel and trucking. The 1950s and 1960s were especially challenging. Completion of the interstate highway system meant that trucking took much of their freight traffic, while Americans took to their cars, lowering short-haul rail passenger traffic. Meanwhile, the airlines took almost all their long-haul passengers. This led to some high-profile bankruptcies, culminating in the Penn Central failure in 1970, then the largest-ever US bankruptcy.

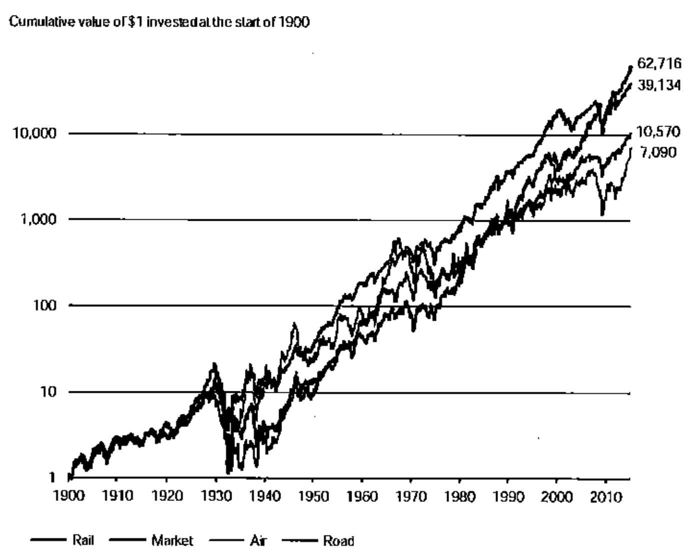
But Figure 6 shows that, since then, railroads have outperformed airlines, road transport and the US market. As Siegel (2005) points out, with hindsight, railroad stock prices had become too depressed following the bankruptcies. The renaissance was also helped by industry rationalization, deregulation and big increases in productivity. The newest technology, airlines, performed the worst. As Warren Buffet said of the Wright Brothers, "If a farsighted capitalist had been present at Kitty Hawk, he would have done his successors a huge favor by shooting Orville down."

Investors should shun neither new nor old industries. There can be times when stock prices in new industries reflect over-enthusiasm about growth, and times when investors become too pessimistic about declining industries. However, it is dangerous to assume that investors persistently make errors in the same direction: they may at times underestimate the value of new technologies and overestimate the survival prospects of moribund industries. There is a role for classic investment analysis to seek out industries and stocks that represent good value, and to avoid those that seem overpriced.

Figure 6

Performance of US transportation stocks: 1900 to date

Source: Elroy Dimson, Paul Marsh and Mike Staunton; DMS USA index; other indices compiled from CRSP



The birth of industries

From the 18th century canal boom to the late 19th century internet revolution, the birth of new industries has been heralded by successive waves of companies joining stock markets via IPOs.

The S&P 500 index began in 1957. It is regularly rebalanced to ensure it continues to represent 500 "leading firms in leading industries." New companies, often representing new industries, enter as IPOs or once they grow large enough. By 2003, 917 new constituents had joined the index (Siegel, 2005). Meanwhile, companies leave if they get acquired, shrink, fail, or otherwise die.

Siegel compared the returns from investing in the actual S&P 500 with a strategy of just holding the original constituents, and reinvesting the proceeds from deaths in the survivors. He found that investors would have been better off if they had stuck with the original firms. This would also have been less risky. Siegel explains, "Investors have a propensity to overpay for the "new" while ignoring the "old" ... growth is so avidly sought after that it lures investors into overpriced stocks in fast-changing and competitive industries, where the few big winners cannot compensate for the myriad of losers."

Siegel's findings are consistent with the large body of evidence on IPOs. Ritter (2014) analyzes 7,793 US IPOs from 1980 to 2012. Investors who bought at the issue price made an average first day return of 17.9%. However, investors then experienced an average market-adjusted loss of 18.6% over the next three years.

The UK findings are similar. Dimson and Marsh (2015) analyze 3,507 IPOs from 2000 to 2014. The market-value weighted average first day return for investors who bought at the issue price was 8.5%. Over the next two years, the average loss, adjusted for market movements, was 9.4%. Gregory, Guermat and Al-Shawawreh (2010) show that post-IPO underperformance lasts even longer. For 2,499 UK IPOs from 1975 to 2004, they find an average underperformance of 31.6% over the five years post-IPO.

Loughran and Ritter (1995) argue that IPOs are systematically overpriced. "For IPOs the prior rapid growth of many of the young companies makes it easy to justify high valuations by investors who want to believe they have identified the next Microsoft." But why don't investors learn, and enforce lower IPO prices? After all, the poor long-run performance of IPOs has been well publicized over the last 30 years. Loughran and Ritter suggest a behavioral explanation, "investors are betting on longshots ... [and] seem to be systematically misestimating the probability of finding a big winner. It is the triumph of hope over experience."

IPOs are typically growth stocks in growth industries, and their performance thus conforms to the extensive evidence that, over the long run, growth stocks have underperformed both the market and value stocks. This evidence is reviewed in the accompanying Sourcebook. There is still controversy over whether the value premium arises from behavioral factors, or is a reward for risk. The main behavioral argument is that investors are too much in love with, and overpay for growth. This is the Loughran and Ritter position.

Given the evidence on post-IPO performance, we might expect the return from stocks to depend on their "seasoning," which is defined as the time that has elapsed since their IPO (see Ibbotson, 1975). Figure 7 shows the impact of seasoning on UK stock returns. The four line plots show the returns over the last 35 years from a strategy of investing in stocks which at the start of each year had three years or less seasoning, 4–7 years, 8–20 years, and more than 20 years. The four portfolios are rebalanced annually to ensure that they always capture the desired range of seasoning.

Figure 7 shows that the greater the seasoning, the higher the returns. The only exception to this was briefly around the dot-com boom and bust. But by the end of the period, terminal wealth was almost three times higher from investing in the most, rather than the least seasoned stocks. At the stock level, old clearly beats new. And since new industries are disproportionately represented among IPOs, this lends credence to Nairn's observation that new industries and new technology often experience periods of over-enthusiasm.

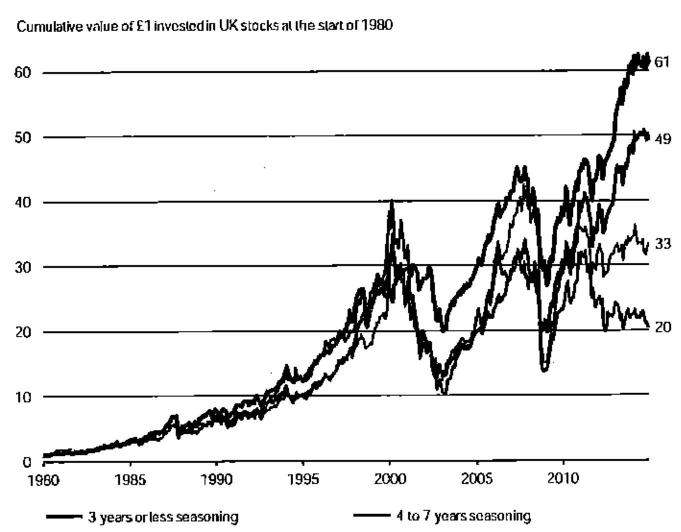
Industry rotation strategies

If industries can experience periods of over- or under-valuation, it may be possible to exploit this through industry rotation. We examine this for the USA and UK using the industry data described above. We focus on two variables, past returns and a value measure. The choice of the latter is driven by data availability, and is either industry book-to-market (USA, 1927 on), or industry yield (USA, prior to 1927 and UK throughout).

Figure 7

Impact of seasoning on UK stock returns, 1980–2014

Source: Elroy Dimson, Paul Marsh and Mike Staunton



Each new year, we rank all then-existing industries by either their past year's return or the value metric. We assign industries to quintiles from the lowest- to the highest-ranked groupings, and invest equal amounts in the industries in each quintile. Industries are re-ranked annually, bringing in new ones that have emerged, and dropping any for which indices are no longer produced. This strategy is repeated annually from 1900 to 2014. Figure 8 summarizes the results.

Industry rotation: Reversals or momentum?

The two sets of bars on the left of Figure 8 relate to rotation based on prior-year returns for the USA (dark blue) and the UK (gray). Each set of bars shows the annualized returns from investing in the previous year's worst performers (losers), through to investing in the best quintile (winners). If industries periodically become over- or undervalued, and then revert to fair value, we might expect reversals, with past losers beating past winners.

Figure 8

Industry rotation strategies in the USA and UK,1900–2014

Source: Elroy Dimson, Paul Marsh and Mike Staunton; Ken French US industry data; Cowles (1938) industry data; FTSE International UK industry indices

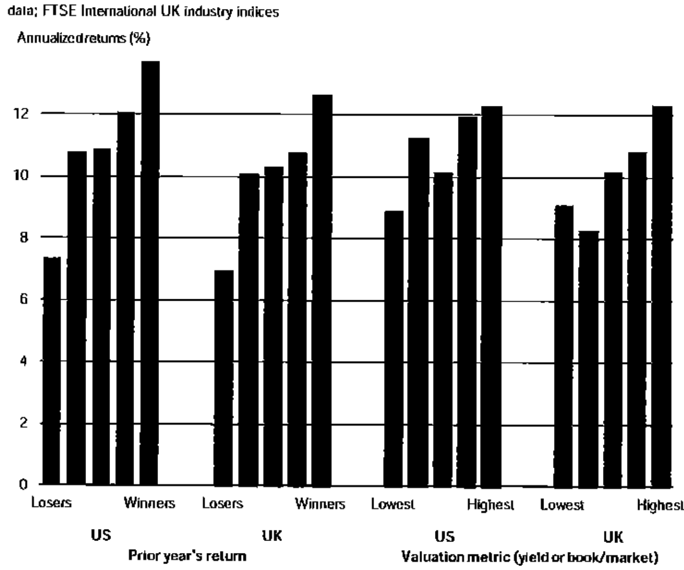


Figure 8 shows the reverse is true. There is substantial industry momentum, with winners tending to continue to win, and losers having a propensity to continue their losses. This is consistent with prior research. Moskowitz and Grinblatt (1999) claim that industry momentum accounts for much of the individual stock momentum anomaly. Grundy and Martin (2001) find the stock-specific component more important. Scowcroft and Sefton (2005) find that for large-caps, momentum is driven mostly by industries, but for small caps, it is largely driven by stocks.

The momentum effect shown in Figure 8 is substantial, and remarkably consistent between the USA and UK. It is not driven by volatility. In the USA, the winner portfolio had the same volatility as the losers, while in the UK, the winners had lower volatility. The Sharpe ratios, as well as the returns, are thus much higher for the winners than the losers. Furthermore, we would expect to find an even stronger effect if we shortened our holding/rebalancing period to the 1–6 months more typically associated with momentum strategies. Figure 8 certainly provides no evidence of an industry reversal effect over a one-year interval.

Industry value rotation

The two sets of bars on the right of Figure 8 relate to rotation based on a simple valuation metric (red bars for the USA and light blue for the UK). Each set of bars shows the annualized returns from investing in the lowest yield or book-to-market industries, through to the highest. Both low yield and low book-to-market are associated with growth industries. Companies from new industries and technologies tend initially to pay low dividends, retaining cash for growth and investment. Mature and declining industries with fewer prospects pay out more. In growth industries, a large part of market value will comprise capitalized future opportunities not yet reflected in book value or assets in place. Thus lower yield and book-to-market industries tend to be newer growth industries, while those with higher valuation metrics tend to be older "value" industries.

Figure 8 shows that, in both the USA and UK, there was an industry value premium, with "value" industries giving higher returns than "growth" industries. This might be because the value industries were riskier. However, in both the USA and UK, the standard deviations of the lowest and highest value quintiles were comparable. Similarly, the return differences are not explained by beta. The outperformance from investing in value-oriented, rather than growth industries is thus robust to standard risk adjustment. There could be other factors that explain the premium, such as tax. But the existence of the premium is consistent with there being periods of overvaluation for growth industries that the rotation strategy helps avoid, and periods of undervaluation for value industries that the rotation strategy helps exploit.

Industries today

Figure 9 shows sector weightings at start-2015 for the world, the USA, UK, Germany, Japan and emerging markets. The weightings are based on the 10 ICB (Industry Classification Benchmark) sectors, which cover broad groupings of industries. The world index provides a benchmark for judging over- or underweight positions elsewhere.

There are big differences between countries. The USA has a heavy weighting in technology (17%), high weightings in oil and gas, health care and consumer services, and lower weightings in basic materials, consumer goods and telecoms. The UK has a tiny weighting in technology (1%), but a high weighting in resources (over 25% in oil and gas plus mining stocks within basic materials) and financials (22%).

Germany and Japan have heavy weightings in manufacturing industries, and low or negligible (Germany) weightings in resources (oil and gas and mining). Germany's heavy weighting in basic materials (23%) is attributable to chemicals. Germany and Japan have high weightings in consumer goods, where automobiles are especially important. Both are underweight in health care and Germany is underweight in consumer services.

Emerging markets have a high weighting in financials (32%, of which two thirds is in banks), and are overweight the world index in oil and gas and basic materials. These three sectors make up almost half of emerging market capitalization. Emerging markets are also overweight in telecoms, very underweight in health care, and underweight in consumer goods and services.

Concentration of industries by country

Figure 10 shows that many industries are concentrated within particular countries. The ICB classification system divides the ten sectors in Figure 9 into 40 industries (although, confusingly, ICB uses different terminology, referring to sectors as industries and vice versa.) Figure 10 shows a subset of these more detailed industry groupings, highlighting the country with the largest weighting (in blue if it is the USA or red otherwise) and the second-largest weighting (in gray). The USA, which accounts for around half the world's capitalization, has the largest weighting in 33 industries.

The red bars in Figure 10 show the seven industries where the USA is not the largest player. Japan leads in automobiles, mobile telecoms, and electronics; Hong Kong in real estate; the UK in mining; China in alternative energy; and Korea in leisure goods. In addition, Figure 10 includes all industries where either the US weighting accounts for over two thirds of the world total, or else the weighting of the second-largest country exceeds 20%. The latter group, indicated by gray bars displaying country names, shows that the UK is a major player in life insurance and tobacco; Japan in industrial engineering and leisure goods; Germany in chemicals; Australia in mining; Switzerland in food; and Denmark in alternative energy.

Clearly, industries are highly concentrated within countries. In 35 of the 40 industries, the two countries with the largest weights account for over half the industry's global capitalization; in 30 industries, the top two countries account for more than 60% of industry weight; in 18 industries, they account for over 70%; and in seven industries, for over 80%.

Figure 9

Sector weightings in key countries and regions, start-2015

Source: FTSE International world index series

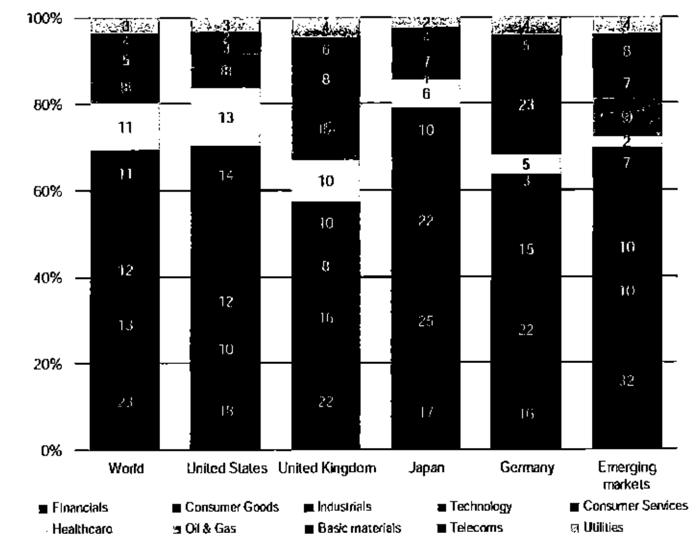
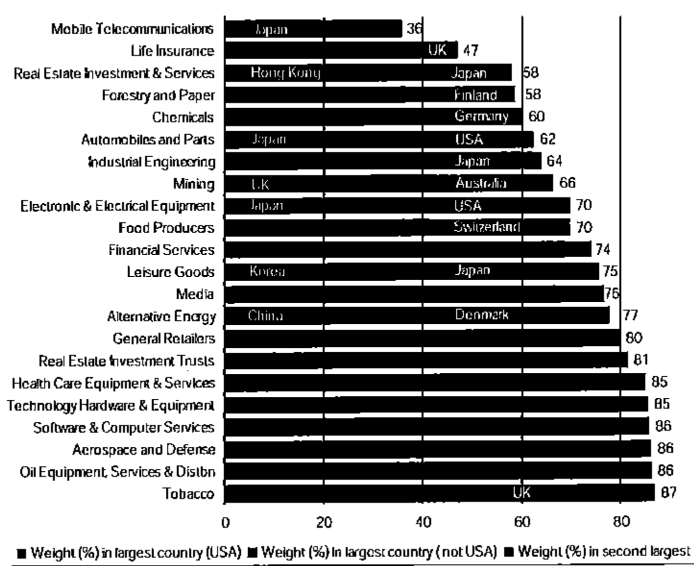


Figure 10

Concentration of industries by country, start-2015

Source: FTSE International world index series



Concentration of countries by industry

Just as industries can be concentrated within a few countries, so countries can be dominated by a handful of industries. Figure 11 shows the weight of the largest and three largest industries in 28 of the 47 countries in the FTSE All World index. It shows the five most concentrated by industry (at the top), the five least concentrated (at the bottom), plus all other Yearbook countries. In the five countries at the top, three or fewer industries make up the country's entire capitalization.

Figure 11

Concentration of countries by industry, start-2015

Source: FTSE International world index series

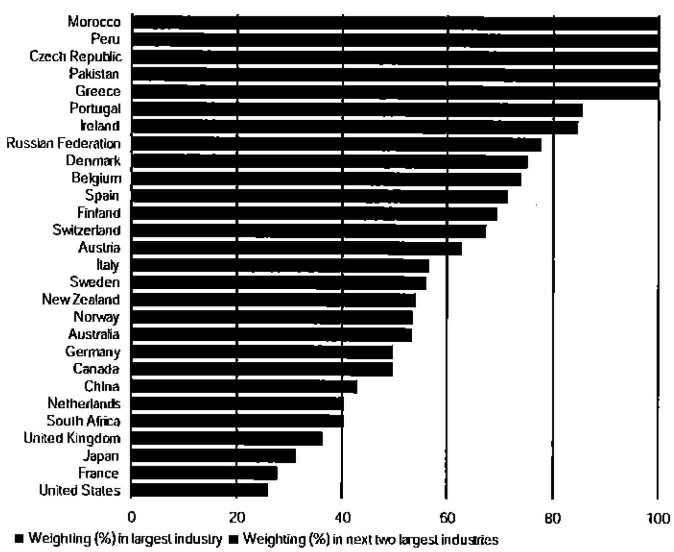
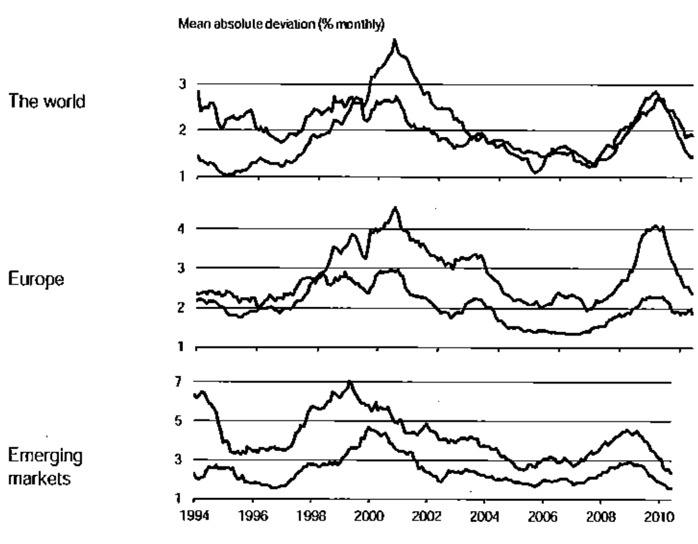


Figure 12

The relative importance of industries versus countries

Source: Jose Menchero and Andrei Morozov (2012)



The USA, Japan, France and UK are the least concentrated. But even here, the three largest industries (out of 40 in total) make up between 26% (USA) and 36% (UK) of country capitalization.

The weighting of the three largest industries accounts for at least 40% of country capitalization for 42 out of the 47 countries (including those not shown in Figure 11); for at least 50% for 33 countries; for at least 60% for 21 countries; and for 70% or more in 15 countries. The implications are clear. Investors in most countries will have poorly diversified portfolios, with heavy industry concentration if they restrict investment to their own country. This reinforces the imperative to diversify internationally. But many industries are concentrated within particular countries. This underlines the need for global diversification across countries in order to diversify effectively across industries.

Do industries or countries matter more?

An understanding of whether industries or countries matter more in impacting stock returns is important to global investors. It dictates whether asset allocation and active positions should be focused primarily on industries or countries; and whether diversification across industries or countries is likely to lead to the greater risk reduction. It also has implications for how research and research departments should be structured.

Early studies such as Lessard (1974) and Heston and Rouwenhorst (1995) found that country factors dominated industries. But globalization and developments such as the Eurozone seem likely to have reduced distinctions between countries. Indeed, more recent research shows that industries have gained in importance relative to countries. The challenge is to disentangle industry from country effects. The UK stock market accounts for 43% of the listed world mining industry. Oil accounts for 56% of the value of the Russian market. Separating out Russia from the oil effect, and mining from the UK is difficult.

A recent and thorough study by Menchero and Morozov (2012) uses a global factor model to address this issue. They investigated a large universe of stocks – all the constituents of the MSCI All Country World Investable Market Index – over the period 1994–2010. Figure 12 reproduces their results for one of their measures, mean absolute deviation, which quantifies the relative strength of industries versus countries.

The top panel of Figure 12 shows Menchero and Morozov's findings for the world as a whole. Until 1999, countries dominated industries, but during the dot-com bubble and bust, industries assumed greater importance. Since 2003, industries and countries have been roughly equally important. The middle panel shows that within Europe, countries dominated until 1998, but since the introduction of the euro in 1999, industries have been more important. The bottom panel shows that for emerging markets, countries have dominated industries throughout, although the difference between them has declined.

Concluding remarks

Industries are a key investment factor. Many countries' stock markets are highly concentrated within a few industries, while many industries are concentrated within a few countries. To exploit diversification opportunities to the full, investors need to diversify across a wide spread of industries and countries. Both matter, although there is evidence that globalization has led to a decline in the relative importance of countries, with industries assuming greater importance.

Industries have risen and fallen over the years as technology has advanced. It is interesting to see which have done best and worst, but this tells us little about the future. The industrial landscape will change during the 21st century perhaps even more radically than in the past. As Charles Duell, commissioner of the US Patent and Trademark office said in 1902, "In my opinion, all previous advances in the various lines of invention will appear totally insignificant when compared with those which the present century will witness." Investors must focus on the future.

What we can say with confidence is that there will continue to be a wide variation between the returns on different industries. It will remain hard to predict the likely winners and losers, but industries and their weightings will continue to matter.

Should investors focus on new industries and shun the old? Or should they be contrarian? We have seen that both new and old industries can reward as well as disappoint. It all depends on whether stock prices correctly embed expectations. New industries are typically born on a wave of IPO activity, and we have seen that investors should be especially cautious about the valuations of IPOs and unseasoned stocks.

One way of leaning against any tendency to overvalue the new and undervalue the old is to follow an industry value rotation strategy. This has historically generated a premium. This is consistent with there being periods of overvaluation for growth industries which this strategy helps avoid; and periods of undervaluation for value industries which the strategy helps exploit.

But momentum appears to be an even more effective rotation strategy. Buying last years' best-performing industries while shorting the quintile of worst performers would, since 1900, have generated an annualized winner-minus-loser premium of 6.1% in the USA and 5.3% in the UK. Before costs, US investors would have grown 870 times richer from buying winning industries rather than losers.

If these rotation strategies were to continue working, which industries appear most and least attractive at the start of 2015? In the USA, utilities, insurance, transport and healthcare are favored, while leisure, software, electrical equipment and beverages look least attractive. In the UK, utilities, tobacco, pharmaceuticals and life assurance rank highest, while the laggards are technology hardware, aerospace, industrial engineering and electronic and electrical equipment.

This is just an illustration. It is emphatically not a recommendation. Historically, the rotation strategies, even when combined, have failed in around one year in three. In the case of momentum, the failures can be especially painful at market turning points. But for the patient, long-run investor who can weather such episodes, the past success of these strategies may provide food for thought.


FOTO: TANUSHA / FOTOLIA.COM

Responsible investing: Does it pay to be bad?

Investors are increasingly concerned about social, environmental and ethical issues, and asset managers are under growing pressure to demonstrate responsible investment behavior. This can take the form of "exit" via ethical screening, or "voice" through engagement and intervention. We show in this article that "sin" can pay, not least because those choosing to exit "sinful" stocks can cause them to offer higher returns to those less troubled by ethical considerations. However, the expected financial impact of modest exclusions is generally small. We also provide evidence that corporate engagement can pay, whether the focus is on environmental and social issues or on corporate governance.

Elroy Dimson, Paul Marsh and Mike Staunton, London Business School

Some investors take a laissez faire approach, investing where returns seem most promising, and ignoring social, environmental and ethical issues. Others take an approach that they regard as more "responsible." There are three reasons for choosing to be a responsible investor. First, the owners of businesses share in responsibility for the firms' actions. Second, they can induce them to improve corporate behavior. And third, long-run returns may be enhanced by ensuring that companies have high standards of behavior.

The owners' responsibility includes events like BP's rig explosion (Deepwater Horizon), Union Carbide's gas leak (Bhopal), Lonmin's labor relations (Marikana), Exxon's oil spill (Exxon Valdez), Tokyo Electric Power's meltdown (Fukushima), Massey Energy's mine explosion (Upper Big Branch), and the Savar Building collapse (Rana Plaza). It also embraces dishonesty and malfeasance. For example, there are a number of documented cases with such well-known companies as Lockheed (bribery), Siemens (corruption), Enron (false accounting), Walmart (child labor), and Mattel (lead paint).

The laissez faire view is losing ground. The world's largest asset owners now devote extensive resources to social and environmental issues and corporate governance, and to engaging with investee companies on these issues.

The extent of engagement is reported to be at an all-time high. The UN-supported Principles for Responsible Investment lists 1,349 signatories with assets of over USD 45 trillion, around half the assets of the global institutional investor market (Shubb, 2014). The Global Sustainable Investment Alliance estimates that worldwide some USD 14 trillion of professionally managed portfolios incorporate environmental, social and governance concerns into their decisions.

Corporations and their executives also wish to be seen as responsible, with a commitment to delivering broader benefits, not just financial rewards. Under the UN Global Compact, more than 12,000 business organizations in 145 countries have committed to responsible and sustainable corporate practices.

Why be good?

The motivations for taking a responsible approach to investing include complicity, influence and universal ownership. The notion of complicity underpins the screening processes followed by the

Norwegian Government Pension Fund Global; see Nystuen, Follesdal and Mestad (2011). For the Norwegians, "owning shares or bonds in a company that can be expected to commit gross unethical actions may be regarded as complicity in these actions" (Graver, 2003). Some faith-based investors veto investing in certain companies (e.g. alcohol) on the grounds that such businesses are offensive to their values.

Influence, or "leverage" in the terminology preferred by Richardson (2013), seeks to persuade companies to behave differently. The leverage that asset owners have may enable them to persuade the executives of businesses that they own – or perhaps their regulators, judiciary or other influencers – to improve their behavior. The improvement may be motivated by social justice and/or the interests of stakeholders.

The third motivation relates to the fact that the very largest asset owners are increasingly "universal owners," a term proposed by Monks and Minow (1995). They are now so large that they essentially own every company in the market. Furthermore, many of them have investment horizons that extend into the distant future. Universal owners cannot escape costly, company-specific factors: if one investee company benefits at the expense of creating additional costs for another, there may be no net gain to an asset owner with shares in both. Logically, universal owners should focus on increasing the size of the cake – the aggregate value of all corporations – rather than being too concerned about how the cake is sliced up between companies.

An example of this broader focus is labor practices. Some investee companies may lower production costs by employing children, or by sourcing from companies that employ children, but they are unlikely to pay the costs of poor child health or under-education. The universal owner may recognize that child labor in one firm reduces the profitability of other firms who do not employ children, and that impaired education may impede broader economic progress. From a long-term perspective, the owner can therefore benefit financially by engaging with companies and regulatory authorities. This is the business case, but there is, of course, also an ethical case. Similar arguments may be put forward in relation to corruption, nuclear proliferation, climate change and other societal issues.

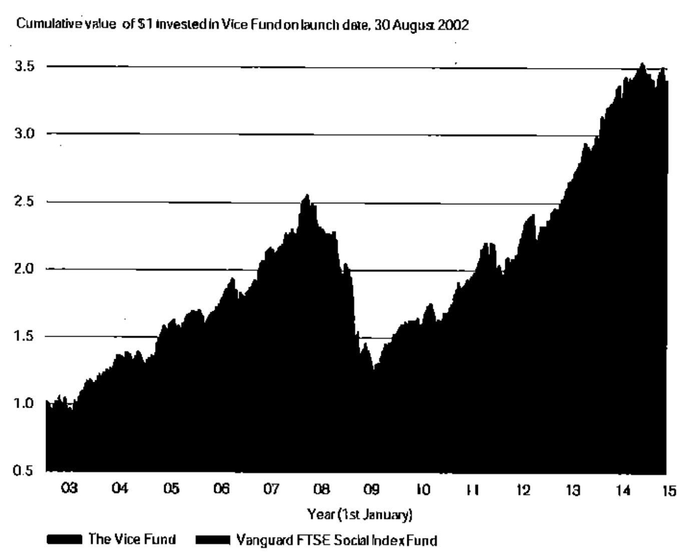
There is a small number of universal owners, if that term is taken literally, such as the Norwegian Government Pension Fund Global, the California Public Employees' Retirement System (CalPERS), the California State Teachers' Retirement System (CalSTRS), New York City Employees' Retirement System (NYCERS), the Universities Superannuation Scheme (USS) and the BT Pension Scheme. There are also many investment managers, notably of globally diversified passive funds, who interact with investee companies in the interests of large numbers of investors with more modest wealth.

The universal ownership approach shares some of the methods associated with complicity and influence. However, it is based on the notion that financial rewards can accrue from taking a broad view of responsible corporate behavior; see Dimson, Kreutzer, Lake, Sjo and Starks (2013). While this may be true of the very largest investors, especially sovereign funds, for most institutions there is a risk that an investor practicing responsible investment to the greatest possible extent could forego immediate investment returns in violation of fiduciary obligations. The gains from pursuing a universal ownership approach may be too unquantifiable or too costly in immediate financial terms. Whether this is the case is an empirical issue, which we address below.

Figure 1

The Vice Fund and Vanguard FTSE Social Index Fund 2002–14

Source: Morningstar. Data from 30 August 2002. Dividends reinvested. Charges not deducted.



Exit and voice

The political scientist and economist Albert Hirschman describes the two responses open to members of an organization when they perceive that it is demonstrating a decrease in benefit to its members. On the one hand, they can "exit" – that is, they can withdraw from the relationship. On the other hand, they can "voice" – in other words, they can speak out in an attempt to improve the relationship through communication of the complaint, grievance or need for change.

Admati and Pfleiderer (2009) refer to exit as the Wall Street Walk'. It may simply be a screening out or selling decision. But for an active owner, exit may be a more political action intended to apply pressure on the company or industry in question. If it is a coordinated activity, exit involves concerted disinvestment intended to persuade a business, industry, or nation to change its policy or regime.

Exit and the wages of sin

Figure 1 plots the cumulative returns, including reinvested dividends, on two US mutual funds launched in the early 2000s. The relative winner was the Vice Fund, whose excellent investment performance (USD 10,000 growing from inception to USD 33,655 at start-2015) earned it a top rating from Lipper and Morningstar. On the other hand, the Vanguard FTSE Social Index Fund, which had lower investment growth over the same interval (USD 10,000 growing to USD 26,788), was the relative loser. During this period, the S&P500 had performance midway between these two funds.

The Vice Fund invests in businesses that are considered by many to be socially irresponsible. Recently renamed the Barrier Fund, it has assets of USD 290 million invested in "industries with significant barriers to entry, including tobacco, alcoholic beverage, gaming and defense/aerospace industries." The Social Index Fund tracks an index screened by social, human rights, and environmental criteria. Constituents have superior environmental policies, strong hiring/promotion records for minorities and women, and a safe workplace. There are no companies involved in tobacco, alcohol, adult entertainment, firearms, gambling, nuclear power, and unfair labor practices. It has assets under management of USD 1.5 billion, over five times that of the Vice Fund.

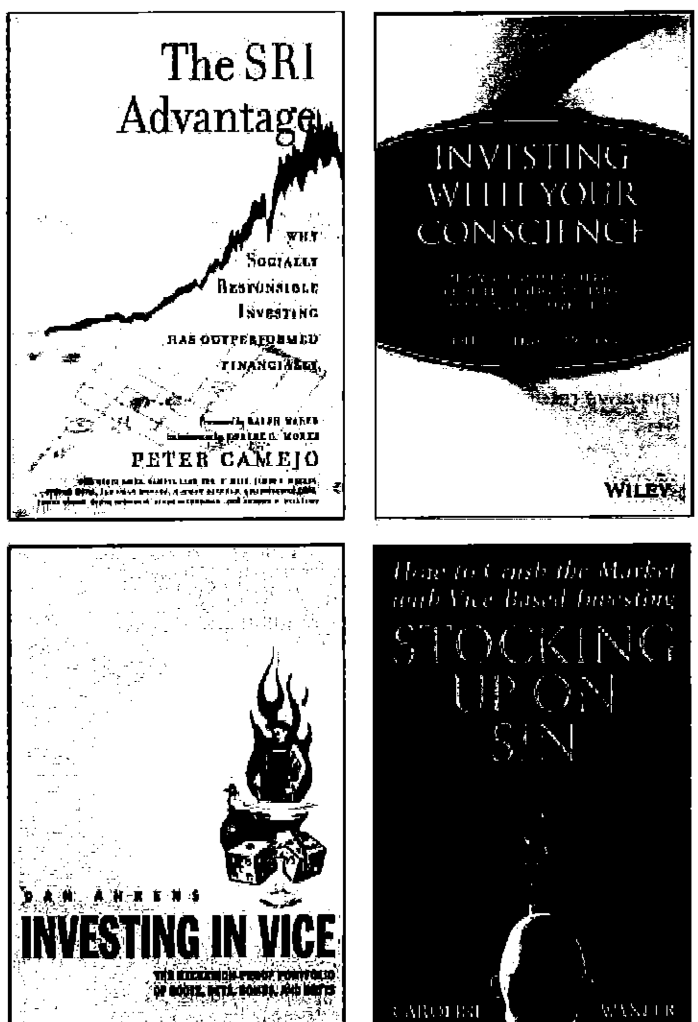
Many ethical investors emphasize "doing well by doing good." They consider that investing in responsible and principled companies is likely to be rewarded in the long run by better stock market performance. In The SRI Advantage: Why Socially Responsible Investing Has Outperformed Financially (Figure 2), Peter Camejo explains that he "presents overwhelming evidence that SRI has outperformed financially, explains in detail why SRI outperforms, and then examines the implications for investment professionals, investors, pension funds, and community/non-profit groups."

John Harrington makes similar claims in his book (also Figure 2) Investing with Your Conscience: How to Achieve High Returns Using Socially Responsible Investing.

Figure 2

Books on responsible and sin-based investing

Source: Publishers. Acknowledgements are included in the References.



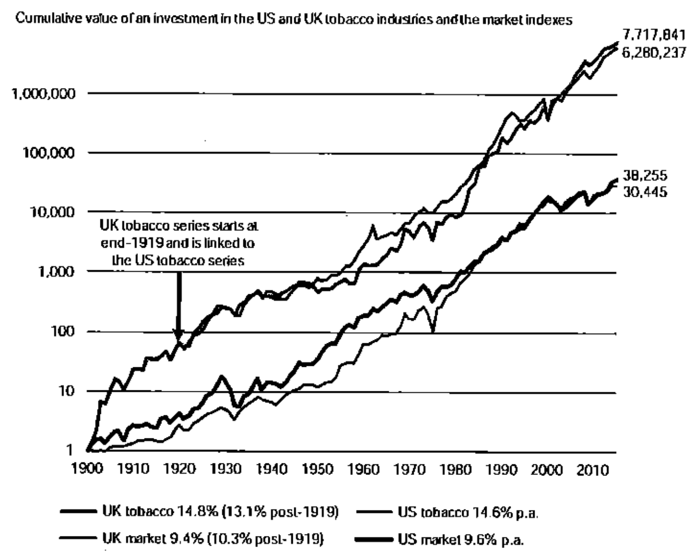
In reality, however, much of the evidence that we review suggests that, as illustrated by the Vice Fund, "sin" pays. Investments in unethical stocks, industries and countries have tended to outperform. For those for whom principles have a price, it is important to know the likely impact screening may have on both performance and diversification. Also, ironically, responsible investors should recognize that they may be partly responsible for the higher returns from sin.

The standard argument is that irresponsible businesses can be disciplined by the threat of divestment of the firm's shares. The assumption is that downward pressure on the share price will make the company less valuable, pushing up its cost of capital to the detriment of its ability to raise finance, and possibly raising the likelihood of a takeover bid. Lower stock prices will also punish executives where it hurts – through their pay – according to this point of view.

Figure 3

Cumulative returns on tobacco and on equities, 1900–2014

Source: Elroy Dimson, Paul Marsh and Mike Staunton. Currencies are nominal USD and nominal GBP



As Dan Ahrens explains in his book, Investing in Vice: The Recession Proof Portfolio of Booze, Bets, Bombs and Butts (Figure 2), it can be profitable to invest in stocks that ethical investors abhor. The rationale for "vice investing" is that these companies have a steady demand for their goods and services regardless of economic conditions, they operate globally ("vice" is a worldwide phenomenon), they tend to be high-margin businesses, and they are in industries with high entry barriers. Yet, if a large enough proportion of investors avoids "vice" businesses, their share prices will be depressed. Appealingly to Dan Ahrens, if companies have a lower stock price, they offer a buying opportunity to investors who are relatively untroubled by ethical considerations. Caroline Waxler has a similar interpretation in Stocking Up on Sin: How to Crush the Market with Vice-Based Investing (also Figure 2).

The paradox, then, is that depressed share prices for what some regard as noxious and nasty businesses may demonstrate that responsible and ethical investors are having an impact on the value of a company whose activities conflict with social norms. If so, the shares will ultimately sell at a lower price relative to fundamentals. For example, they may trade at a lower price/earnings or lower price/dividend ratio. Buying them would then offer a superior expected financial return which, for some investors, compensates for the emotional "cost" of exposure to offensive companies.

Exit from companies and industries

A number of companies have been excluded by investors where there is a record of contributing to severe environmental damage, serious violations of societal norms, or systematic human rights abuses. Even for investors who tolerate aerospace and defense, there can be exclusions of businesses involved, sometimes indirectly, in nuclear-arms manufacture, antipersonnel landmines and cluster munitions. Several large funds follow the recommendations of the Norwegian Council on Ethics.

It can be hard to find universally shared evaluations. Despite high ratings from Kinder, Lydenberg, Domini & Co (KLD), Walmart was divested by Norway's sovereign fund because of its unacceptable labor practices. Despite its credentials as a purveyor of Fair Trade coffee, Starbucks has become a boycott target because of its UK tax avoidance practices. Despite its inclusion in the Dow Jones Sustainability World Index and FTSE4Good Index, Medtronic has been excoriated in the USA for its tax-inversion scheme. Amazon, a company praised for environmental initiatives, is accused in Europe and the USA of anticompetitive tax arrangements. The sin of theft has now been extended to being perceived to be robbing the state of the tax that societies believe companies ought to pay.

For a long-term perspective, we can gain a deeper insight by looking at the impact of eliminating businesses that violate established norms. In their paper, The Price of Sin, Harrison Hong and Marcin Kacperczyk examine "sin stocks" traded in the USA over the period 1926–2006; they also look at the European experience over the period 1985–2006. They define sin stocks as companies that are involved in the "triumvirate of sin" (alcohol, tobacco and gambling) and, in some additional tests, weaponry. The authors show that institutional investors tend to avoid sin stocks, which typically sell at a lower price in relation to fundamentals. They report larger expected returns for these shares.

Tobacco companies are particularly informative. For the first half of the 20th century, tobacco was not widely regarded as harmful. But by the mid-1950s, there was a confluence of four diverse strands of evidence – epidemiology, animal experiments, cellular pathology and carcinogens in tobacco smoke – and the causal association between smoking, particularly cigarette smoking, and lung cancer was established. By then, at least some investors were shunning the tobacco business.

Hong and Kacperczyk date the transition of tobacco companies from neutral to "sinful" status as occurring in the USA during 1947–1965. Over this interval, they observe an underperformance of 3% per year. After 1965, when the health impact of tobacco became well known, US tobacco companies outperformed comparable firms by more than +3% per year over the period 1965–2006. Moreover, even though US tobacco companies faced a barrage of litigation during this period, they outperformed their international peers.

We use our own industry indices – taken from the accompanying article – to estimate the outperformance of tobacco stocks over a complete 115-year period. As Figure 3 shows, tobacco companies beat the overall equity market by an annualized 4.5% in the US and by 2.6% in the UK (over the slightly shorter 85-year period of 1920–2014). Over the entire 81 years of the Hong-Kacperczyk study, US sin stocks provided an annualized excess return, relative to non-sin stocks of 3%–4% per year. During 1985–2006, international sin stocks outperformed by around 2½% per year.

In another study entitled Sin Stock Returns, Fabozzi, Ma and Oliphant examine a larger number of sin stocks, drawn from multiple markets from 1970–2007. The authors used a carefully enunciated definition of sinful activity, and included only the 267 stocks that were deemed sufficiently liquid to be investible. Figure 4 sets out their findings. Each stock has a unique start- and end-date in the sample, so the authors compute the excess return for each stock relative to the return on the capitalization-weighted index for the market over the interval for which it trades. They find a high level of performance from investing internationally in sin stocks. Averaged within sin categories, the mean excess return varies from a low of 5.3% (alcohol), through 9.6% (biotech), 10.0% (adult services), 14.7% (tobacco) and 24.6% (weapons), to a high of 26.4% (gaming).

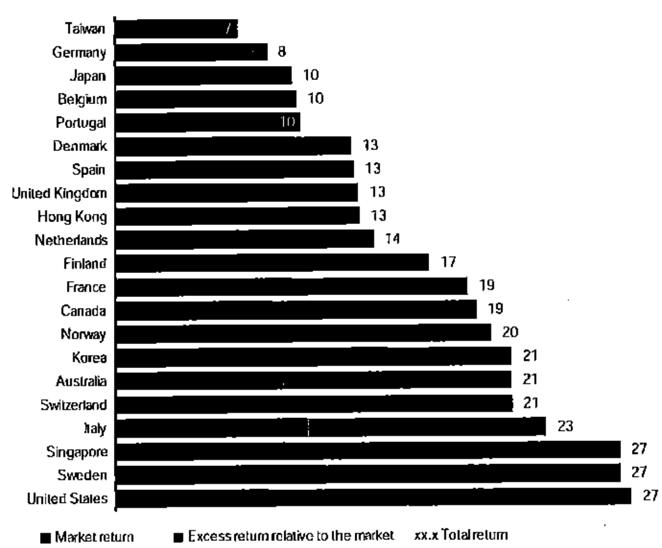
Their average performance varies by country, but is systematically high. In only two cases is the excess return negative (Taiwan –2%; Portugal –1%). The excess return is statistically significant in all but three markets. Both papers find that their measures of performance are barely impacted by the choice of performance measurement criterion.

Apart from the capacity limits of equal-weighted strategies, there are other impediments to profiting from sin stocks. First, there are not many "pure play" sin stocks: out of thousands in the US investment universe, Hong and Kacperczyk identify only 193 examples in the 81 years they study, only 56 of them alive by 2006. Second, a sin-stock portfolio is undiversified. Third, vice investing is unconventional (no competitors have emerged for the Vice Fund). Fourth, there are no examples of a sin-stock Exchange Traded Fund (FocusShares launched one based on the ISE's *SIN*dex, but it failed to attract investors). We have discussed exit from individual companies and industries. In addition, there can be an even broader approach to exit. We consider next the idea of boycotting an entire market.

Figure 4

Annual returns on sin stocks in 21 countries, 1970–2007

Source: Fabozzi, Ma and Oliphant (2008). For Taiwan and Portugal the excess return was negative



Exit from countries

Various countries have at some time, recently or decades ago, been subject to divestment pressures and investment boycotts. They include Burma/Myanmar (EU sanctions), Cuba (Helms-Burton Act), Iran (Sanctions Enabling Act), Israel (BDS movement), Russian Federation (EU sanctions) and South Africa (Anti-Apartheid Movement). There are about 150 countries with stock exchanges, but about half of them are omitted by all the major index providers. Most global indices omit frontier markets, but some investors may also select benchmarks that omit secondary emerging markets or even primary emerging markets, while some investors simply decide for themselves which markets are acceptable. Our point is this: for almost all investors, individual national markets are screened out of their portfolios.

We examine the impact of screening out countries based on their degree of corruption. Countries are evaluated using the Worldwide Governance Indicators compiled by Kaufmann, Kraay and Mastruzzi (2010) and supported by the World Bank. The indicators comprise annual scores on six broad dimensions of governance in 215 countries from 1996 to date. While we have singled out the corruption indicator, it is highly correlated with the five other measures. The main score is a percentile ranking across all nations that runs from zero to 100.

Figure 5 shows the geographical distribution for the corruption score. Corruption is not a major feature among the 21 markets with a complete Yearbook history; it is more prevalent among developing markets.

Country exclusion based on corruption

We estimate the total return since 2000 for stock markets in countries ranked by their corruption score at that date. To measure investment returns, we use the indices for each of the 47 countries in the FTSE Global Equity Index Series (GEIS). We compute the equally weighted average of the equity market returns within a grouping, where each index falls in one, and only one, grouping. All returns are expressed in common currency (US dollars) and they include the performance of markets that were deleted from GEIS, or ceased to exist. We deem Venezuela, which was removed on 20 June 2003 with a total return index value on the 19th of 94.78 and on the 20th of zero, to have lost 99.99% of its value.

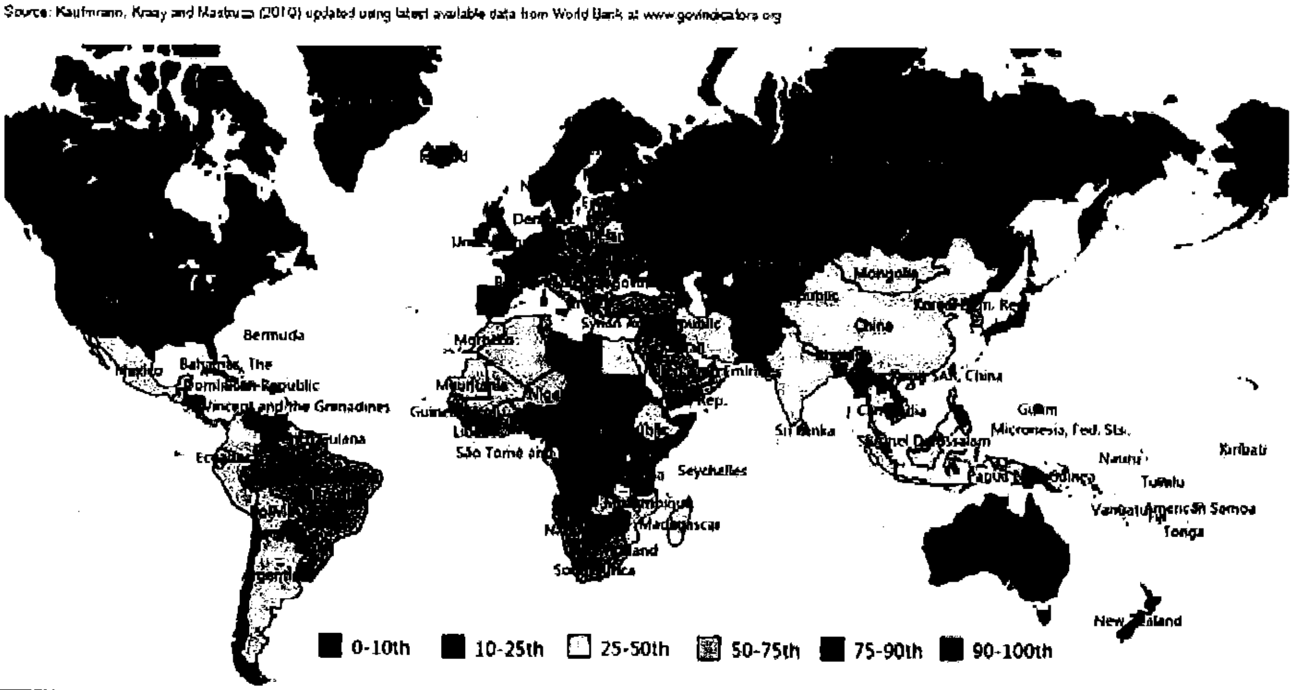
Figure 6 shows the results for three groupings: Yearbook countries with a continuous history since 1900 versus others; countries grouped by the date of eligibility for entry to the FTSE GEIS series; and the World Bank's corruption indicator. We set the scene with the left-hand (red) bars, which report the average return since start-2001 of the 21 Yearbook markets, which provided an average annualized return of 5.9%. Indexes with a shorter history had a 10.4% annualized return.

The middle (blue) bars present information on countries ranked by the timing of entry to the FTSE series, which began on 31 December 1986 with 23 countries. Seven were added during the 1990s

Figure 5

Control of corruption scores around the world, 2014

Source: Kaufmann, Kraay and Mastruzzi (2010) updated using latest available data from World Bank at www.govindicators.org



(strictly, six in the 1990s plus one in 1988), and a further 18 during 2000. The average returns were 7.4% per year for the original countries, 8.1% for the additions of the 1990s, and 10.4% for those that entered in 2000. The latter cohort included a substantial number of emerging markets with lower standards of governance.

The right-hand (gray) bars report the post-2000 returns for markets ranked by the corruption indicator. There are 14 with a poor score, 12 that are acceptable, 12 that are good, and 11 with excellent scores. The average returns for the last three groups were between 5.3% and 7.7%. In contrast, the markets with poor control of corruption had an average return of 11.0%. Realized returns were higher for equity investments in jurisdictions that were more likely to be characterized by corrupt behaviors. This pattern is time specific, and there are sub-periods when more "saintly" markets did better than the sinners. Because the interval we study is short, our results may simply reflect a period when emerging markets outperformed.

While our findings could be attributable to many factors other than corruption, we are sympathetic to the view that low standards of governance may be regarded as a priced risk factor. Luo and Balvers (2014) introduce a boycott factor into asset pricing that reflects the extent to which subgroups of investors have a non-pecuniary preference to avoid certain groups of stocks. Their model is supported by carefully conducted empirical testing based on US data, and could easily be extended to country exclusions.

An application of country exclusion was adopted in 2002 by CalPERS, whose Permissible Emerging Market Policy blacklisted entire countries that fell short of a minimal threshold on factors such as political stability, democratic institutions, transparency, labor practices, corporate responsibility and disclosure. The resulting restriction on investing in Russia, China and other (then) high-performing emerging markets was costly: "by late 2006, CalPERS' emerging market portfolio had been subject to 2.6% in annual opportunity cost of foregone return, totaling over USD 400 million in losses from the time of the policy's inception" (Huppé and Hebb, 2011). In 2007, CalPERS dropped its emerging-market country withdrawal strategy, and switched to a principles-based approach to selecting companies in the developing world. They chose to use voice rather than exit within emerging markets, and embraced dialogue, engagement and shareholder activism.

Using voice

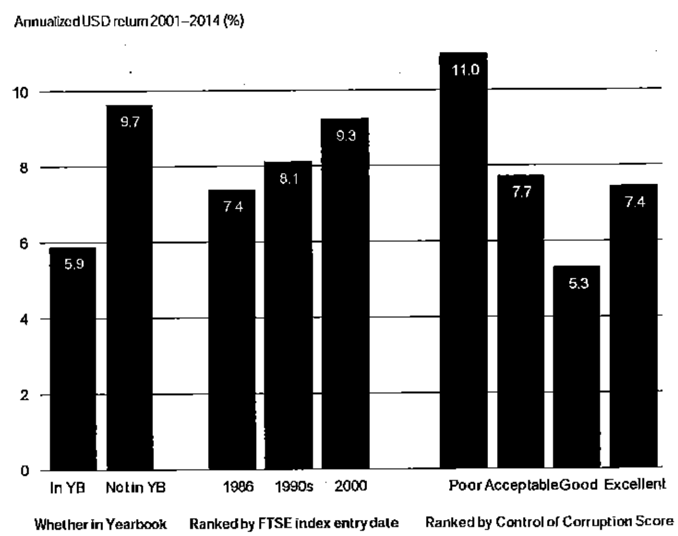
Investors have a voice, which is lost when they exit from a company. They use their voice by means of activism, with the aim of influencing corporate behavior. They may hold discussions with executives, send written communications, submit and vote on proxy proposals, and seek to influence regulators and standard setters. These activities are undertaken in the belief that responsible investors can guide management toward improved financial performance and/or enhanced social conditions for stakeholders and communities. Hedge funds are also major activist investors, mostly seeking to maximize investment returns.

The most visible activists seek to achieve stock market performance by improving corporate governance or migrating ownership to a management willing to work more effectively in the interests of shareholders. These activists, who focus on investor returns rather than corporate social responsibility, eschew exit from unsatisfactory companies, and indeed often target "entrance" to such companies where they use voice to facilitate change.

Figure 6

Return on markets ranked by corruption tendency, 2001–14

Source: Elroy Dimson, Paul Marsh and Mike Staunton, DMS data, FTSE indexes, World Bank



There is a considerable body of research on the use of voice in corporate engagement. Greenwood and Schor (2009) examine activist-driven takeovers in the USA and suggest that hedge funds may be better suited to identifying undervalued targets and prompting a takeover, rather than at engaging in long-term corporate governance and improving companies' operations.

More recently, Becht, Franks, Grant and Wagner (2014) examine 1796 cases of public activism across three regions (Europe, North America and Asia). They find that market-adjusted returns at block-disclosure are 4.5%–7.5% (depending on the region), followed by outcome-announcement returns of 3.0%–9.3% (also depending on the region, Asia being the lowest). Returns are higher for successful engagements than for those that fail to achieve their objectives.

Figure 7

Cumulative return around engagement disclosure, 2000–14

Source: Becht, Franks, Grant and Wagner (2014)

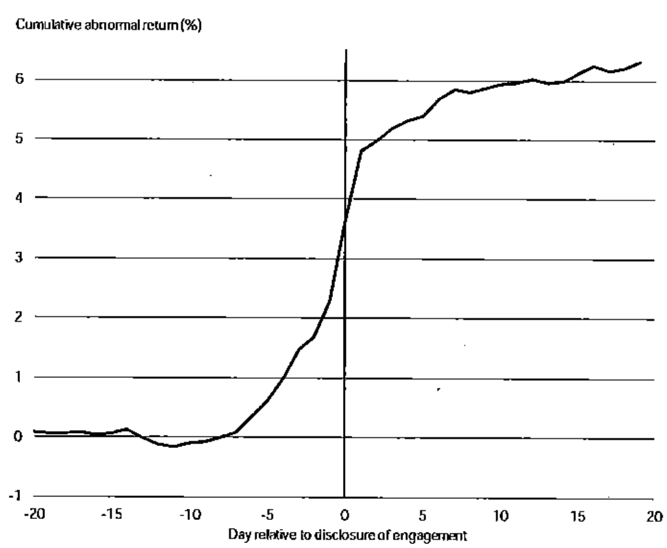


Figure 7 plots the cumulative market-adjusted returns for the pooled sample from all three regions, from 20 days before to 20 days after the activist engagement is disclosed through initial filing or press coverage. There is some pre-disclosure drift in returns that anticipates the engagement and some post-disclosure drift that reflects the consensus view on the likely outcome.

In addition to the positive market reaction to the announcement of an activist block holding, there is also a further reaction on the announcement of a successful outcome. Combining both components, the average returns exceed 15%. The study reports that coordinated activism tends to generate higher returns than individual activism. Use of voice on corporate strategy has a positive financial value, but what can we say about social and environmental issues? Does the evidence we have on corporate governance interventions carry over to engagements on issues to do with responsible investing?

Voice on social issues

As dialogue with investee companies has expanded, there has been a corresponding growth in articles on shareholder engagement. Unfortunately, however, there has been little high-quality research on the impact of interventions on corporate social responsibility. Most studies rely on static and delimited measures of corporate social responsibility, such as the social-responsibility scores produced by KLD, now known as MSCI ESG. Data like this can help establish whether a company that is profitable tends to spend more on corporate social responsibility activities; it cannot establish that spending on corporate social responsibility tends to make a company more profitable. Documenting correlations leaves some of the most basic questions about corporate social responsibility activism unanswered.

Active ownership

To provide more evidence, Dimson, Karakas and Li (2015) draw on a large, proprietary database of environmental, social and governance engagements. Their study, which examines US public companies, addresses questions including: which firms do active owners engage and how are these engagements executed? Do active owners compete or collaborate with other shareholders and with what effect? How do engaged firms respond? What determines the success of engagements? How does the market react to engagements on social responsibility issues? Do active owners succeed in implementing their objectives? And how do these activities affect financial performance?

The dataset analyzed in this study is unusual in being a point-in-time record of active engagements. It was provided by an institutional investor that actively engages in dialogue with target companies (4,000 of them in 2013) and exercises ownership rights at shareholders' meetings (voting on 60,000 resolutions in that year), achieving the change sought in 7% of cases. The primary sample consists of 2,152 engagement sequences (1,252 social, environmental and ethical, and 900 corporate-governance-based sequences) for 613 public firms between 1999 and 2009. The success rate for engagements is 18% and, on average, it takes a sequence of 2–3 engagements over 1–2 years until success can be recorded.

Compared to a matched sample of companies, firms are more likely to be engaged if they are large, mature, and performing poorly. The likelihood of being engaged is further increased if the asset manager has a large shareholding in a particular firm, if other socially conscious institutional investors (e.g. pension activists or ethical funds) are shareholders, if there are reputational concerns for the target company and if it has inferior standards of governance.

An analysis of the engagement features and tactics shows that successful prior engagement experience with the same target firm increases the likelihood of subsequent engagements being successful. In addition, collaboration among the asset manager and other active investors and/or stakeholders contributes positively to the success of engagements, particularly for the social, environmental and ethical engagements.

Figure 8 provides evidence on how the market reacts to activism on corporate social responsibility, showing post-engagement performance for the entire sample. Cumulative abnormal returns (CARs) are based on total returns, which are adjusted for Fama-French size-decile matched returns. The sample is partitioned into the CAR for successful engagements (i.e. those that accomplished the objectives set out prior to engagement) and the CAR for unsuccessful engagements (i.e. those that did not).

Dimson, Karakas and Li find that corporate social responsibility engagements generate a cumulative size-adjusted abnormal return of +2.3% over the year following the initial engagement (see the blue line in Figure 8). Investment performance is superior for successful engagements (+7.1%) and gradually flattens out after a year (red line) when the objective is accomplished for the median firm in our sample. There is a neutral market reaction to unsuccessful engagements (gray line).

The abnormal returns are similar for successful environmental/social and successful corporate governance engagement; and similar for unsuccessful environmental/social and unsuccessful corporate governance engagements. In other words, investors placed much the same financial value on successful social activism as on successful governance interventions. Active ownership provided stakeholder benefits and did not destroy firm value even when engagements were unsuccessful. Of course, this study focusses on a single and, in hindsight, successful example of shareholder activism. Activism by an under-skilled or under-resourced team risks a lower payoff. It should also be noted that engagements by an active owner relate to a small part of a large portfolio, and their impact on overall performance will be modest. Finally, a caveat: these rewards from active ownership may be period-specific and, in a subsequent period, the benefits from engagement could wane.

The choice of exit or voice

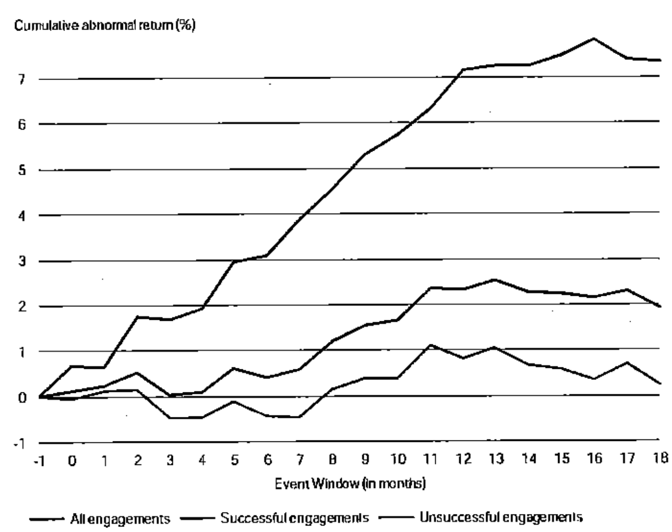
Investors already recognize there is a corporate governance dividend attached to firms: they command a higher market valuation, have cheaper access to capital and benefit from a strong shareholder base. We have reported here on a study that indicates there is a corporate social responsibility premium attached to firms: they are likely to attract additional investors, avoid environmental and social mishaps, and sell at a higher multiple.

For corporations and shareholders, adherence to superior ethical principles is important, and it impacts on their overall performance. Investors increasingly demand greater transparency from companies about their governance principles as well as their environmental policies and practices, their record of protecting human rights and their contribution to the communities in which they operate.

Figure 8

Cumulative abnormal returns (CARs) after engagement

Source: Dimson, Karakas and Li (2015). Fama-French size decile returns from Professor French's website



Consistent with responsible companies trading at a premium price, companies that violate social norms sell for lower stock prices, and there have consequently been favorable investment returns from sin stocks. As long as those investor preferences persist, one should expect higher total returns from sin stocks. If the "sin" discount stays constant, the expected capital gain is the same for sin and non-sin stocks: the excess returns to sin stocks should then come in the form of higher dividends over time. Contrary to conventional wisdom, this gives "sin" investors an incentive to have longer investment horizons.

Investors therefore face a challenge: should they exit from objectionable stocks, or should they use voice to make target companies more acceptable? The decision will often depend on how strongly they feel about loyalty – retaining their stake in the company – and about the scope for changing the company for the better. It will also depend on the potential return loss from omitting objectionable stocks as well as the loss of diversification. For large institutions, an added consideration may be the cost of exiting a very large position.

Figure 9

The forward-looking impact of exclusions on the World index

Source: Ang (2014, p.109). Non-BDT denotes non-bank, non-defense, non-tobacco.

Industries admitted to the index portfolio				Optimized portfolio	
Non-BDT	Banks	Defence	Tobacco	Minimum volatility (%)	Maximum Sharpe Ratio
✓	✓	✓	✓	12.05	.4853
✓	✓	✓	✗	12.10	.4852
✓	✓	✗	✗	12.10	.4852
✓	✗	✗	✗	12.10	.4853

Andrew Ang (2014) estimates the impact of exclusions on forward-looking estimates of risk and return. He analyses the FTSE All World index portfolio which, in 2012, comprised 39 industries and 2871 stocks. He estimates industry risk and correlation (for the technically minded, using Ledoit-Wolf estimation) and expected returns (using Black-Litterman forecasts). Using a risk-free interest rate of 2%, he then computes optimal portfolio composition with various industry exclusions, and reports the risk of a minimum-volatility portfolio and the maximum reward-to-risk ratio (the Sharpe Ratio) for each screened portfolio.

Figure 9 reports his results. The first row shows the minimum risk and maximum Sharpe Ratio when all industries are deemed investible. The next three rows shows the impact of exclusions, which raise the risk of the minimum volatility portfolio from 12.05% to 12.10%; and reduce the maximum Sharpe Ratio from 0.4853 to 0.4852. In terms of expected risk and return, the penalty for screening is small.

We have shown in the companion article of this Yearbook that industries can have discernible factor exposures, and this raises the possibility that sin stocks are nevertheless likely to behave differently from their more attractive counterparts. The realized return from exiting may therefore turn out to be a financial disappointment or an investment success.

Following a rather different methodology, Humphrey and Tan (2014) simulate portfolios that mimic the characteristics of socially responsible mutual funds by using stock-level data and avoiding the confounding impact of cost-drag on performance. They find no indication that screening has a material impact on portfolio risk or return. On average, investors neither experience harm nor benefit from investing in a portfolio of socially responsible securities.

Screening out securities from a portfolio must always reduce the benefits of diversification. Why, then, does omission of up to three groups of companies have a tiny impact on expected future performance? The explanation is that, in Ang's analysis, the marginal loss of diversification from exposure to 36 rather than 39 industries is small, and expected returns in his analysis are related solely to the beta of each industry relative to the market index (i.e. there is no expected premium from sin industries).

In the Humphrey and Tan work, the impact of pure-play sin stocks (and of pure-play responsible companies) is weakened because they are dominated by the overwhelming value of core holdings that defy classification as sinful or saintly. Stierli (2014), in his Credit Suisse research, shows that a stringent application of screening criteria reduces the worldwide investment universe by over 65% and, in this case, the impact could be bigger.

Positive screening

We have seen that companies which achieve a high standard of corporate social responsibility have generated superior stock market performance. There is a school of thought that tilting an investor's portfolio toward responsible companies may be rewarded with a positive contribution to investment performance. This tilting strategy is sometimes referred to as positive screening.

A dilemma is that we do not always have clear evidence on how corporate social responsibility influences stock prices. When a company's behavior improves, the perceived risk of the stock may fall, in which case investors will require a lower rate of return. To deliver a lower return in the future, the stock price must rise. So as the company becomes more "responsible," its share price should go up, and once it is a more desirable company, its equities should subsequently deliver a lower return than less responsible alternative investments. Positive screening, which involves selecting companies with standards of behavior that are established as high, may therefore be associated with lower investment returns.

The widely-cited study by Gompers, Ishii and Metrick (2003) provides a case study of this effect. The authors reported that the governance index they compiled was positively correlated with stock market performance during the course of the 1990s: the better governed the company, the more its shares appreciated. Their results encouraged many institutions to emphasize good governance as a criterion for selecting stocks for investment.

Subsequent research by Bebchuk, Cohen and Wang (2013) found that the superior return from well-governed companies could be ascribed to gradual learning during the 1990s by market participants about the benefits of good governance. The stock prices of better-governed companies drifted up, to the extent that during the following decade (2000–08) the governance premium in US stock returns disappeared.

Borgers, Derwall, Koedijk and Ter Horst (2013) demonstrate that there was an analogous period of learning about the importance of being sensitive to the interests of corporate stakeholders. During 1992–2004, companies with positive strategies from a corporate-responsibility perspective outperformed, but subsequently (2005–09) they failed to generate superior returns.

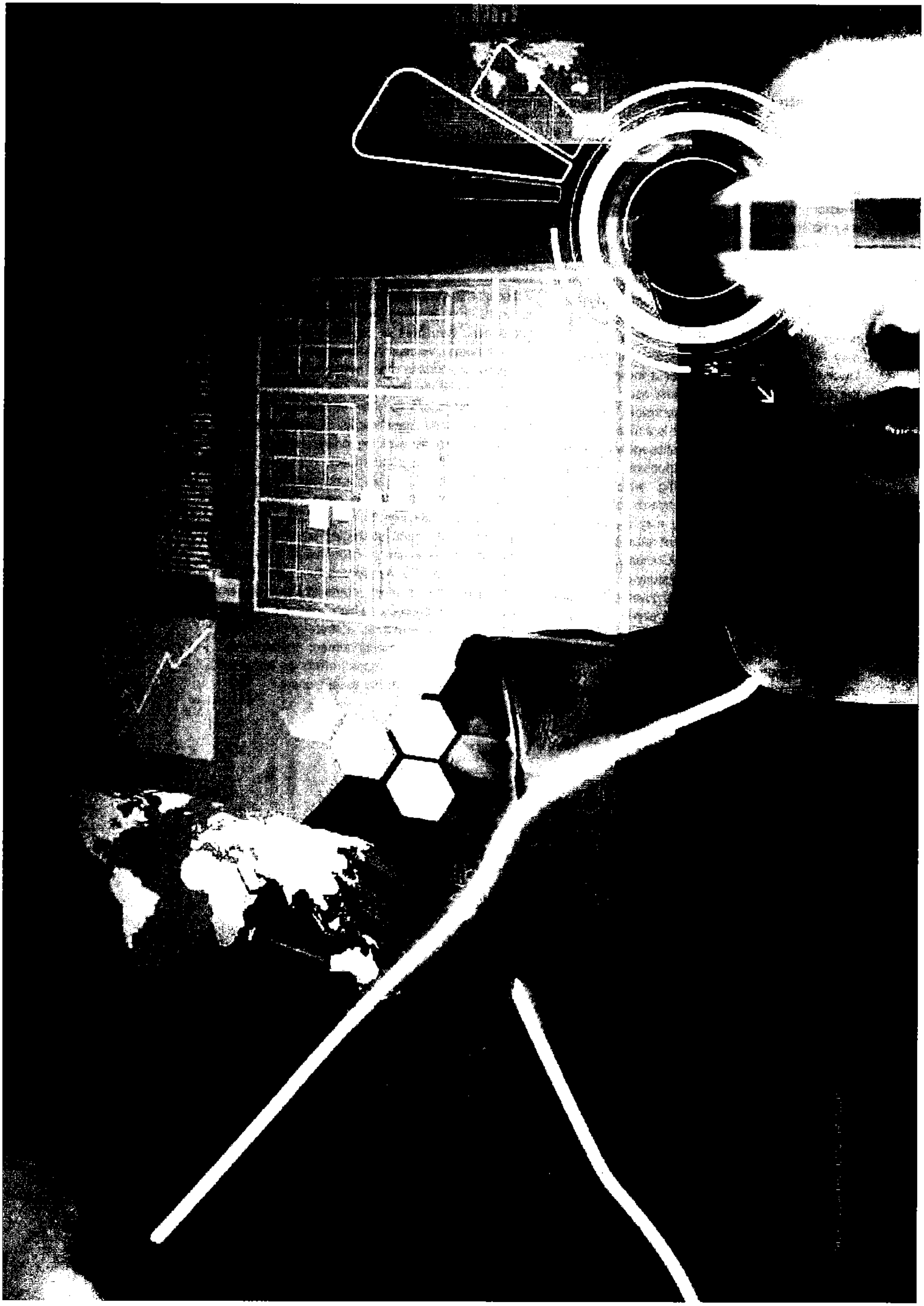
The implication is that buying shares in responsibly managed companies cannot be seen as a strategy that will necessarily be rewarded in the financial markets. Financially, a better course of action may be for investors to engage with the firms whose shares they own or wish to buy. This may facilitate more substantial changes for the better than can be accomplished through positive screening or through exit.

How should that be implemented? To maximize the probability of success as an activist, asset owners might consider the "washing machine" strategy advocated by Gollier and Pouget (2014). They argue that a large investor can generate continuing outperformance by buying non-responsible companies and turning them into more responsible businesses. After they have been cleaned up, the shares may then be sold at a price that reflects the accomplishments of the activist.

Conclusion

Investors have a choice between responding to unacceptable corporate behavior by means of exit or voice. Exit involves excluding the shares of companies, industries or countries with unattractive attributes. Voice involves engaging with the company or pursuing other methods for amending the behavior of the firm. Exit can give rise to performance deviations (positive or negative) relative to unconstrained benchmarks. Large-scale divestment, such as avoiding entire markets, can have a particularly marked impact.

As well as being socially responsible, engaging with investee companies has been shown to be profitable in a number of cases. We find that responsible investment strategies are more likely to pay off when action is coordinated with like-minded activists. To be successful, responsible investing requires a major commitment in asset management resources, which can be costly. While it is not for everyone, a strategy of rotating attention to successive engagement opportunities (the "washing machine" model) offers an interesting direction for responsible asset owners.



Do equity discount rates mean revert?

Mean reversion is a natural phenomenon that provides contrarian investors with a powerful rationale for making and justifying their investment choices. Well-behaved macro signals are highly prized but elusive. Investors often ask us if changes in the cost of capital for equity markets can be predicted, and if the likelihood and magnitude of those changes can be quantified. They would like to know if there is an equilibrium or mean-reverting level for the cost of capital and, if so, how quickly does it revert?

David A. Holland, Bryant Matthews and Pratyasha Rath, Credit Suisse HOLT Valuation & Analytics

"The future ain't what it used to be" – Yogi Berra

HOLT is a division of Credit Suisse that for over 30 years has been offering institutional clients a unique perspective on equity valuation and risk. The HOLT discount rate represents a real, market-implied cost of capital for listed equities, adjusted for regional and sector risk, as well as company size and leverage. Because it links today's equity prices with forward estimates of cash flow, it is a market-implied measure. When investor risk appetite is high (low), the discount rate decreases (increases). Since the bankruptcy of Lehman Brothers in September 2008, when the discount rate skyrocketed, it has been a signal closely followed by many fund managers. A 100 basis-point increase in the discount rate typically amounts to a drop in equity valuation of 20%–25%! Understanding the behavior of the discount rate can be highly beneficial and potentially lucrative.

A recent history of the US discount rate

After discount rates shot up during the 2008 credit crisis, we received numerous calls for our view on the appropriate discount rate to use in valuing corporations. Would discount rates revert back to pre-crisis levels, or would a new, higher level persist? How long would it take to "normalize"? To answer these questions requires an understanding of the discount rate's dynamic behavior.

The weighted-average real discount rate time series for US industrial and service companies since 1976 is plotted in Figure 1. The median is 5.6% with 25th and 75th percentile values of 4.6% and 6.8%, respectively. Today's discount rate is 4.2%, which indicates the US stock market is relatively expensive (nearly at the 10th percentile of historical observations).

Figure 1

Monthly time-series of the weighted-average real discount rate for US industrial and service companies (1976 to present)

Source: Credit Suisse HOLT

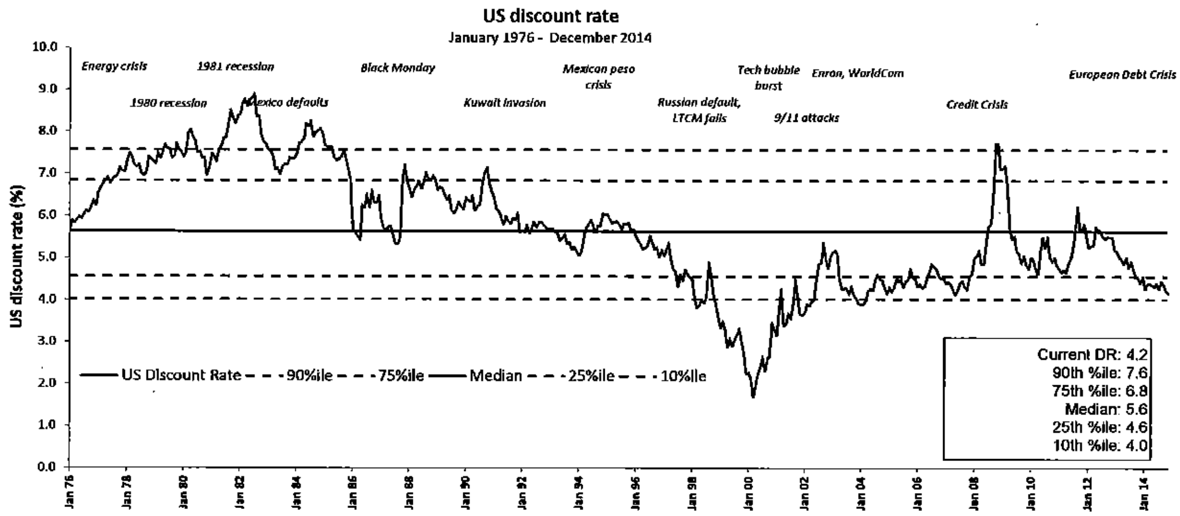


Figure 1 reveals multi-year trends that can be associated with economic developments in the USA, and spikes that can be associated with specific market and macroeconomic events. Spikes indicate bouts of market panic and risk aversion. The large spike in late 2008 highlights the widespread panic of the credit crisis. Conversely, the extraordinary risk appetite preceding the dotcom bubble which peaked in 2000, and accompanying the commodity super cycle during the mid-2000s, is associated with very low discount rates during these periods.

As a general rule of thumb, a discount rate below 5% indicates that investors might be too euphoric and above 7% that investors might be too pessimistic. It is clear from Figure 1 that the market can remain relatively cheap or expensive for many years at a time. As John Maynard Keynes is credited saying, "Markets can remain irrational a lot longer than you and I can remain solvent."[1]

What are the latest discount rates for key equity markets?

The market-implied discount rates for key equity markets are calculated on a weekly basis and used by our clients to obtain a relative sense of value and risk appetite in each market. As a general rule, when the discount rate exceeds or falls below its 75th or 25th percentile, the market has entered either pessimistic or optimistic territory. This can help fund managers decide which markets to gain exposure to, and which to avoid.

Market-implied discount rates as of 10 January 2015 are graphed as triangles from lowest to highest in Figure 2. Indonesia, China and Switzerland have the lowest discount rates (risk on) while Russia, Argentina and Italy have the highest (risk off). The blue vertical bars indicate the interquartile range for each country over the past decade. The black line is the 10-year median. These are useful for relative observations. Ten countries are trading in their bottom quartile (risk on), while only Russia is trading in its top quartile (risk off). Fifteen countries out of 23 are at or below their 10-year medians. While this chart gives us an excellent bird's eye view of regional risk appetite, it does not indicate if and how quickly mean reversion occurs.

General observations about annual changes in the US discount rate

Let us assume the discount rate is mean-reverting. This suggests a rounded discount rate of 6% for US industrial and service companies. Using this rate today in a discounted cash flow model would show that most US stocks are expensive. Due to the highly auto-correlated nature of the discount rate, the best guess for next month's discount rate is not the mean-reverting level, but rather the most recent observation. Fund managers are paid to be in the market, so using the most recent market-implied discount rate is rational, but care should be taken.

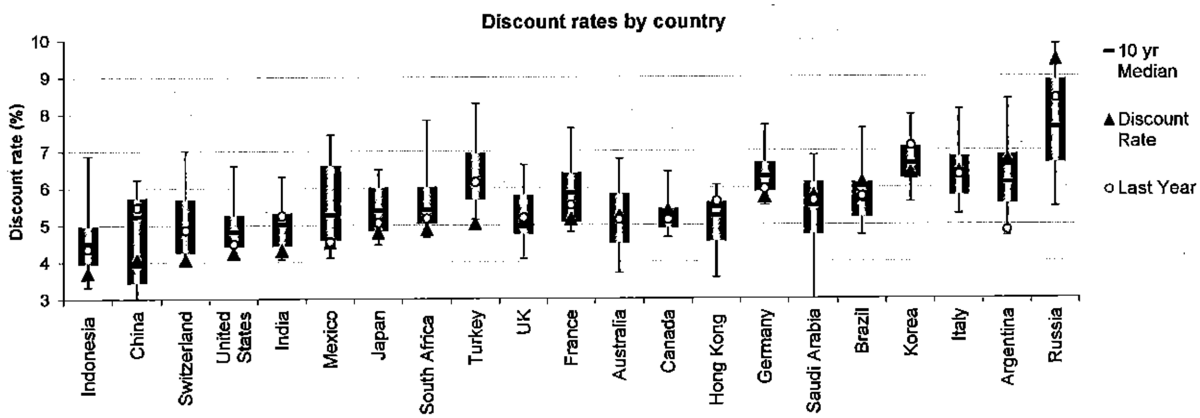
Because of the importance of the discount rate in determining value, it is beneficial to understand whether it is mean-reverting or random walk. If it is mean-reverting, what is the level and rate of mean reversion? (There is no need to pick stocks

[1] The earliest reference to this quote is A. Gary Schilling in Forbes magazine in February 1993, but it is often credited to Keynes.

Figure 2

Market-implied discount rates for industrial and service firms in key equity markets

Source: Credit Suisse HOLT as of 10 January 2015



when you know which way the discount rate is heading.) A great place to start is to understand how the discount rate has changed over 12-month periods. This information can be used to construct probability trees comprised of worst, base and best-case scenarios. Summary statistics since 1976 are shown in Table 1.

Since 1976, the median US discount rate has been 5.6% with dramatic swings. The discount rate was below 4.0% for 10% of the time and over 7.6% for 10% of the time. These outer values could be used to quantify best and worst cases, but that would ignore the fact that the best indicator of next month's discount rate is this month's discount rate.

To take advantage of this property, we calculated the 12-month change in discount rate since 1976. The median change is minus ten basis points. This suggests that in one year the discount rate will essentially be the same as today. The discount rate dropped by 110 basis points or more 10% of the time and increased by 100 basis points or more 10% of the time. A general rule of thumb is to use the 10th, 50th and 90th percentiles for the worst, base and best cases when constructing probability trees. Thus a change of ±100 basis points in the discount rate over the next 12 months is a perfectly sensible assumption for worst (+100 basis points) and best (–100 basis points) cases.

Is the likely change in the discount rate different when the market is euphoric or highly pessimistic? Table 1 shows the 12-month percentage point change when the starting discount rate is less than 5% (euphoric market) and when it is greater than 7% (pessimistic market). Note the asymmetry in changes, particularly for pessimistic markets. Significant drops in discount rate have been more likely for pessimistic markets. This asymmetry suggests that mean reversion might be at work when the discount rate wags its tail.

How does the monthly change in the US discount rate behave?

To better understand the discount rate's behavior, it is helpful to look at the distribution of monthly changes, shown in Figure 3. The median monthly change is a negligible –1 basis point with a 10th percentile change of –26 basis points or less, and a 90th percentile change of 26 basis points or more. The standard deviation is 24.6 basis points which annualizes to 85 basis points.

As is so often the case for financial data, the observations indicate more bunching in the center and fatter tails (leptokurtosis) than that predicted by a normal distribution (red line).

Table 1

US discount rate and 12-month changes in discount rate since 1976

Source: Credit Suisse HOLT data and analysis

Percentile	p(10%)	p(25%)	p(50%)	p(75%)	p(90%)
US discount rate	4.0%	4.6%	5.6%	6.8%	7.6%
12-month change in discount rate (all)	-1.1%	-0.6%	-0.1%	0.5%	1.0%
12-month change in discount rate if < 5%	-0.9%	-0.4%	0.2%	0.7%	1.3%
12-month change in discount rate if > 7%	-1.7%	-1.0%	-0.1%	0.3%	1.0%

In summary, the best guess for next month's discount rate is this month's value with a 10% chance it could drop by 26 basis points or more (best-case scenario for those anticipating an increase in risk appetite) and a 10% chance it could increase by 26 basis points or more (worst-case scenario for those anticipating an increase in risk appetite).

Does the discount rate mean-revert? We tested for this by plotting the discount rate versus its value one month earlier. Random walk behavior is indicated by a slope of one and mean reversion results in a slope less than one. The results are shown in Figure 4.

The slope of 0.986 and correlation coefficient of 0.985 give a weak indication of mean reversion (and a strong display of the auto-correlation). The mean-reverting point for this sample is 5.6%, which means that 98.5% of the spread between today's discount rate and the mean-reverting level is expected to remain in one month's time with a standard deviation of 0.25%. If today's discount rate were 4.20%, then next month's most likely value would be 4.22% plus or minus 0.25%.[2] The volatility and noise of 0.25% swamps any possible mean reversion and signal of 0.02%. If anything is clear, it is that noise dominates signal! This makes the discount rate behavior appear random, and muddies the identification of a clear mean-reverting level. For all intents and purposes, the monthly change in discount rate is a random walk process.

Since HOLT's discount rate is a measure of aggregate risk appetite, and stock price changes are characterized as random walk, it makes sense that HOLT's discount rate approximates a random walk process. The apparent non-existence of mean reversion should not be dismissed by those whose investment horizons are secular, i.e. just because the discount rate seems far too high or low does not mean it will not persist at stretched values. Forward estimates of the discount rate can be generated by incorporating an error term in a simple predictive model. The probability of a given level can be determined, where t is in months, DR(0) is today's discount rate, and LTDR is the long-term discount rate.

$$DR(t) = LTDR + 0.985t \times [DR(0) - LTDR] + \varepsilon(t) \qquad (1)$$

The US discount rate on 10 January 2015 was 4.2%, which places it firmly in the lower quartile of historical observations. The above equation can be used to generate a probability table for the evolution of today's discount rate. The top row in Table 2 indicates the number of months forward and the left-hand column indicates the cumulative probability of achieving a discount rate at or below the value indicated. Note how the median slowly drifts toward an assumed mean-reverting point of 6% as time rolls on. Looking one month ahead, the most likely US discount rate is 4.2% with a 10% chance of being 3.9% or lower, and a 10% chance of being 4.5% or higher. Looking 12 months ahead, the expected discount rate is 4.5% with a 10% probability of being 3.4% or less (which is deep bull territory), and a 10% probability of being 5.6% or higher (which is tantamount to full mean reversion). The market is a noisy system!

Figure 3

Distribution of 1-month changes in the US discount rate since 1976

Source: Credit Suisse HOLT data and analysis

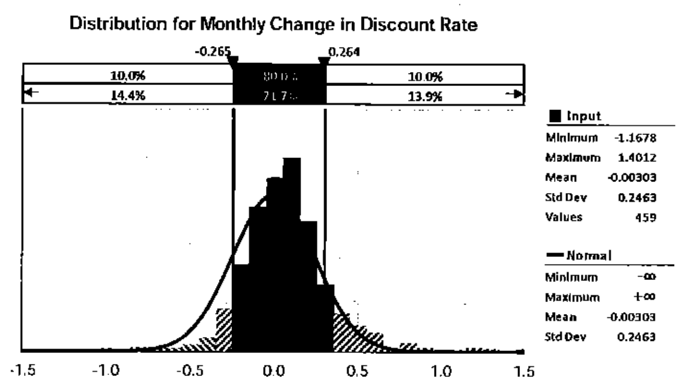
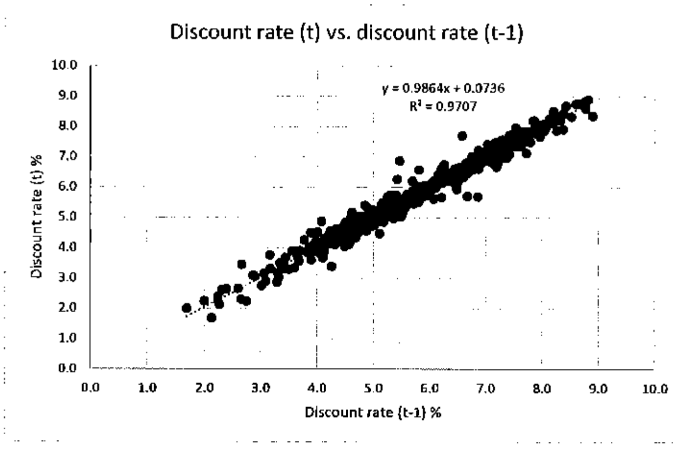


Figure 4

A plot of the US discount rate versus its value in the preceding month since 1976

Source: Credit Suisse HOLT data and analysis

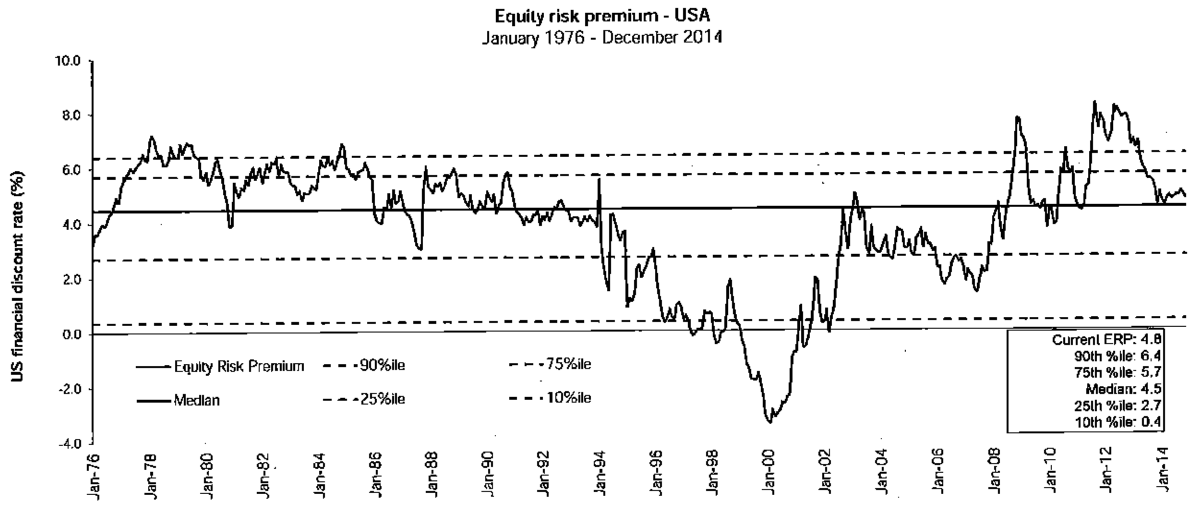


[2]The most-likely value would be 5.6% + 0.985 x (4.20% – 5.6%) = 4.22% for a very small expected increase of 2 basis points in one month. This is insignificant relative to volatility of 25 basis points; the signal-to-noise ratio is 2/25 = 0.08. Imagine trying to track a faint signal in that sea of noise.

Figure 5

Time series of the market-implied ERP for US industrial and service companies

Source: Credit Suisse HOLT



How do changes in the discount rate manifest in the equity risk premium?

The market-implied equity risk premium (ERP) can be estimated from the weighted-average discount rate. The results for US industrial and service firms are charted in Figure 5. Investors are risk averse and demand a premium for the riskiness of equity yields relative to safe "risk-free" yields on government bonds. The median market-implied ERP since 1976 is 4.5%, which is in line with the historical ERP of 4.2% from 1928 to 2012 that Mauboussin and Callahan report, and the 4.5% reported by Dimson, Marsh and Staunton from 1963 to 2012 (all values are relative to US Treasury bonds).

The market-implied ERP is highly volatile, reflecting the vicissitudes of the market's risk appetite for equities. An ERP of 0% suggests risk-neutral behavior and a value less than 0%, which accompanied the dotcom bubble, implies aggressive risk-seeking behavior. The high level of ERP since the credit crisis has been amplified by depressed yields on risk-free treasuries due to quantitative easing and fears of slower growth. Today's ERP of 4.8% is in line with the median. An increase in the risk-free rate due to an ending of quantitative easing would likely reduce the ERP.

The bitter truth about mean reversion

"Predictions are hazardous, especially about the future." – Danish proverb

Expectations of reversion to the mean drive many investments. Well-behaved macro signals are highly prized but elusive. By assuming the stock market is in aggregate fairly priced, Credit Suisse HOLT determines a market-implied discount rate and then uses this to value individual stocks. For investors who have to be invested in equities, or believe the market is approximately right in the aggregate, it is sensible to take a market-neutral approach and use the most current market-implied discount rate when valuing stocks.

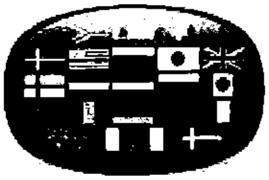
Asset allocators and strategists, however, need to take a view on the attractiveness of markets, and can use market-implied discount rates as a signal. Their job is difficult. Any hints of mean reversion in the market-implied discount rate and ERP are swamped by volatility, suggesting that macro predictions based on imminent mean reversion are precarious at best. For all intents and purposes, monthly changes in the discount rate and ERP are random walk. The market can remain seemingly irrational for long periods, debunking naïve arguments for near-term mean reversion.

Table 2

Probabilistic evolution of the 10 January 2015 US discount rate of 4.2% as a function of months forward

Source: Credit Suisse data and analysis

		Months forward						
	4.2%	1	3	6	12	24	36	60
Cumulative probability	10%	3.9%	3.7%	3.6%	3.4%	3.2%	3.1%	2.8%
	25%	4.1%	4.0%	3.9%	3.9%	3.9%	3.9%	4.0%
	50%	4.2%	4.3%	4.4%	4.5%	4.7%	4.9%	5.3%
	75%	4.4%	4.6%	4.8%	5.1%	5.6%	5.9%	6.5%
	90%	4.5%	4.8%	5.1%	5.6%	6.3%	6.8%	7.7%



All markets

Country profiles

The coverage of the Credit Suisse Global Investment Returns Yearbook comprises 23 countries and three regions, all with index series that start in 1900. Three countries were added in 2013 (Austria, now with a 115-year record, plus Russia and China, which have a gap in their financial market histories from the start of their communist régimes until securities trading recommenced) and one more in 2014 (Portugal, with a 115-year record). There is a 23-country world region, a 22-country world ex-US region, and a 16-country European region. For each region, there are stock and bond indices, measured in USD and weighted by equity market capitalization and GDP, respectively.

Figure 1 shows the relative market capitalizations of world equity markets at our base date of end-1899. Figure 2 shows how they had changed by end-2014. Markets that are not included in the Yearbook dataset are colored black. As these pie charts show, the Yearbook covered 98% of the world equity market in 1900 and 91% at end-2014.

In the country pages that follow, there are three charts for each country or region with an unbroken history. The upper chart reports the cumulative real value of an initial investment in equities, long-term government bonds, and Treasury bills, with income reinvested for the last 115 years. The middle chart reports the annualized real returns on equities, bonds, and bills over this century, the last 50 years, and since 1900. The bottom chart reports the annualized premia achieved by equities relative to bonds and bills, by bonds relative to bills, and by the real exchange rate relative to the US dollar for the latter two periods.

Countries are listed alphabetically, starting on the next page, and followed by three regional groups. Extensive additional information is available in the Credit Suisse Global Investment Returns Sourcebook 2015. This hard-copy reference book of over 220 pages, which is available through London Business School, also contains bibliographic information on the data sources for each country. The underlying annual returns data are redistributed by Morningstar Inc.

The Yearbook's global coverage

The Yearbook contains annual returns on stocks, bonds, bills, inflation, and currencies for 23 countries from 1900 to 2014. The countries comprise two North American nations (Canada and the USA), ten Eurozone states (Austria, Belgium, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal, and Spain), six European markets that are outside the euro area (Denmark, Norway, Russia, Sweden, Switzerland, and the UK), four Asia-Pacific countries (Australia, China, Japan and New Zealand), and one African market (South Africa). These countries covered 98% of the global stock market in 1900 and 91% of its market capitalization by the start of 2015.

Figure 1

Relative sizes of world stock markets, end-1899

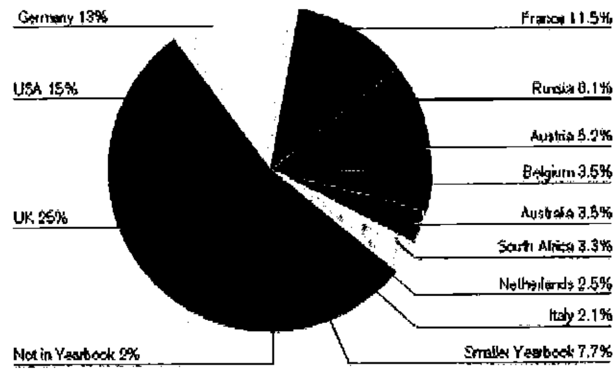
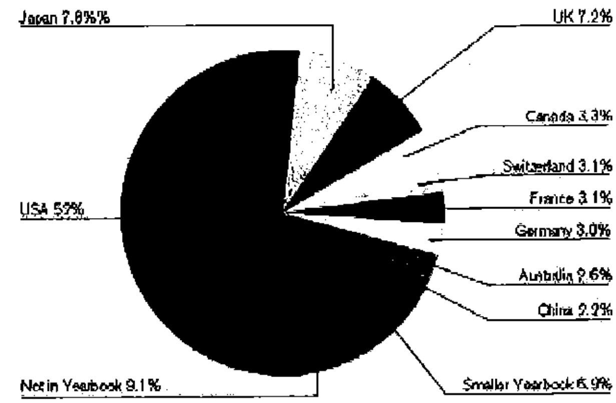


Figure 2

Relative sizes of world stock markets, end-2014



Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Data sources

1. Dimson, E., P. R. Marsh and M. Staunton, 2002, Triumph of the Optimists, NJ: Princeton University Press
2. Dimson, E., P. R. Marsh and M. Staunton, 2007, The worldwide equity premium: a smaller puzzle, R Mehra (Ed.) The Handbook of the Equity Risk Premium, Amsterdam: Elsevier
3. Dimson, E., P. R. Marsh and M. Staunton, 2015, Credit Suisse Global Investment Returns Sourcebook 2015, Zurich: Credit Suisse Research Institute
4. Dimson, E., P. R. Marsh and M. Staunton, 2015, The Dimson-Marsh-Staunton (DMS) Global Investment Returns Database, Morningstar Inc.

Selected data sources for each country are listed in the country profiles below. Detailed attributions, references, and acknowledgements are in the Sourcebook (reference 3).

Australia

The lucky country

Australia is often described as "The Lucky Country" with reference to its natural resources, weather, and distance from problems elsewhere in the world. But maybe Australians make their own luck. The Heritage Foundation ranked Australia as the Yearbook country with the highest economic freedom, while the Charities Aid Foundation study of World Giving ranked Australia as the most generous out of 146 countries in the world.

Whether it is down to economic management, a resource advantage or a generous spirit, Australia has been the second-best performing equity market over the 115 years since 1900, with a real return of 7.3% per year. Regardless of whether it is measured relative to bonds or bills, Australia's long-term equity risk premium has been higher than for any other Yearbook country.

The Australian Securities Exchange (ASX) has its origins in six separate exchanges, established as early as 1861 in Melbourne and 1871 in Sydney, well before the federation of the Australian colonies formed the Commonwealth of Australia in 1901. The ASX ranks among the world's top ten stock exchanges by value and turnover. Half the index is represented by banks (35%) and mining (11%), while the largest stocks at the start of 2015 are BHP Billiton, Commonwealth Bank of Australia, National Australia Bank, Australia & New Zealand Banking Group, and Westpac Banking Corporation.

Australia also has a significant government and corporate bond market, and is home to the largest financial futures and options exchange in the Asia-Pacific region.

Capital market returns for Australia

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 3,441 as compared to 6.8 for bonds and 2.2 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 7.3%, bonds 1.7%, and bills 0.7%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 6.6%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

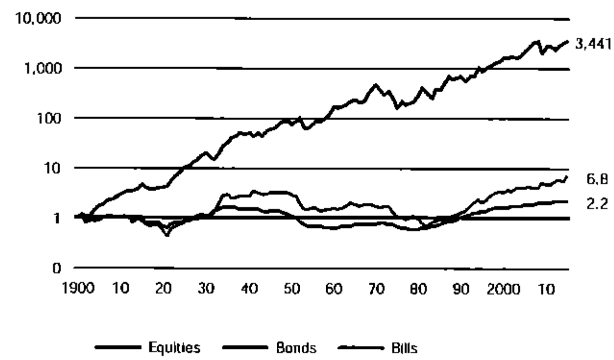


Figure 2

Annualized real returns on major asset classes (%)

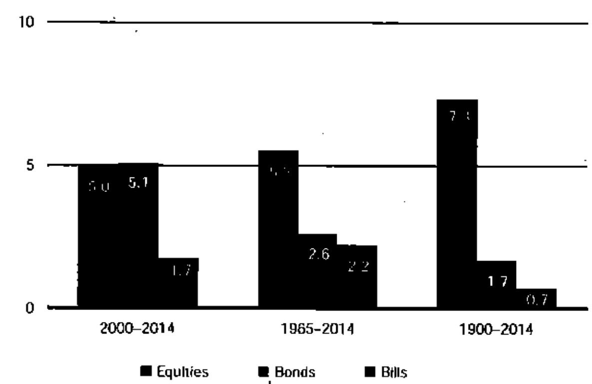
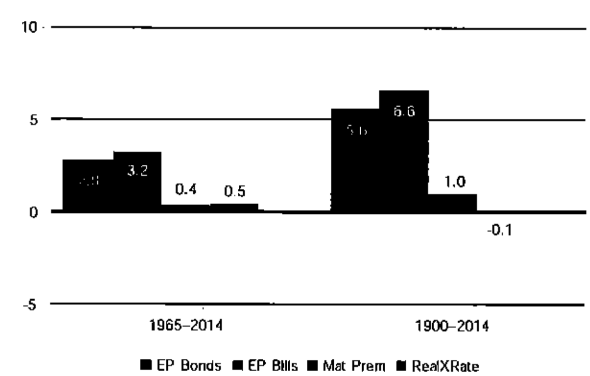


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Austria

Lost empire

The Austrian Empire was re-formed in the 19th century into Austria-Hungary, which, by 1900, was the second-largest country in Europe. It comprised modern-day Austria, Bosnia-Herzegovina, Croatia, Czech Republic, Hungary, Slovakia, Slovenia; large parts of Romania and Serbia; and small parts of Italy, Montenegro, Poland, and Ukraine. At the end of World War I and the break-up of the Habsburg Empire, the first Austrian republic was established.

Although Austria did not pay reparations after World War I, the country suffered hyperinflation during 1921–22 similar to that of Germany. In 1938, Austria was annexed by Germany and ceased to exist as an independent country until after World War II. In 1955, Austria became an independent sovereign state again, and was admitted as a member of the European Union in 1995, and a member of the Eurozone in 1999. Today, Austria is prosperous, enjoying high per capita GDP.

Bonds were traded on the Wiener Börse from 1771 and shares from 1818 onward. Trading was interrupted by the world wars and, after the stock exchange reopened in 1948, share trading was sluggish, and there was not a single IPO in the 1960s or 1970s. From the mid-1980s, building on Austria's gateway to Eastern Europe, the Exchange's activity expanded. Still, over the last 115 years, real stock market returns (0.6% per year) have been lower for Austria than for any other country with a record from 1900 to date.

Financials represents half (47%) of the Austrian equity market. At the start of 2015, the largest Austrian company is Erste Group Bank (25% of the market), followed by OMV, Voestalpine, Andritz, and Immofinanz.

Capital market returns for Austria

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 1.9 as compared to 0.0117 for bonds and 0.0001 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 0.6%, bonds –3.8%, and bills –8.1%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 5.5%. The premia series exclude 1921–22. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

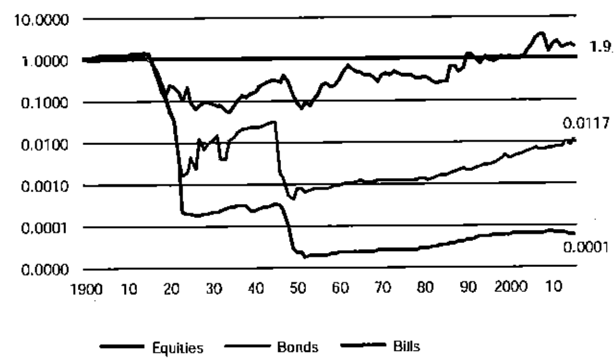


Figure 2

Annualized real returns on major asset classes (%)

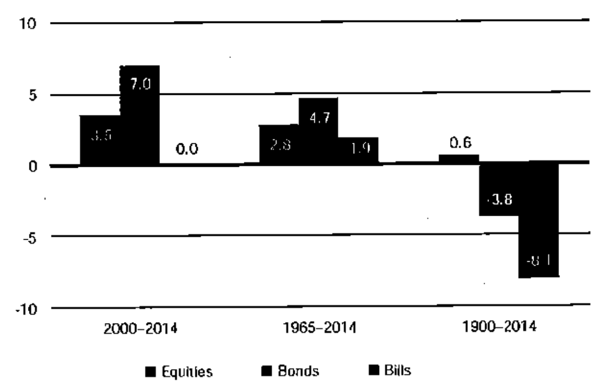
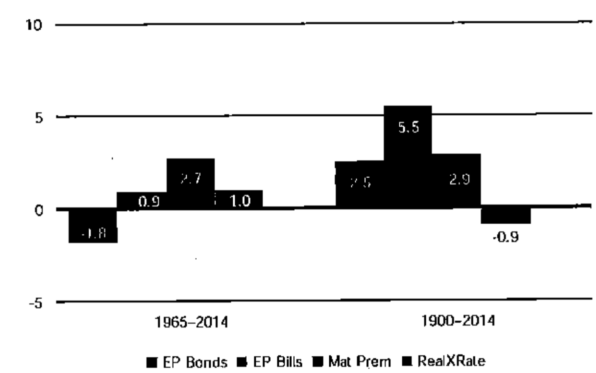


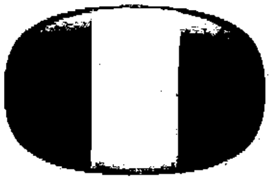
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



Belgium

At the heart of Europe

Belgium lies at the centre of Europe's economic backbone and its key transport and trade corridors, and is the headquarters of the European Union. Belgium has been ranked the most globalized of the 208 nations that are scored in the KOF Index.

Belgium's strategic location has been a mixed blessing, making it a major battleground in international wars, including the Battle of Waterloo, 200 years ago, and the two world wars of the 20th century. The ravages of war and attendant high inflation rates are an important contributory factor to its poor long-run investment returns – Belgium has been one of the three worst-performing equity markets and the seventh worst-performing bond market out of all those with a complete history. Its equity risk premium over 115 years was the worst of the Yearbook countries when measured relative to bills, and fourth-lowest when measured relative to bonds.

The Brussels Stock Exchange was established in 1801 under French Napoleonic rule. Brussels rapidly grew into a major financial center, specializing during the early 20th century in tramways and other urban transport.

Its importance has gradually declined, and what became Euronext Brussels suffered badly during the banking crisis. Three large banks made up a majority of its market capitalization at the start of 2008, but the banking sector now represents less than 10% of its index. By the start of 2015, most of the index (57%) was invested in just one company, Anheuser-Busch InBev, the leading global brewer and one of the world's top five consumer products companies.

The Belgian data draws on work by Annaert, Buelens and Deloof (2012), whom we cite in the Credit Suisse Global Investment Returns Sourcebook 2015.

Capital market returns for Belgium

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 21.6 as compared to 1.6 for bonds and 0.7 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 2.7%, bonds 0.4%, and bills –0.3%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.0%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

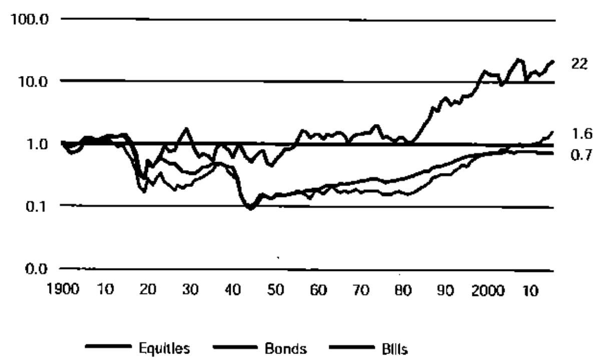


Figure 2

Annualized real returns on major asset classes (%)

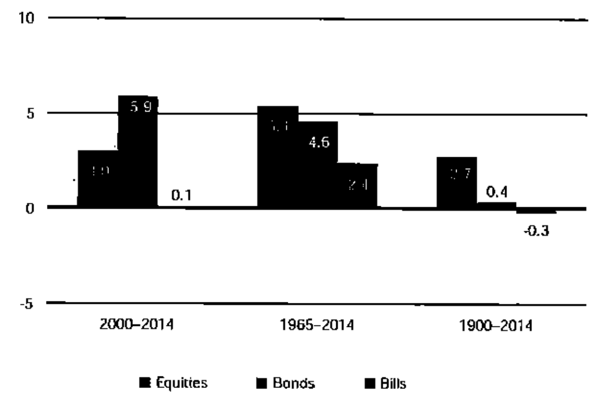
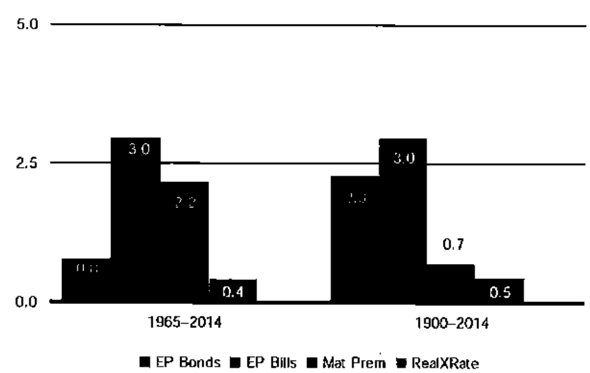


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds: EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Canada

Resourceful country

Canada is the world's second-largest country by land mass (after Russia), and its economy is the tenth-largest. As a brand, it is rated number two out of all the countries monitored in the Country Brand Index. It is blessed with natural resources, having the world's second-largest oil reserves, while its mines are leading producers of nickel, gold, diamonds, uranium and lead. It is also a major exporter of soft commodities, especially grains and wheat, as well as lumber, pulp and paper.

The Canadian equity market dates back to the opening of the Toronto Stock Exchange in 1861 and – as can be seen in the pie chart on the first page of the country profiles section of this report – it is now the world's fourth-largest stock market by capitalization. Canada's bond market also ranks among the world's top ten.

Given Canada's natural endowment, it is no surprise that oil and gas has a 21% weighting, with a further 4% in mining stocks. Banks comprise 29% of the Canadian market. The largest stocks are currently Royal Bank of Canada, Toronto-Dominion Bank, Bank of Nova Scotia, and Suncor Energy.

Canadian equities have performed well over the long run, with a real return of 5.8% per year. The real return on bonds has been 2.2% per year. These figures are close to those for the United States.

Capital market returns for Canada

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 636 as compared to 12.8 for bonds and 5.6 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 5.8%, bonds 2.2%, and bills 1.5%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 4.2%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

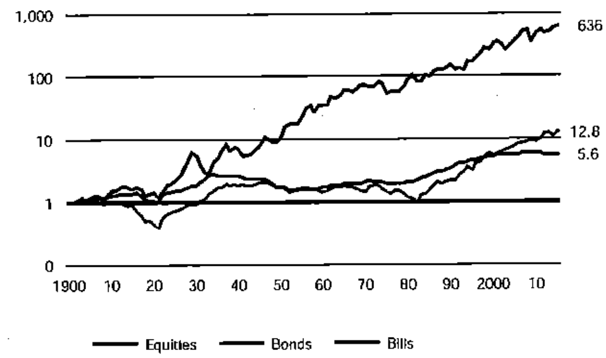


Figure 2

Annualized real returns on major asset classes (%)

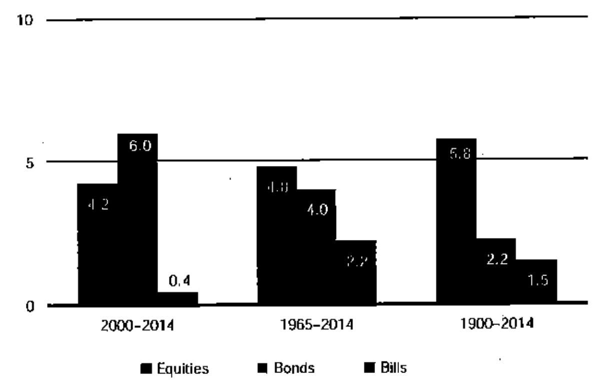
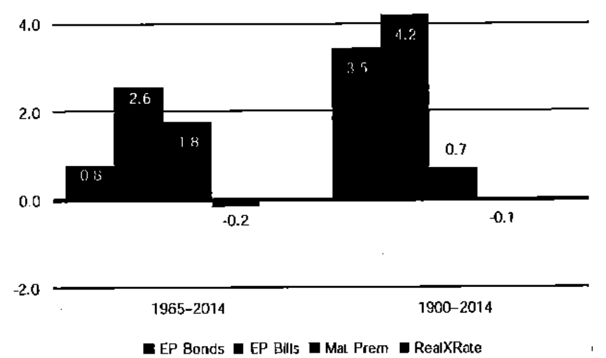


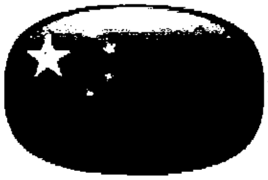
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



China

Emerging powerhouse

The world's most populous country, China has over 1.3 billion inhabitants. After the Qing Dynasty, it became the Republic of China (ROC) in 1911. The ROC nationalists lost control of the mainland at the end of the 1946–49 civil war, after which their jurisdiction was limited to Taiwan and a few islands.

Following the communist victory in 1949, privately owned assets were expropriated and government debt was repudiated, and the People's Republic of China (PRC) has been a single-party state since then. We therefore distinguish among three periods. First, the Qing period and the ROC. Second, the PRC until economic reforms were introduced. Third, the modern period following the second stage of China's economic reforms of the late 1980s and early 1990s.

Though a tiny proportion of assets held outside the mainland may have retained value, and some UK bondholders received a small settlement in 1987 for outstanding claims, we assume the communist takeover generated total losses for domestic investors. After 1940, we hold the nominal value of assets constant until 1949. This gives rise to a collapse in real values during the early 1940s. Chinese returns from 1900 are incorporated into the world and world ex-US indices.

China's economic growth since the reforms has been rapid, and it is now seen as an engine for the global economy. As we discussed in some detail in the 2014 Yearbook, China's fast GDP growth has not been accompanied by superior investment returns. Nearly half (42%) of the Chinese stock market's free-float capitalization is represented by financials, mainly banks and insurers. The largest companies are Tencent Holdings (8% of the index) and China Mobile and China Construction Bank (each 7%), followed by the Industrial and Commercial Bank of China (6%).

Capital market returns for China

In addition to the performance from 1900 to the 1940s, Figure 1 shows that, over 1993-2014, the real value of equities, with income reinvested, declined to 0.5 as compared to a rise to 1.5 for bonds and 1.1 for bills. Figure 2 displays the 1993–2014 real index levels as annualized returns, with equities giving –3.2%, bonds 1.9%, and bills 0.5%. Figure 3 expresses the annualized long-term real returns as premia. Since 1993, the annualized equity risk premium relative to bills has been –3.7%. For more explanations, see page 35.

Figure 1
Cumulative real returns from 1900 to 2014

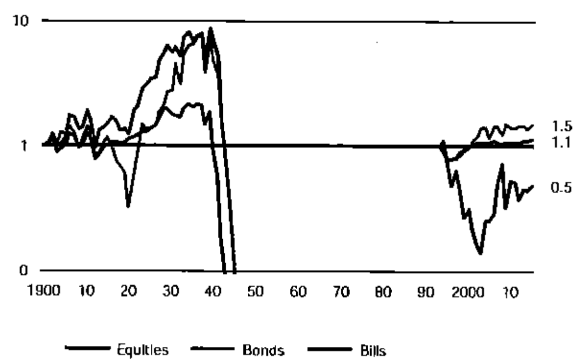


Figure 2
Annualized real returns on major asset classes (%)

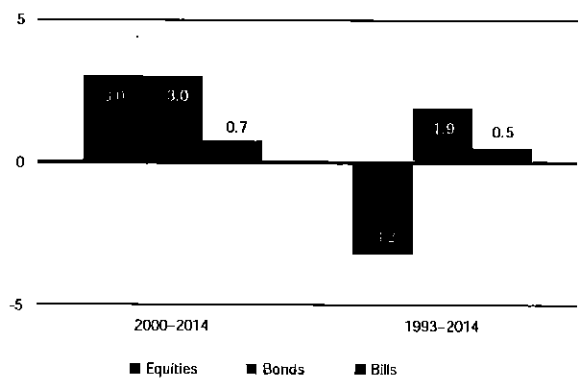
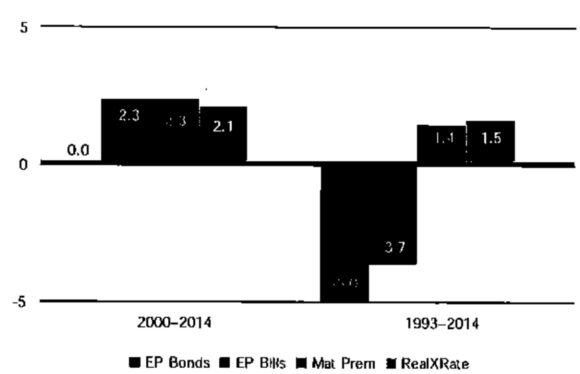


Figure 3
Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Denmark

Happiest nation

The United Nations World Happiness Report, published by Columbia University's Earth Institute, rated Denmark the happiest nation on earth, ahead of Finland, Norway and the Netherlands. The Global Peace Index 2014 rates the country as the most peaceful in the world (jointly with Iceland). And, according to Transparency International, Denmark also ranked joint top with Finland and New Zealand as the least corrupt country in the world.

Whatever the source of Danish happiness and tranquility, it does not appear to spring from outstanding equity returns. Since 1900, Danish equities have given an annualized real return of 5.3%, which is close to the performance of the world equity index.

In contrast, Danish bonds gave an annualized real return of 3.3%, the highest among the Yearbook countries. This is because our Danish bond returns, unlike those for other Yearbook countries, include an element of credit risk. The returns are taken from a study by Claus Parum (see the reference list in the accompanying Credit Suisse Global Investment Returns Sourcebook 2015), who felt it was more appropriate to use mortgage bonds, rather than more thinly traded government bonds.

The Copenhagen Stock Exchange was formally established in 1808, but traces its roots back to the late 17th century. The Danish equity market is relatively small. It has a high weighting in healthcare (54%) and industrials (16%). Nearly one half (41%) of the Danish equity market is represented by one company, Novo-Nordisk. Other large companies include Danske Bank and AP Møller-Mærsk.

Capital market returns for Denmark

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 388.5 as compared to 39.8 for bonds and 11.0 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 5.3%, bonds 3.3%, and bills 2.1%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.1%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

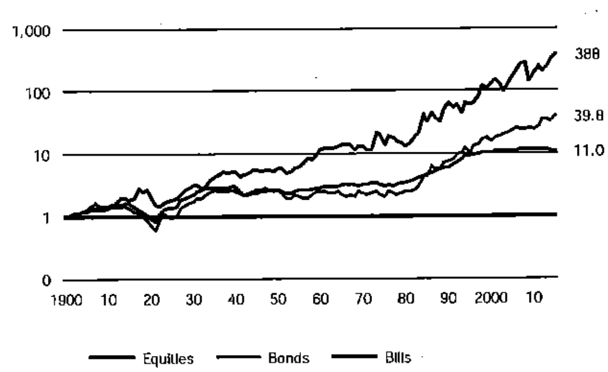


Figure 2

Annualized real returns on major asset classes (%)

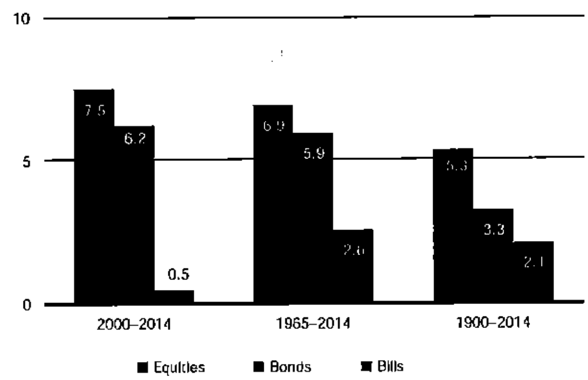
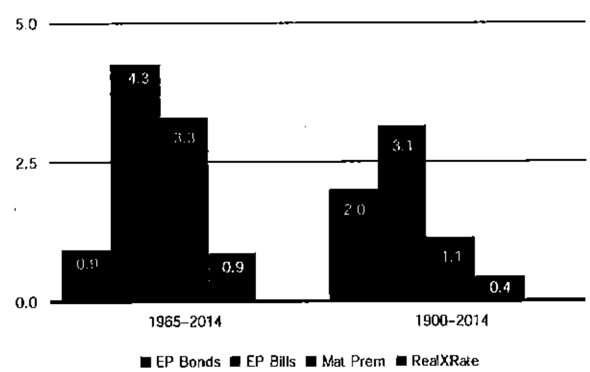


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Finland

East meets West

With its proximity to the Baltics and Russia, Finland is a meeting place for Eastern and Western European cultures. This country of snow, swamps and forests – one of Europe's most sparsely populated nations – was part of the Kingdom of Sweden until sovereignty transferred in 1809 to the Russian Empire. In 1917, Finland became an independent country.

The Fund for Peace ranked Finland as the most stable country, while The Economist Intelligence Unit ranked the Finnish educational system as the world's best. According to Transparency International, Finland is joint top with Denmark and New Zealand as the least corrupt countries. A member of the European Union since 1995, Finland is the only Nordic state in the Eurozone. The country has shifted from a farm and forestry community to a more industrial economy. Per capita income is among the highest in Western Europe.

Finland excels in high-tech exports and is the home country of Nokia. Following Microsoft's acquisition of Nokia's mobile phone business in November 2014, Nokia announced plans to license product designs to third-party manufacturers. Forestry provides a secondary occupation for Finland's rural population.

Finnish securities were initially traded over-the-counter or overseas, and trading began at the Helsinki Stock Exchange in 1912. Since 2003, the Helsinki exchange has been part of the OMX family of Nordic markets. At its peak, Nokia represented 72% of the value-weighted HEX All Shares Index, and Finland was a particularly concentrated stock market. Today, the largest Finnish companies are currently Nokia (26% of the market), Sampo (19% of the market), and Kone (15%).

We have made enhancements to our Finnish equity series, drawing on work by Nyberg and Vaihekoski (2014), whom we acknowledge in the Credit Suisse Global Investment Returns Sourcebook 2015.

Capital market returns for Finland

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 400 as compared to 1.3 for bonds and 0.5 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 5.3%, bonds 0.2%, and bills –0.5%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 5.9%. For additional explanations of these figures, see page 35.

Figure 1
Cumulative real returns from 1900 to 2014

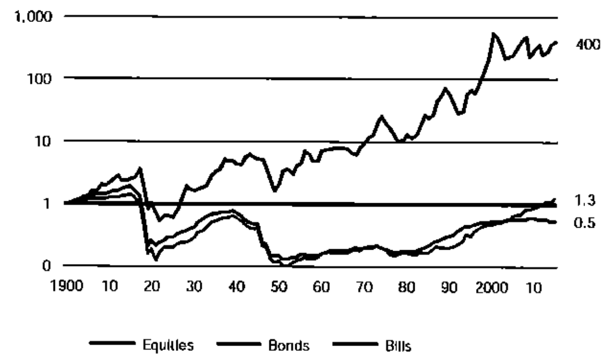


Figure 2
Annualized real returns on major asset classes (%)

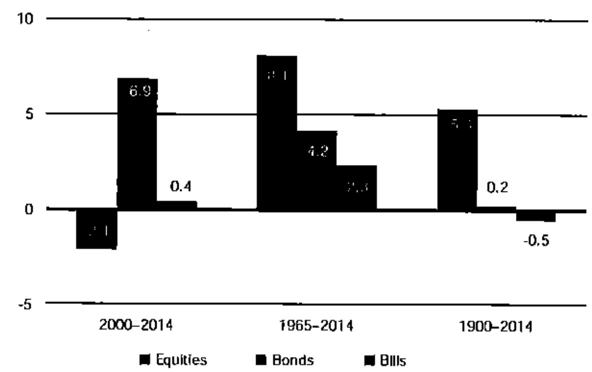
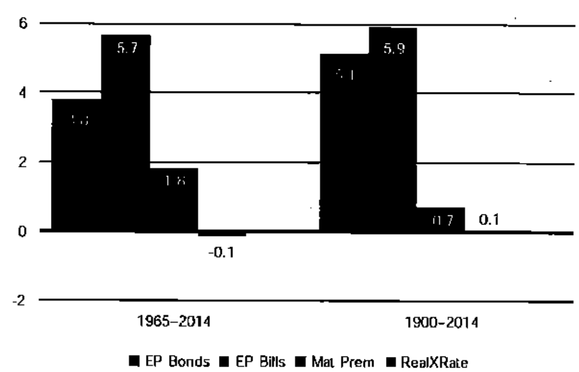
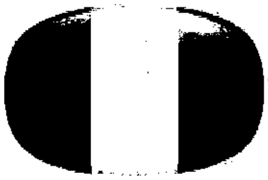


Figure 3
Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



France

European center

Paris and London competed vigorously as financial centers in the 19th century. After the Franco-Prussian War in 1870, London achieved domination. But Paris remained important, especially, to its later disadvantage, in loans to Russia and the Mediterranean region, including the Ottoman Empire. As Kindelberger, the economic historian put it: "London was a world financial center; Paris was a European financial center."

Paris has continued to be an important financial center, while France has remained at the center of Europe, being a founder member of the European Union and the euro. France is Europe's second-largest economy. It has the largest equity market in Continental Europe and one of the largest bond markets in the world. At the start of 2015, France's largest listed companies were Sanofi, Total, and BNP Paribas.

Long-run French asset returns have been disappointing. France ranks in the bottom quartile of countries with a complete history for equity performance, for bonds and for bills, but in the top quartile for inflation – hence the poor fixed income returns. However, the inflationary episodes and poor performance date back to the first half of the 20th century and are linked to the world wars. Since 1950, French equities have achieved mid-ranking returns.

Capital market returns for France

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 36.3 as compared to 1.3 for bonds and 0.0 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 3.2%, bonds 0.2%, and bills –2.8%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 6.1%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

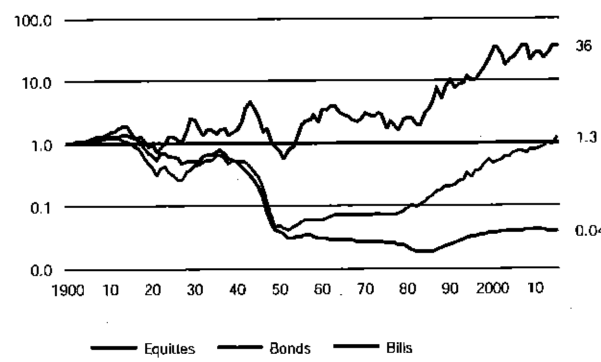


Figure 2

Annualized real returns on major asset classes (%)

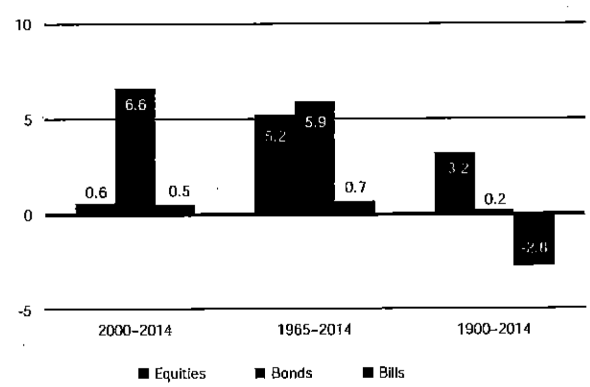
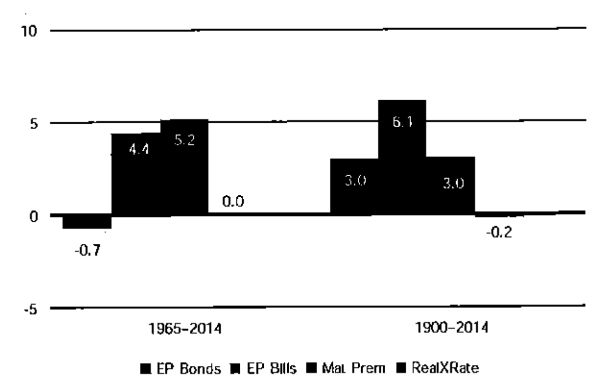


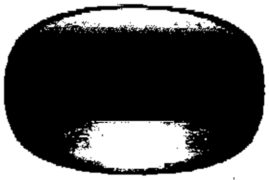
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



Germany

Locomotive of Europe

German capital market history changed radically after World War II. In the first half of the 20th century, German equities lost two thirds of their value in World War I. In the hyperinflation of 1922–23, inflation hit 209 billion percent, and holders of fixed income securities were wiped out. In World War II and its immediate aftermath, equities fell by 88% in real terms, while bonds fell by 91%.

There was then a remarkable transformation. In the early stages of its "economic miracle," German equities rose by 4,373% in real terms from 1949 to 1959. Germany rapidly became known as the "locomotive of Europe." Meanwhile, it built a reputation for fiscal and monetary prudence. From 1949 to date, it has enjoyed the world's second-lowest inflation rate, its strongest currency (now the euro), and an especially strong bond market.

Today, Germany is Europe's largest economy. Formerly the world's top exporter, it has now been overtaken by China. Its stock market, which dates back to 1685, ranks seventh in the world by size, while its bond market is among the world's largest.

The German stock market retains its bias toward manufacturing, with weightings of 23% in basic materials, 22% in consumer goods, and 15% in industrials. The largest stocks are Bayer, Siemens, BASF, Allianz, and SAP.

Our German data incorporates new estimates of historical returns provided to us by Richard Stehle, whose work is cited in the Credit Suisse Global Investment Returns Sourcebook 2015.

Capital market returns for Germany

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 38 as compared to 0.2 for bonds and 0.1 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 3.2% since 1900 and 5.0% since 1965. Figure 3 expresses the annualized long-term real returns as premia. Since 1965, the annualized equity risk premium relative to bills has been 3.3%. Bond, bill and premia series exclude 1922–23. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

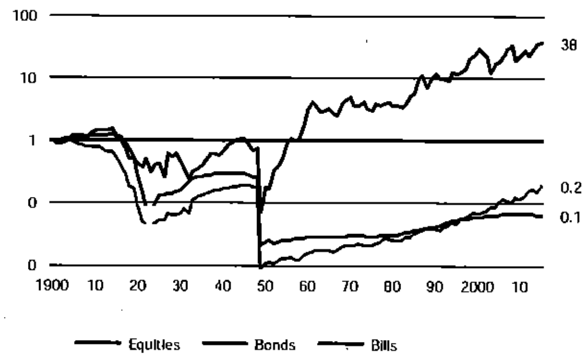


Figure 2

Annualized real returns on major asset classes (%)

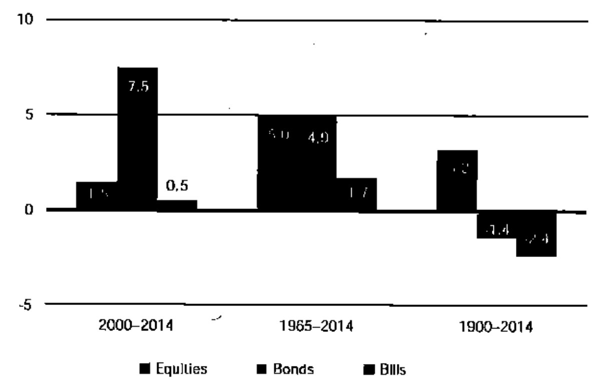
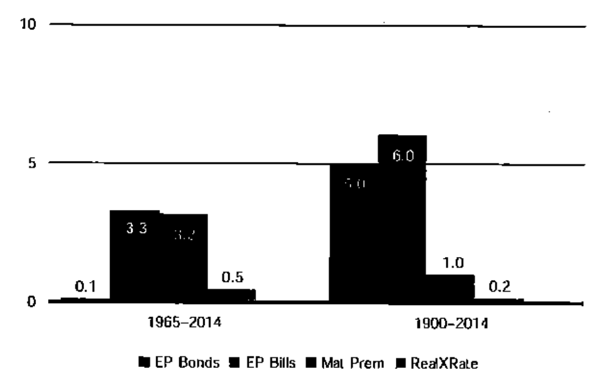


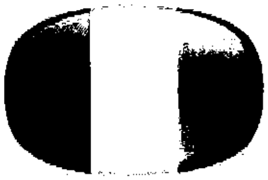
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



Ireland

Born free

Stock exchanges had existed from 1793 in Dublin and Cork, but Ireland was born as an independent country in 1922 as the Irish Free State, released from 700 years of Norman and later British control. In the period following independence, economic growth and stock market performance were weak and, during the 1950s, the country experienced large-scale emigration.

Ireland joined the European Union in 1973 and, from 1987, the economy improved. By the 1990s and early 2000s, Ireland experienced great economic success and became known as the Celtic Tiger. By 2007, Ireland had become the world's fifth-richest country in terms of GDP per capita, the second-richest in the EU, and was experiencing net immigration.

Over the period 1987–2006, Ireland had experienced the second-highest real equity return of any Yearbook country. The financial crisis changed that, and the country still faces hardship. Just as the Born Free Foundation aims to free tigers from being held captive, Ireland now needs to be saved from being a captive of the economic system.

The country is one of the smallest Yearbook markets and, sadly, it has become smaller. Too much of the boom was based on real estate, financials and leverage, and Irish stocks were decimated after 2006. The captive tiger now has a smaller bite.

To monitor Irish stocks from 1900, we constructed an index for Ireland based on stocks traded on the country's two stock exchanges. Ireland adopted the euro from the outset of the Eurozone, and our return series then became euro-denominated.

Capital market returns for Ireland

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 113 as compared to 6.2 for bonds and 2.2 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 4.2%, bonds 1.6%, and bills 0.7%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.5%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

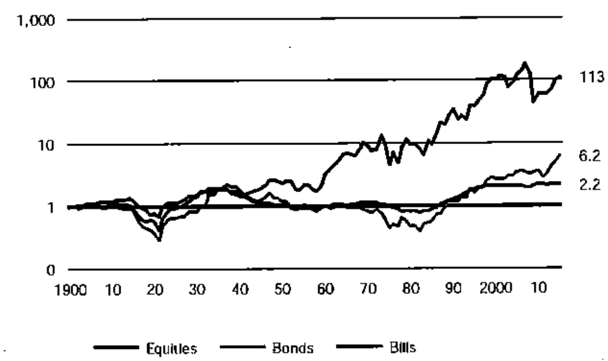


Figure 2

Annualized real returns on major asset classes (%)

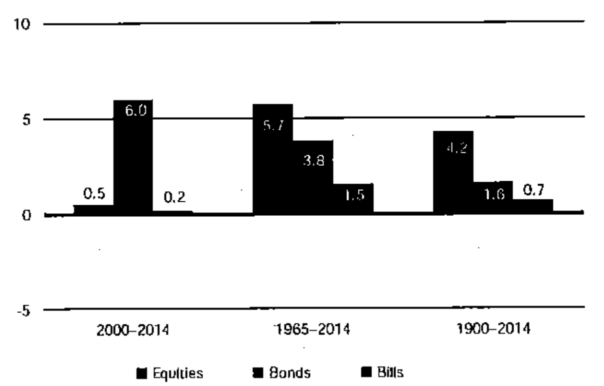
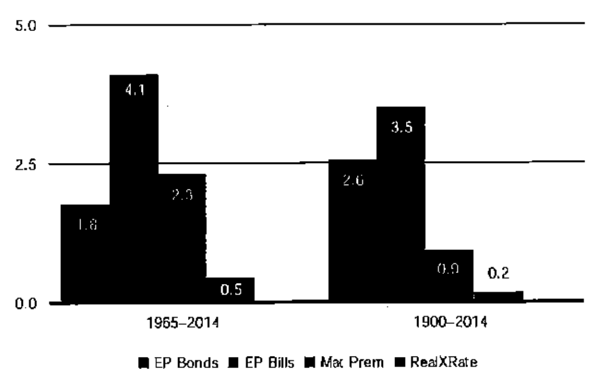


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Italy

Banking innovators

While banking can trace its roots back to Biblical times, Italy can claim a key role in the early development of modern banking. North Italian bankers, including the Medici family, dominated lending and trade financing throughout Europe in the Middle Ages. These bankers were known as Lombards, a name that was then synonymous with Italians.

Italy retains a large banking sector to this day, with banks still accounting for over a quarter (28%) of the Italian equity market, and insurance for a further 10%. Oil and gas accounts for 15%, and the largest stocks traded on the Milan Stock Exchange are Eni, Enel, Intesa Sanpaolo, Unicredit, and Generali.

Italy has experienced some of the poorest asset returns of any Yearbook country. Since 1900, the annualized real return from equities has been 1.9%, which is one of the three lowest returns out of the Yearbook countries. After Germany and Austria, which experienced especially severe hyperinflations, Italy has suffered the poorest real bond and real bill returns of any Yearbook country, the highest inflation rate, and the weakest currency.

Capital market returns for Italy

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 8.8 as compared to 0.3 for bonds and 0.0 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 1.9%, bonds –1.2%, and bills –3.5%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 5.7%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

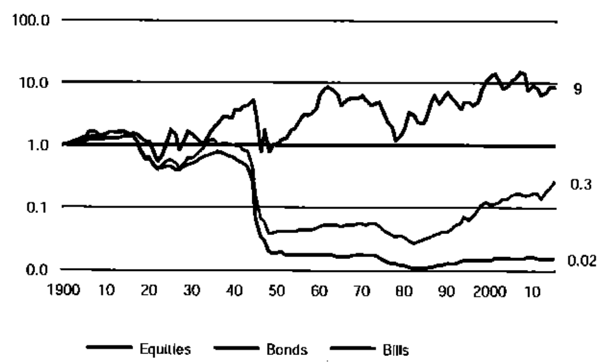


Figure 2

Annualized real returns on major asset classes (%)

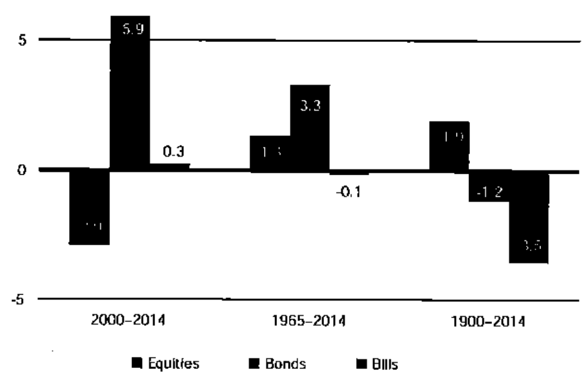
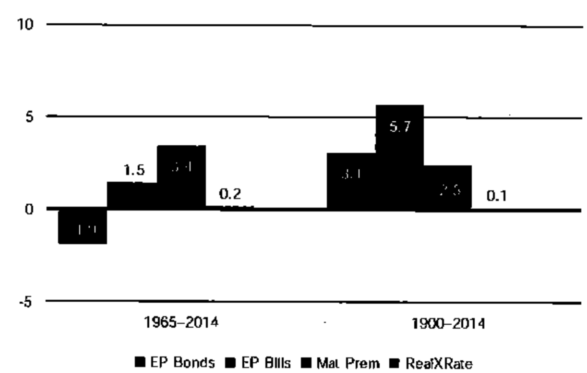


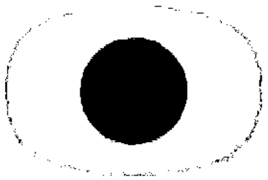
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



Japan

Birthplace of futures

Japan has a long heritage in financial markets. Trading in rice futures had been initiated around 1730 in Osaka, which created its stock exchange in 1878. Osaka was to become the leading derivatives exchange in Japan (and the world's largest futures market in 1990 and 1991), while the Tokyo Stock Exchange, also founded in 1878, was to become the leading market for spot trading.

From 1900 to 1939, Japan was the world's second-best equity performer. But World War II was disastrous and Japanese stocks lost 96% of their real value. From 1949 to 1959, Japan's "economic miracle" began and equities gave a real return of 1,565%. With one or two setbacks, equities kept rising for another 30 years.

By the start of the 1990s, the Japanese equity market was the largest in the world, with a 41% weighting in the world index, as compared to 30% for the USA. Real estate values were also riding high: a 1993 article in the Journal of Economic Perspectives reported that, in late 1991, the land under the Emperor's Palace in Tokyo was worth about the same as all the land in California.

Then the bubble burst. From 1990 to the start of 2009, Japan was the worst-performing stock market. At the start of 2015, its capital value is still close to one third of its value at the beginning of the 1990s. Its weighting in the world index fell from 41% to 8%. Meanwhile, Japan suffered a prolonged period of stagnation, banking crises and deflation. Hopefully, this will not form the blueprint for other countries facing a financial crisis.

Despite the fallout after the asset bubble burst, Japan remains a major economic power. It has the world's second-largest equity market as well as its second-biggest bond market. It is a world leader in technology, automobiles, electronics, machinery and robotics, and this is reflected in the composition of its equity market.

Capital market returns for Japan

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 106 as compared to 0.3 for bonds and 0.1 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 4.1%, bonds –0.9%, and bills –1.9%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 6.1%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

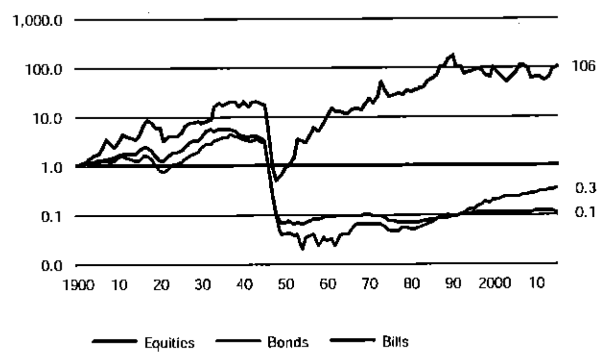


Figure 2

Annualized real returns on major asset classes

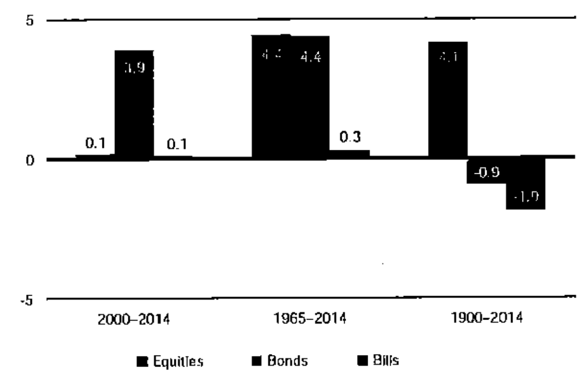
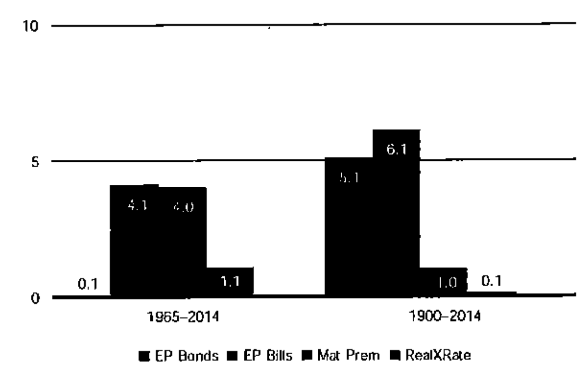


Figure 3

Annualized equity, bond, and currency premia



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Netherlands

Exchange pioneer

Although some forms of stock trading occurred in Roman times and 14th century Toulouse mill companies' securities were traded, transferable securities appeared in the 17th century. The Amsterdam market, which started in 1611, was the world's main center of stock trading in the 17th and 18th centuries.

A book written in 1688 by a Spaniard living in Amsterdam (appropriately entitled Confusion de Confusiones) describes the amazingly diverse tactics used by investors. Even though only one stock was traded – the Dutch East India Company – they had bulls, bears, panics, bubbles and other features of modern exchanges.

The Amsterdam Exchange continues to prosper today as part of Euronext. Over the years, Dutch equities have generated a mid-ranking real return of 5.0% per year. The Netherlands has traditionally been a low inflation country and, since 1900, has enjoyed the lowest inflation rate among the EU countries and the second lowest (after Switzerland) from among all the countries covered in the Yearbook.

The Netherlands has a prosperous open economy. Although Royal Dutch Shell now has its primary listing in London, and a secondary listing in Amsterdam, the Amsterdam exchange still hosts more than its share of major multinationals, including Unilever, Koninklijke Philips, ING Group, Akzo Nobel, Heineken, and ASML Holding.

Capital market returns for the Netherlands

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 268 as compared to 7.1 for bonds and 2.0 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 5.0%, bonds 1.7%, and bills 0.6%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 4.4%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

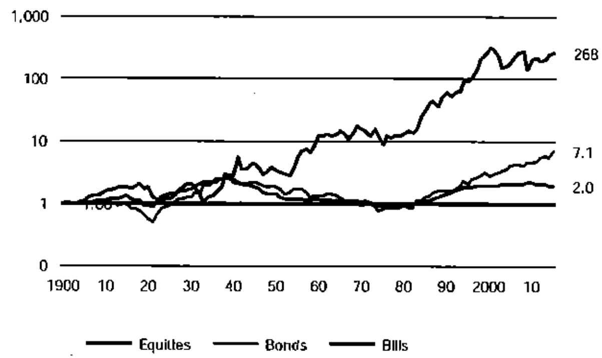


Figure 2

Annualized real returns on major asset classes (%)

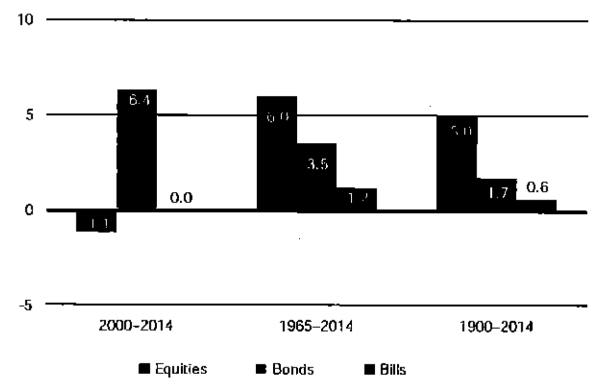
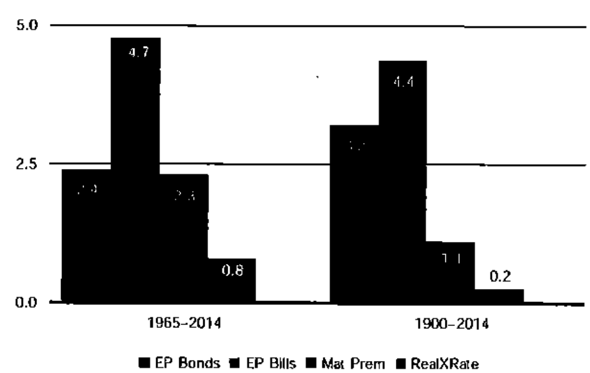


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

New Zealand

Purity and integrity

For a decade, New Zealand has been promoting itself to the world as "100% pure" and Forbes calls this marketing drive one of the world's top ten travel campaigns. But the country also prides itself on honesty, openness, good governance, and freedom to run businesses. According to Transparency International, New Zealand ranked joint top in 2013 with Denmark and Finland as the least corrupt country in the world. The Wall Street Journal ranks New Zealand as the best in the world for business freedom.

The British colony of New Zealand became an independent dominion in 1907. Traditionally, New Zealand's economy was built upon a few primary products, notably wool, meat and dairy products. It was dependent on concessionary access to British markets until UK accession to the European Union.

Over the last two decades, New Zealand has evolved into a more industrialized, free market economy. It competes globally as an export-led nation through efficient ports, airline services, and submarine fiber-optic communications.

The New Zealand Exchange traces its roots to the Gold Rush of the 1870s. In 1974, the regional stock markets merged to form the New Zealand Stock Exchange. In 2003, the Exchange demutualized and officially became the New Zealand Exchange Limited. The largest firms traded on the exchange are Fletcher Building (17% of the index), Spark New Zealand (17%), and Auckland International Airport (11%).

Capital market returns for New Zealand

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 906 as compared to 10.8 for bonds and 6.7 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 6.1%, bonds 2.1%, and bills 1.7%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 4.4%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

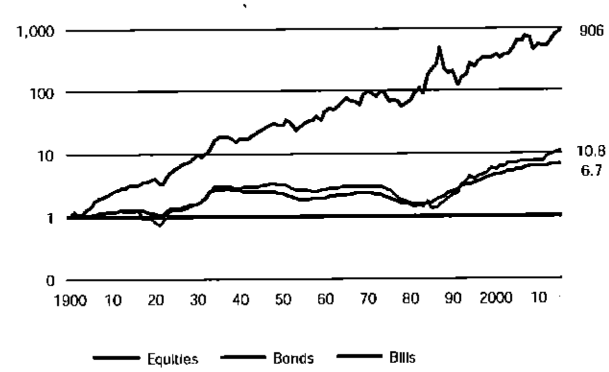


Figure 2

Annualized real returns on major asset classes (%)

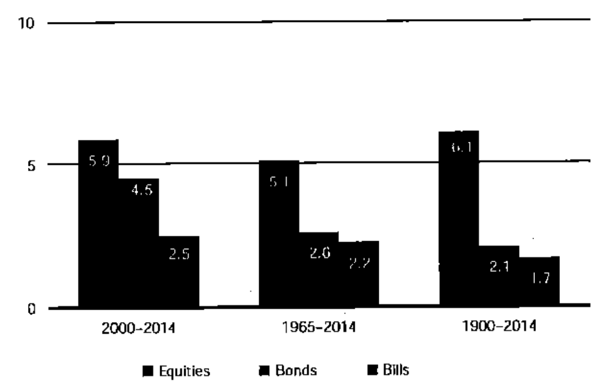
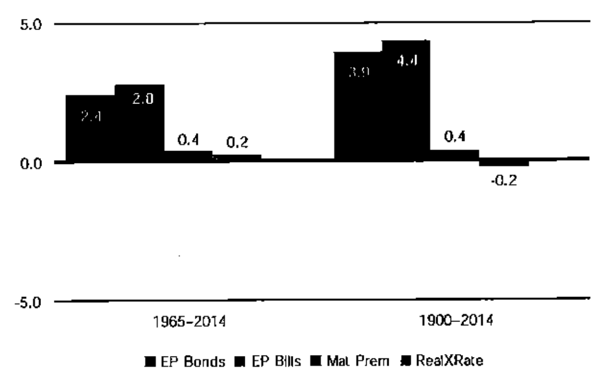


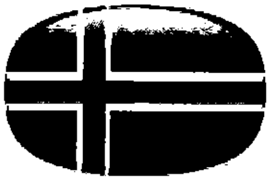
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



Norway

Nordic oil kingdom

Norway is a small country, ranked 115th by population and 61st by land area. However, it is blessed with large natural resources. It is the only country that is self sufficient in electricity production (through hydro power) and it is one of the world's largest exporters of oil. Norway is the second-largest exporter of fish.

The population of 4.9 million enjoys the largest GDP per capita in the world, apart from a few city states. Norwegians live under a constitutional monarchy outside the eurozone. Prices are high: The Economist's Big Mac Index recently reported that a burger in Norway was more expensive than in any other country. The United Nations, through its Human Development Index, ranks Norway the best country in the world for life expectancy, education and overall standard of living.

The Oslo Stock Exchange was founded as Christiania Bors in 1819 for auctioning ships, commodities, and currencies. Later, this extended to trading in stocks and shares. The exchange now forms part of the OMX grouping of Scandinavian exchanges.

In the 1990s, the Government established its petroleum fund to invest the surplus wealth from oil revenues. This has grown to become the largest fund in the world, with a market value approaching USD 0.9 trillion. The fund invests predominantly in equities and, on average, it owns 1.3% of every listed company in the world.

The largest Oslo Stock Exchange stocks are Statoil (19% of the index), DNB (18%), and Telenor (16%).

Capital market returns for Norway

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 117 as compared to 8.4 for bonds and 3.6 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 4.2%, bonds 1.9%, and bills 1.1%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.1%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

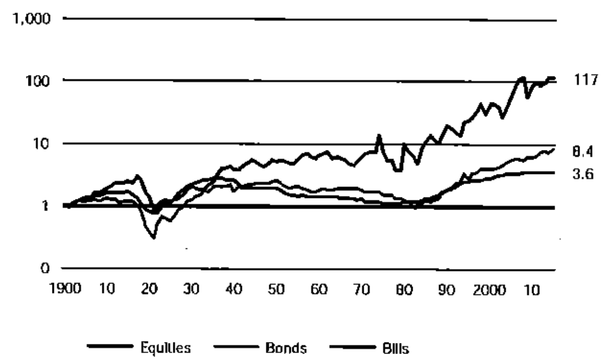


Figure 2

Annualized real returns on major asset classes (%)

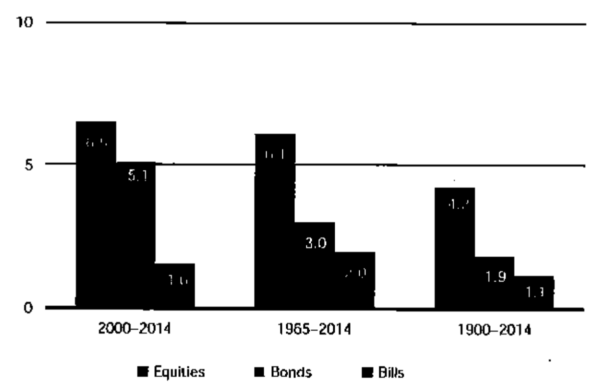
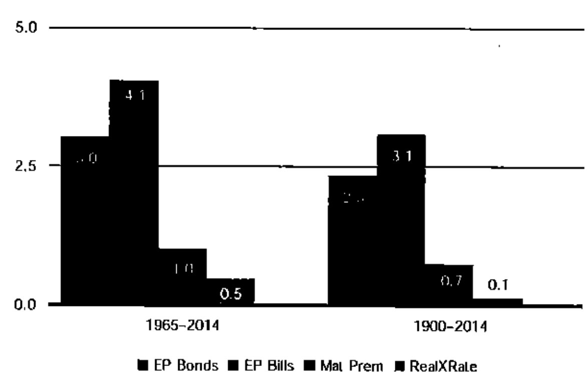


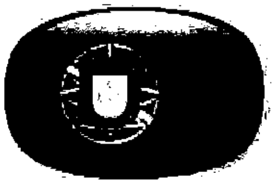
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



Portugal

Land of discoverers

In the 15th century, during The Age of the Discoveries, a rudimentary form of centralized market existed in Lisbon. It solved two problems: how to assemble the large amounts of money necessary to finance the fleets and the voyages; and how to agree the premia for insurance contracts to cover the associated risks. In general, this was not a formally organized market, and transactions were conducted in the open air at a corner of a main street in downtown Lisbon. Nevertheless, that market offered opportunities to trade commodities, in particular those brought by this nation of mariners from recently discovered countries.

Modern Portugal emerged in 1974 from the Carnation Revolution, a bloodless military coup which overthrew the former regime. The country joined the European Union in 1986 and was among the first to adopt the euro. In the second decade of the 21st century the Portuguese economy suffered its most severe recession since the 1970s, and unemployment still remains high.

The companies with the largest market capitalizations are in the utility and energy groups – comprising 53% in utilities and 18% in oil and gas. The largest companies traded in Lisbon are EDP, Galp Energia, BC Portugues, and Jeronimo Martins.

The data for Portuguese equities come from a recently completed study by da Costa and Mata (2014), whose research is cited in full in the Credit Suisse Global Investment Returns Sourcebook 2015.

Capital market returns for Portugal

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 48.1 as compared to 2.5 for bonds and 0.4 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 3.4%, bonds 0.8%, and bills –0.9%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 4.3%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

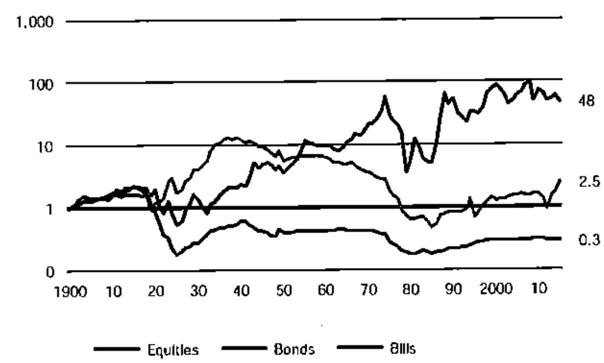


Figure 2

Annualized real returns on major asset classes (%)

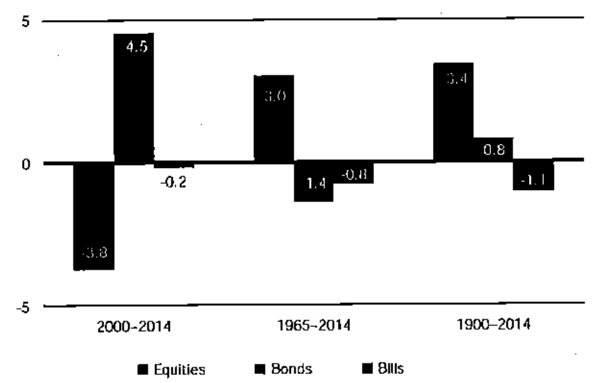
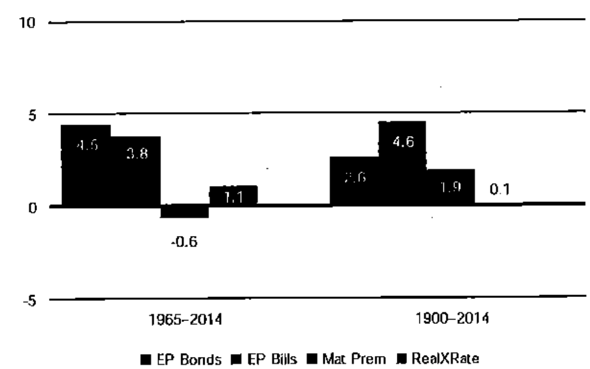


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Russia

Wealth of resources

Russia is the world's largest country, covering more than one-eighth of the Earth's inhabited land area, spanning nine time zones, and located in both Europe and Asia. Formerly, it even owned one-sixth of what is now the USA. It is the world's leading oil producer, second-largest natural gas producer, and third-largest steel and aluminium exporter. It has the biggest natural gas and forestry reserves and the second-biggest coal reserves.

After the 1917 revolution, Russia ceased to be a market economy. We therefore distinguish among three periods. First, the Russian Empire up to 1917. Second, the long interlude following Soviet expropriation of private assets and the repudiation of Russia's government debt. Third, the Russian Federation, following the dissolution of the Soviet Union in 1991.

Very limited compensation was eventually paid to British and French bondholders in the 1980s and 1990s, but investors in aggregate still lost more than 99% in present value terms. The 1917 revolution is deemed to have resulted in complete losses for domestic stock- and bondholders. Russian returns are incorporated into the world, world ex-US, and Europe indices.

In 1998, Russia experienced a severe financial crisis, with government debt default, currency devaluation, hyperinflation, and an economic meltdown. However, there was a surpisngly swift recovery and, in the decade after the 1998 crisis, the economy averaged 7% annual growth. In 2008–09, there was a major reaction to global setbacks and commodity price swings. Fuelled by a persistently volatile political situation, Russian stock market performance has likewise been volatile.

By the beginning of 2015, over half (56%) of the Russian stock market comprised oil and gas companies, the largest being Gazprom and Lukoil. Adding in basic materials, resources are over two-thirds of market capitalization.

Capital market returns for Russia

In addition to the performance from 1900 to 1917, Figure 1 shows that, over 1995-2014, the real value of equities, with income reinvested, grew by a factor of 2.0 as compared to 1.9 for bonds and a decline to 0.6 for bills. Figure 2 displays the 1995-2014 real index levels as annualized returns, with equities giving 3.5%, bonds 3.2%, and bills –2.2%. Figure 3 expresses the annualized long-term real returns as premia. Since 1995, the annualized equity risk premium relative to bills has been 5.8%. For more explanations, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

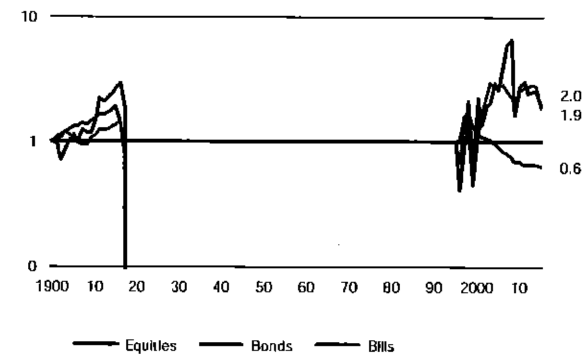


Figure 2

Annualized real returns on major asset classes (%)

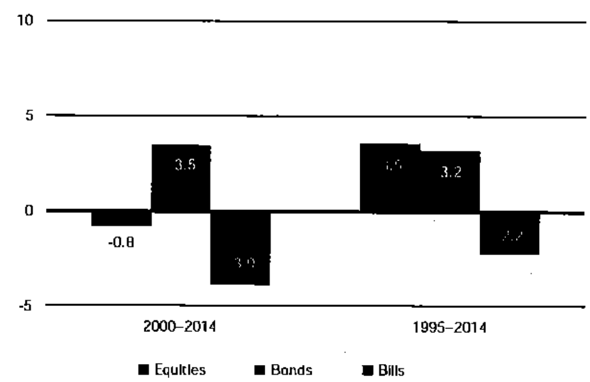
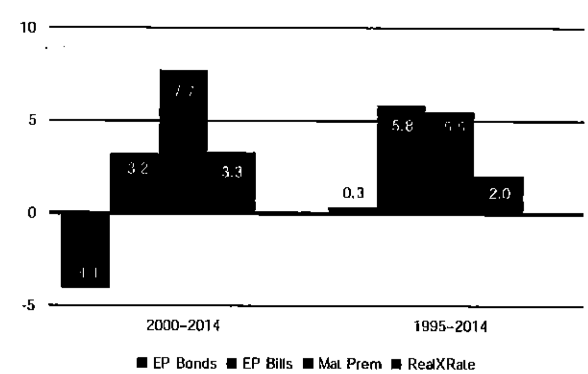


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

South Africa

Golden opportunity

The discovery of diamonds at Kimberley in 1870 and the Witwatersrand gold rush of 1886 had a profound impact on South Africa's subsequent history. Today, South Africa has 90% of the world's platinum, 80% of its manganese, 75% of its chrome and 41% of its gold, as well as vital deposits of diamonds, vanadium, and coal.

The 1886 gold rush led to many mining and financing companies opening up. To cater to their needs, the Johannesburg Stock Exchange (JSE) opened in 1887. Over the years since 1900, the South African equity market has been one of the world's most successful, generating a real equity return of 7.4% per year, which is the highest return among the Yearbook countries.

Today, South Africa is the largest economy in Africa, with a sophisticated financial structure. Back in 1900, South Africa, together with several other Yearbook countries, would have been deemed an emerging market. According to index compilers, it has not yet emerged and today ranks as the fifth-largest emerging market.

Gold, once the keystone of South Africa's economy, has declined in importance as the economy has diversified. Financials account for 24%, while basic minerals lag behind with only 8% of the market capitalization. Taken together, media and mobile telecommunications account for 26% of the market index. The largest JSE stocks are Naspers, MTN, and Sasol.

Capital market returns for South Africa

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 3,551 as compared to 8.6 for bonds and 3.0 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 7.4%, bonds 1.9%, and bills 1.0%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 6.3%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

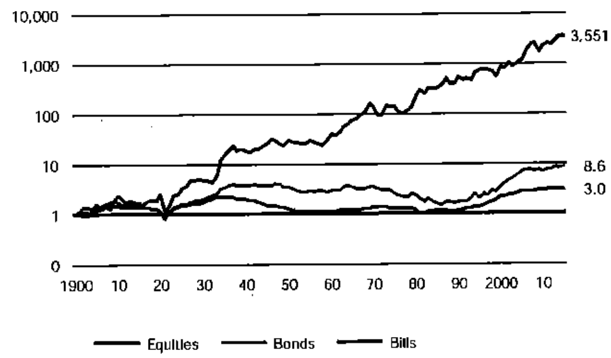


Figure 2

Annualized real returns on major asset classes (%)

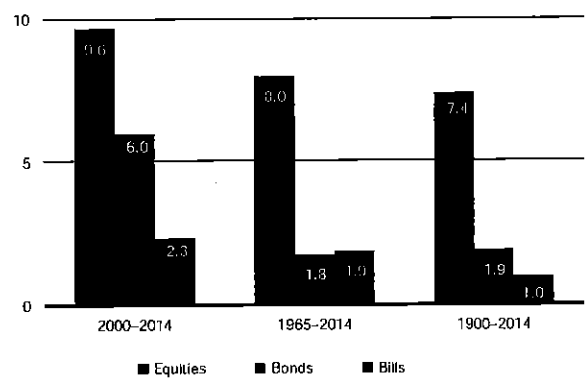
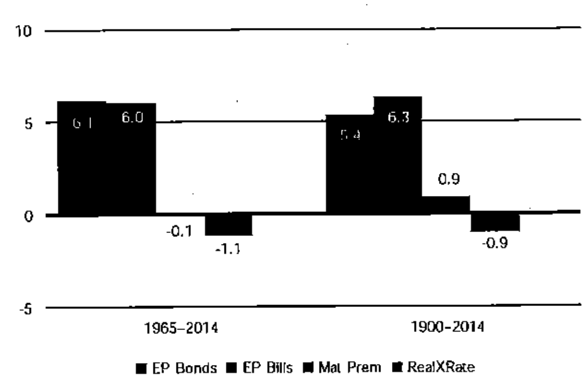


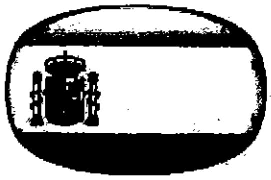
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



Spain

Key to Latin America

Spanish is the most widely spoken international language after English, and has the fourth-largest number of native speakers after Chinese, Hindi and English. Partly for this reason, Spain has a visibility and influence that extends far beyond its Southern European borders, and carries weight throughout Latin America.

While the 1960s and 1980s saw Spanish real equity returns enjoying a bull market and ranked second in the world, the 1930s and 1970s witnessed the very worst returns among our countries. Over the entire 115 years covered by the Yearbook, Spain's long-term equity premium (measured relative to bonds) was 1.9%, which is lower than for any other country that we cover over the same period.

Though Spain stayed on the sidelines during the two world wars, Spanish stocks lost much of their real value over the period of the civil war during 1936–39, while the return to democracy in the 1970s coincided with the quadrupling of oil prices, heightened by Spain's dependence on imports for 70% of its energy needs.

The Madrid Stock Exchange was founded in 1831 and is now the fourteenth-largest in the world, helped by strong economic growth since the 1980s. The major Spanish companies retain strong presences in Latin America combined with increasing strength in banking and infrastructure across Europe. The largest stocks are Banco Santander (24% of the index), Telefonica, BBVA, and Inditex.

Capital market returns for Spain

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 63.2 as compared to 7.7 for bonds and 1.4 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 3.7%, bonds 1.8%, and bills 0.3%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.4%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

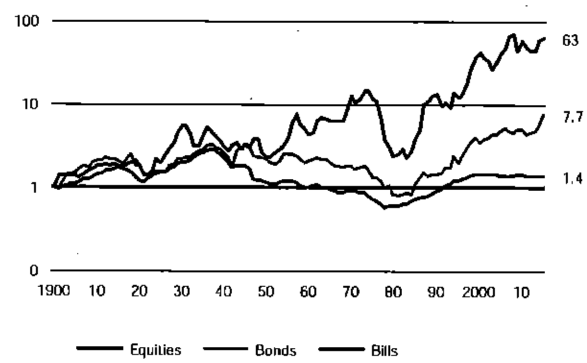


Figure 2

Annualized real returns on major asset classes (%)

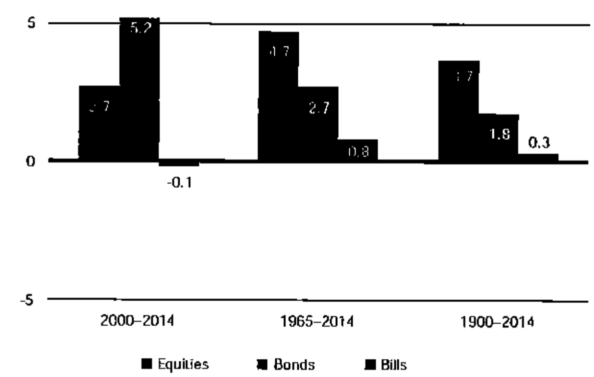
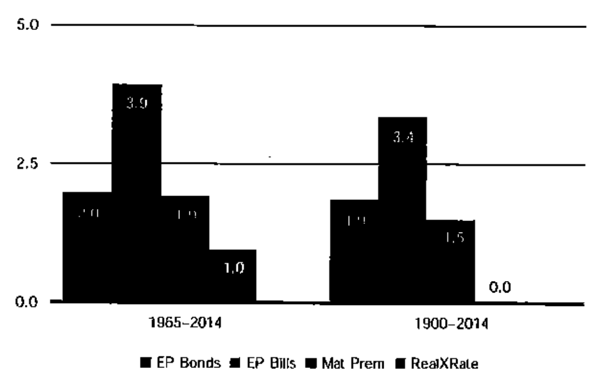


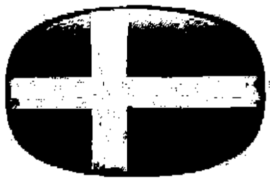
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



Sweden

Nobel prize returns

Alfred Nobel bequeathed 94% of his wealth to establish and endow the five Nobel Prizes (first awarded in 1901), instructing that the capital be invested in safe securities. Were a Nobel prize to be awarded for investment returns, it would be given to Sweden for its achievement as the only country to have real returns for equities, bonds and bills all ranked in the top six.

Real Swedish equity returns have been supported by a policy of neutrality through two world wars, and the benefits of resource wealth and the development of industrial holding companies in the 1980s. Overall, they have returned 5.8% per year. Details on our Swedish index data and sources are provided in the Credit Suisse Global Investment Returns Sourcebook 2015.

The Stockholm Stock Exchange was founded in 1863 and is the primary securities exchange of the Nordic countries. Since 1998, it has been part of the OMX grouping.

In Sweden, the financial sector accounts for a third (35%) of equity market capitalization. The largest single company is Hennes and Mauritz, followed by Nordea Bank and Ericsson.

In 2014, we made enhancements to our series for Swedish equities, drawing on work by Daniel Waldenström (2014), whom we acknowledge in the Credit Suisse Global Investment Returns Sourcebook 2015.

Capital market returns for Sweden

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 684 as compared to 22.9 for bonds and 8.5 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 5.8%, bonds 2.8%, and bills 1.9%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.9%. For additional explanations of these figures, see page 35.

Figure 1
Cumulative real returns from 1900 to 2014

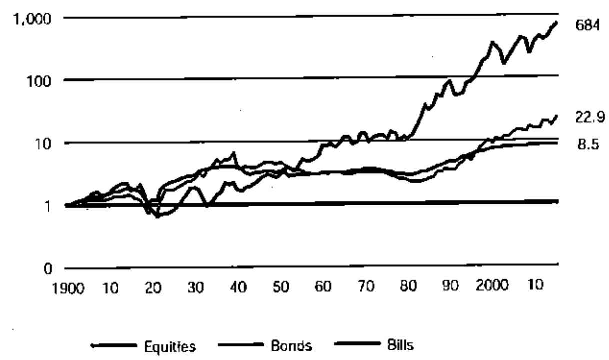


Figure 2
Annualized real returns on major asset classes (%)

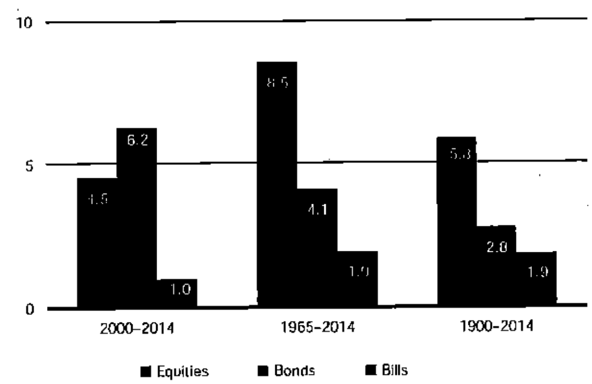
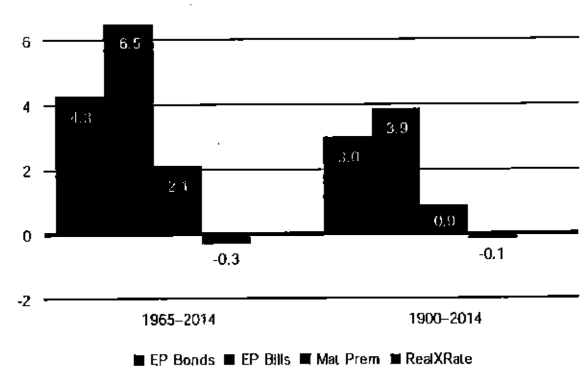


Figure 3
Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Switzerland

Traditional "safe haven"

For a small country with just 0.1% of the world's population and less than 0.01% of its land mass, Switzerland punches well above its weight financially and wins several gold medals in the global financial stakes. In the Global Competitiveness Report 2012–2013, Switzerland is top ranked in the world. It is also ranked by Future Brand Index as the world's number one country brand.

The Swiss stock market traces its origins to exchanges in Geneva (1850), Zurich (1873), and Basel (1876). It is now the world's fifth-largest equity market, accounting for 3.1% of total world value.

Since 1900, Swiss equities have achieved an acceptable real return of 4.5%, while Switzerland has been one of the world's four best-performing government bond markets, with an annualized real return of 2.3%. Switzerland has also enjoyed the world's lowest inflation rate: just 2.2% per year since 1900. Meanwhile, the Swiss franc has been the world's strongest currency.

Switzerland is, of course, one of the world's most important banking centers, and private banking has been a major Swiss competence for over 300 years. Swiss neutrality, sound economic policy, low inflation and a strong currency have all bolstered the country's reputation as a safe haven. Today, close to 30% of all cross-border private assets invested worldwide are managed in Switzerland.

Switzerland's pharmaceutical sector accounts for a third (36%) of the equity market. Listed companies include world leaders such as pharma companies Novartis and Roche, plus Nestle – a trio that together comprise more than half of the equity market capitalization of Switerland.

Capital market returns for Switzerland

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 155 as compared to 14.1 for bonds and 2.5 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 4.5%, bonds 2.3%, and bills 0.8%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.7%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

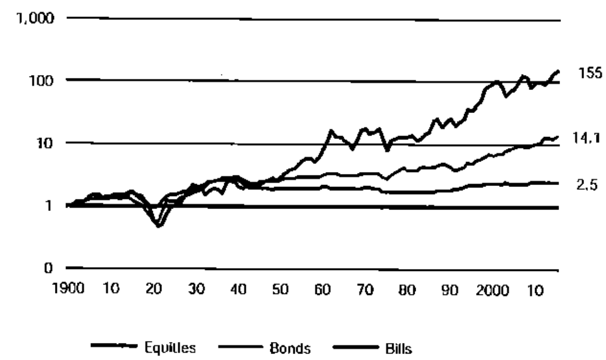


Figure 2

Annualized real returns on major asset classes (%)

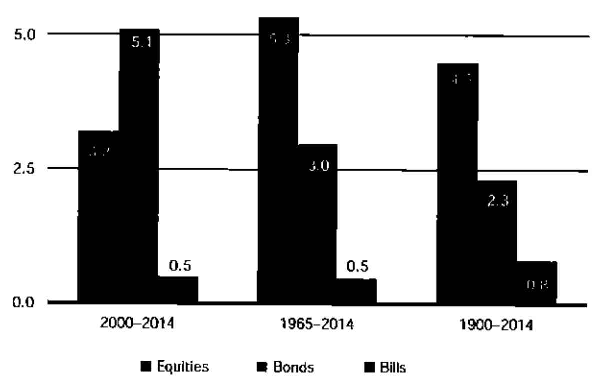
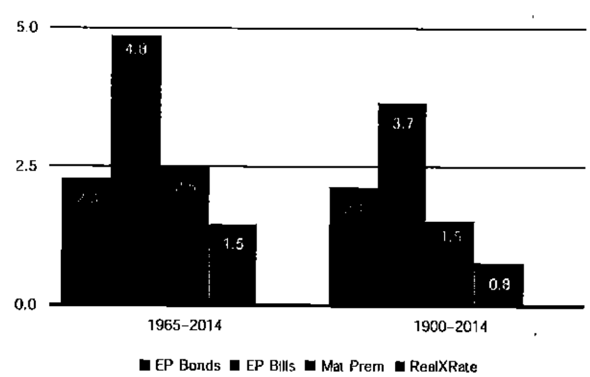


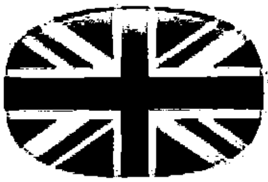
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



United Kingdom

Global center for finance

Organized stock trading in the United Kingdom dates from 1698, and the London Stock Exchange was formally established in 1801. By 1900, the UK equity market was the largest in the world, and London was the world's leading financial center, specializing in global and cross-border finance.

Early in the 20th century, the US equity market overtook the UK and, nowadays, New York is a larger financial center than London. What continues to set London apart, and justifies its claim to be the world's leading international financial center, is the global, cross-border nature of much of its business.

Today, London is ranked as the top financial center in the Global Financial Centres Index, Worldwide Centres of Commerce Index, and Forbes' ranking of powerful cities. It is the world's banking center, with 550 international banks and 170 global securities firms having offices in London. The UK's foreign exchange market is the largest in the world, and Britain has the world's third-largest stock market, third-largest insurance market, and seventh-largest bond market.

London is the world's largest fund management center, managing almost half of Europe's institutional equity capital, and three-quarters of Europe's hedge fund assets. More than three-quarters of Eurobond deals are originated and executed there. More than a third of the world's swap transactions and more than a quarter of global foreign exchange transactions take place in London, which is also a major center for commodities trading, shipping and many other services.

Royal Dutch Shell now has its primary listing in the UK. Other major companies include HSBC, BP, Vodafone, British American Tobacco, and GlaxoSmithKline.

Capital market returns for the United Kingdom

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 367 as compared to 5.9 for bonds and 2.8 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 5.3%, bonds 1.6%, and bills 0.9%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 4.3%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

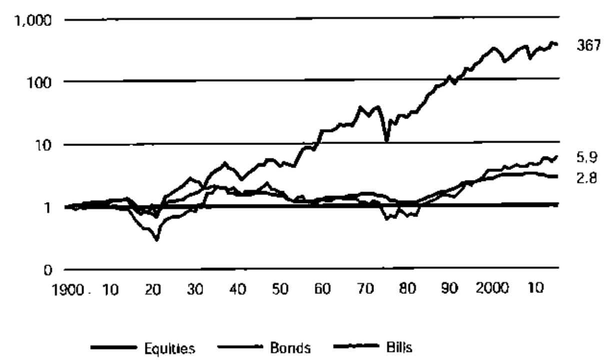


Figure 2

Annualized real returns on major asset classes (%)

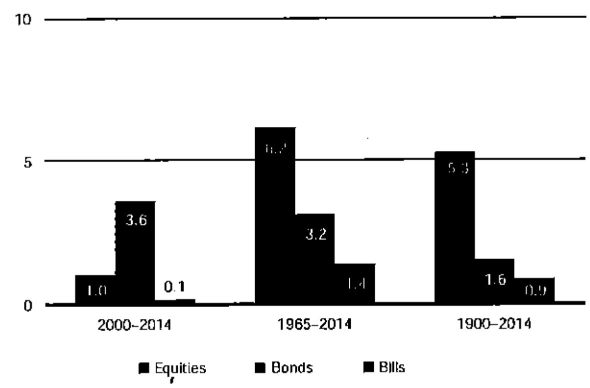
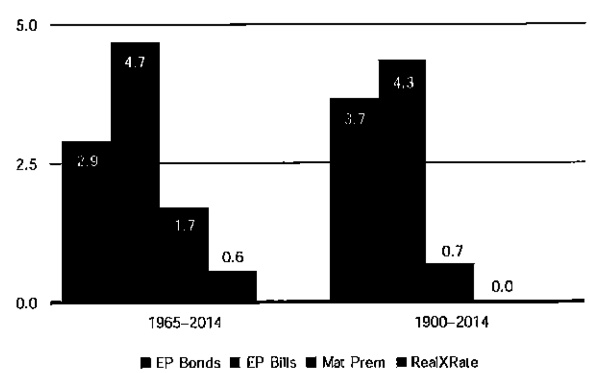


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term government bonds; EP Bills denotes the equity premium relative to Treasury bills; Mat Prem denotes the maturity premium for government bond returns relative to bill returns; and RealXRate denotes the real (inflation-adjusted) change in the exchange rate against the US dollar.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

United States

Financial superpower

In the 20th century, the United States rapidly became the world's foremost political, military, and economic power. After the fall of communism, it became the world's sole superpower. The International Energy Agency predicts that the USA will be the world's largest oil producer by 2017.

The USA is also a financial superpower. It has the world's largest economy, and the dollar is the world's reserve currency. Its stock market accounts for 52% of total world value, which is more than six times as large as Japan, its closest rival. The USA also has the world's largest bond market.

US financial markets are by far the best-documented in the world and, until recently, most of the long-run evidence cited on historical asset returns drew almost exclusively on the US experience. Since 1900, US equities and US bonds have given real returns of 6.5% and 2.0%, respectively.

There is an obvious danger of placing too much reliance on the excellent long-run past performance of US stocks. The New York Stock Exchange traces its origins back to 1792. At that time, the Dutch and UK stock markets were already nearly 200 and 100 years old, respectively. Thus, in just a little over 200 years, the USA has gone from zero to more than a one-half share of the world's equity markets.

Extrapolating from such a successful market can lead to "success" bias. Investors can gain a misleading view of equity returns elsewhere, or of future equity returns for the USA itself. That is why this Yearbook focuses on global returns, rather than just those from the USA.

Capital market returns for the United States

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 1,396 as compared to 10.1 for bonds and 2.7 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 6.5%, bonds 2.0%, and bills 0.9%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 5.6%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

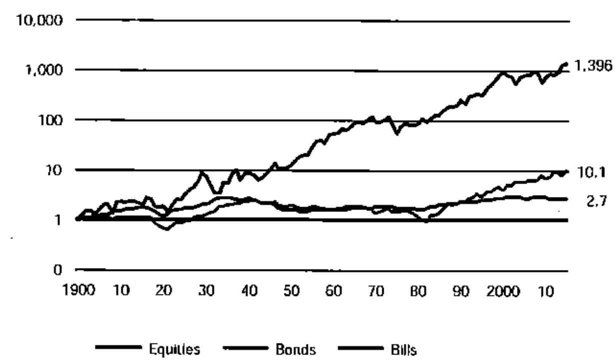


Figure 2

Annualized real returns on major asset classes (%)

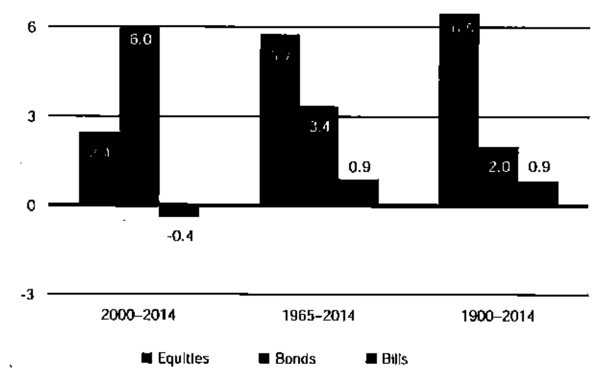
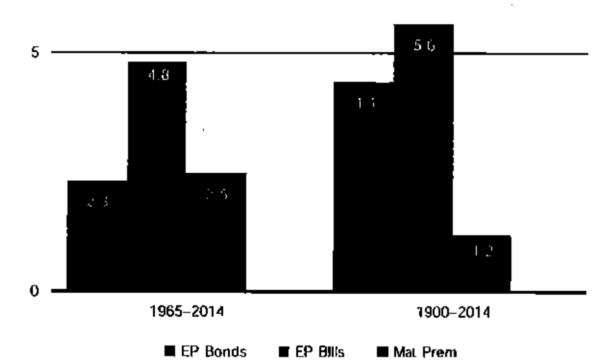


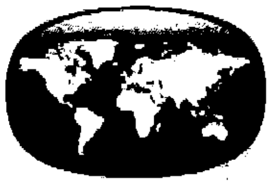
Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term US government bonds; EP Bills denotes the equity premium relative to US Treasury bills; and Mat Prem denotes the maturity premium for US government bond returns relative to US bill returns.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.



World

Globally diversified

It is interesting to see how the Yearbook countries have performed in aggregate over the long run. We have therefore created an all-country world equity index denominated in a common currency, in which each of the 23 countries is weighted by its starting-year equity market capitalization. We also compute a similar world bond index, weighted by GDP.

These indices represent the long-run returns on a globally diversified portfolio from the perspective of an investor in a given country. The charts opposite show the returns for a US global investor. The world indices are expressed in US dollars; real returns are measured relative to US inflation; and the equity premium versus bills is measured relative to US treasury bills.

Over the 115 years from 1900 to 2014, the middle chart shows that the real return on the world index was 5.2% per year for equities, and 1.9% per year for bonds. The bottom chart also shows that the world equity index had an annualized equity risk premium, relative to Treasury bills, of 4.3% over the last 115 years, and an identical premium over the most recent 50 years.

We follow a policy of continuous improvement with our data sources, introducing new countries when feasible, and switching to superior index series as they become available. In 2013, we added Austria, China and Russia; and in 2014, Portugal. Austria and Portugal have a continuous history, but China and Russia do not. To avoid survivorship bias, all these countries are fully included in the world indices from 1900 onward. Two markets register a total loss – Russia in 1917 and China in 1949. These countries then re-enter the world indices after their markets reopened in the 1990s.

Capital market returns for World (in USD)

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 325 as compared to 8.4 for bonds and 2.7 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 5.2%, bonds 1.9%, and bills 0.9%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 4.3%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

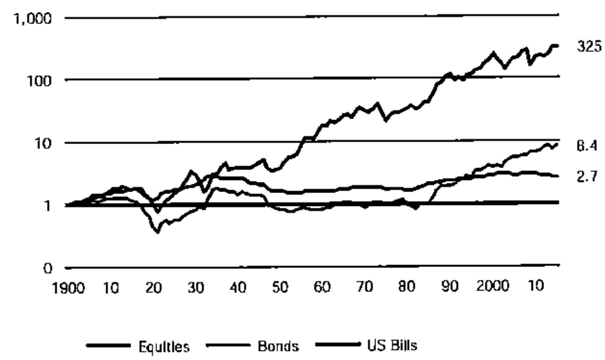


Figure 2

Annualized real returns on major asset classes (%)

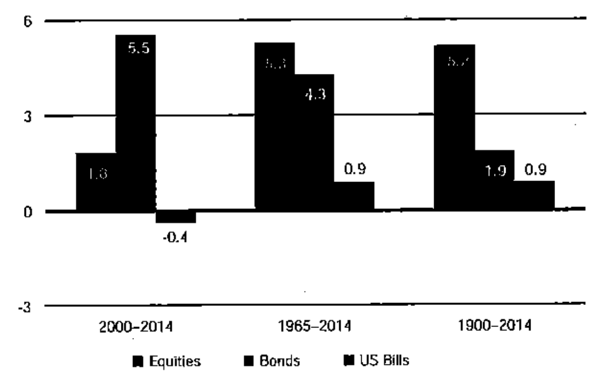
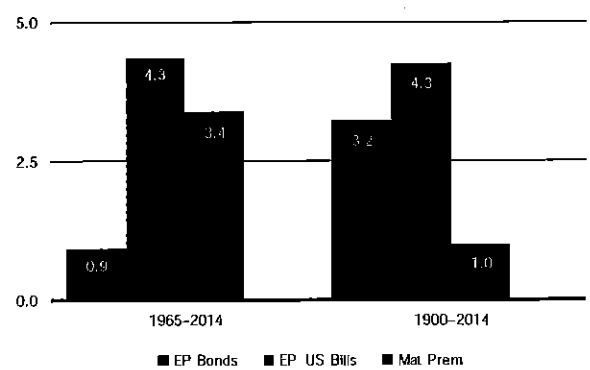


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term US government bonds; EP Bills denotes the equity premium relative to US Treasury bills; and Mat Prem denotes the maturity premium for US government bond returns relative to US bill returns.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

World ex-USA

Beyond America

In addition to the two world indices, we also construct two world indices that exclude the USA, using exactly the same principles. Although we are excluding just one out of 23 countries, the USA accounts for over half the total stock market capitalization of the Yearbook countries, so that the 22-country, world ex-US equity index represents less than half the total value of the world index today.

We noted above that, until relatively recently, most of the long-run evidence cited on historical asset returns drew almost exclusively on the US experience. We argued that focusing on such a successful economy can lead to "success" bias. Investors can gain a misleading view of equity returns elsewhere, or of future equity returns for the USA itself.

The charts opposite confirm this concern. They show that, from the perspective of a US-based international investor, the real return on the world ex-US equity index was 4.4% per year, which is 2.1% per year below that for the USA. This suggests that, although the USA has not been the most extreme of outliers, it is nevertheless important to look at global returns, rather than just focusing on the USA.

We follow a policy of continuous improvement with our data sources, introducing new countries when feasible, and switching to superior index series as they become available. In 2013 and 2014, we added Portugal, Austria, China and Russia. Portugal and Austria have a continuous history, but China and Russia do not. To avoid survivorship bias, the additional countries are fully included in the world indices from 1900 onward. Two markets register a total loss: Russia in 1917 and China in 1949. These countries then re-enter the world and world ex-USA indices after their markets reopened in the 1990s.

Capital market returns for World ex-US (in USD)

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 148 as compared to 6.0 for bonds and 2.7 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 4.4%, bonds 1.6%, and bills 0.9%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.6%. For additional explanations of these figures, see page 35.

Figure 1

Cumulative real returns from 1900 to 2014

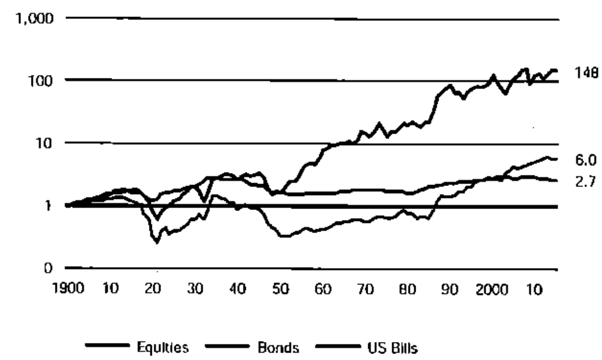


Figure 2

Annualized real returns on major asset classes (%)

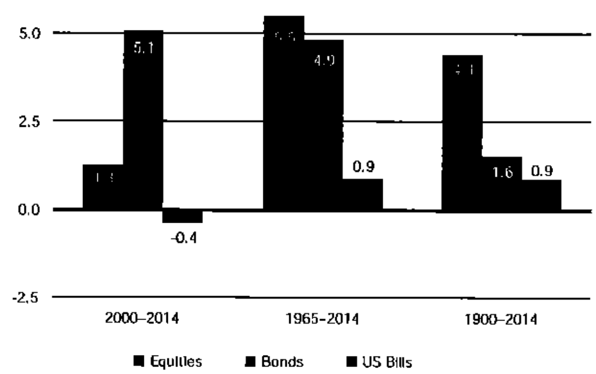
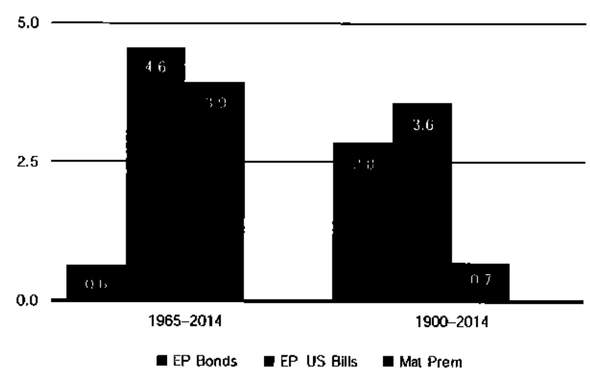


Figure 3

Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term US government bonds; EP Bills denotes the equity premium relative to US Treasury bills; and Mat Prem denotes the maturity premium for US government bond returns relative to US bill returns.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.

Europe

The Old World

The Yearbook documents investment returns for 16 European countries, most (but not all) of which are in the European Union. They comprise 10 EU states in the Eurozone (Austria, Belgium, Finland, France, Germany, Ireland, Italy, the Netherlands, Portugal, and Spain), three EU states outside the Eurozone (Denmark, Sweden and the UK), two European Free Trade Association states (Norway and Switzerland), and the Russian Federation. Loosely, we might argue that these 16 EU/EFTA countries represent the Old World.

It is interesting to assess how well European countries as a group have performed, compared with our world index. We have therefore constructed a 16-country European index using the same methodology as for the world index. As with the latter, this European index can be designated in any desired common currency. For consistency, the figures opposite are in US dollars from the perspective of a US international investor.

The middle chart opposite shows that the real equity return on European equities was 4.3%. This compares with 5.2% for the world index, indicating that the Old World countries have underperformed. This may relate to the destruction from the two world wars (where Europe was at the epicenter) or to the fact that many of the New World countries were resource-rich, or perhaps to the greater vibrancy of New World economies.

We follow a policy of continuous improvement with our data sources, introducing new countries when feasible, and switching to superior index series as they become available. This year and last year, we added three new European countries, Austria, Russia, and Portugal. Two of these countries have a continuous history, but Russia does not. To avoid survivorship bias, these countries are fully included in the Europe indices from 1900 onward, even though Russia registered a total loss in 1917. Russia re-enters the Europe indices after her markets reopened in the 1990s.

Capital market returns for Europe (in USD)

Figure 1 shows that, over the last 115 years, the real value of equities, with income reinvested, grew by a factor of 125 as compared to 3.6 for bonds and 2.7 for bills. Figure 2 displays the long-term real index levels as annualized returns, with equities giving 4.3%, bonds 1.1%, and bills 0.9%. Figure 3 expresses the annualized long-term real returns as premia. Since 1900, the annualized equity risk premium relative to bills has been 3.4%. For additional explanations of these figures, see page 35.

Figure 1
Cumulative real returns from 1900 to 2014

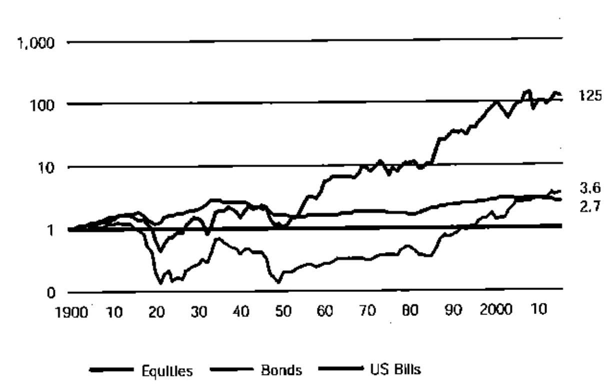


Figure 2
Annualized real returns on major asset classes (%)

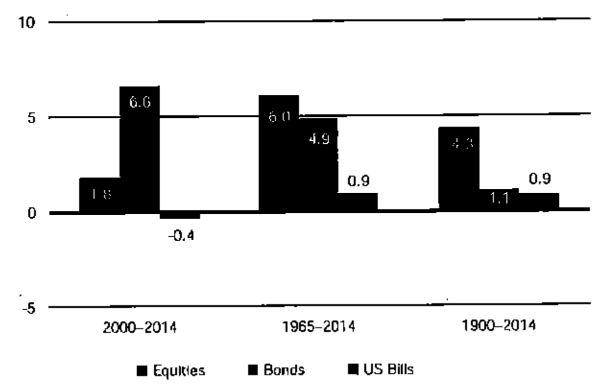
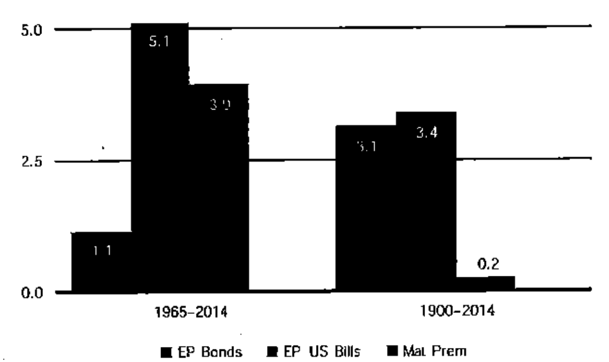
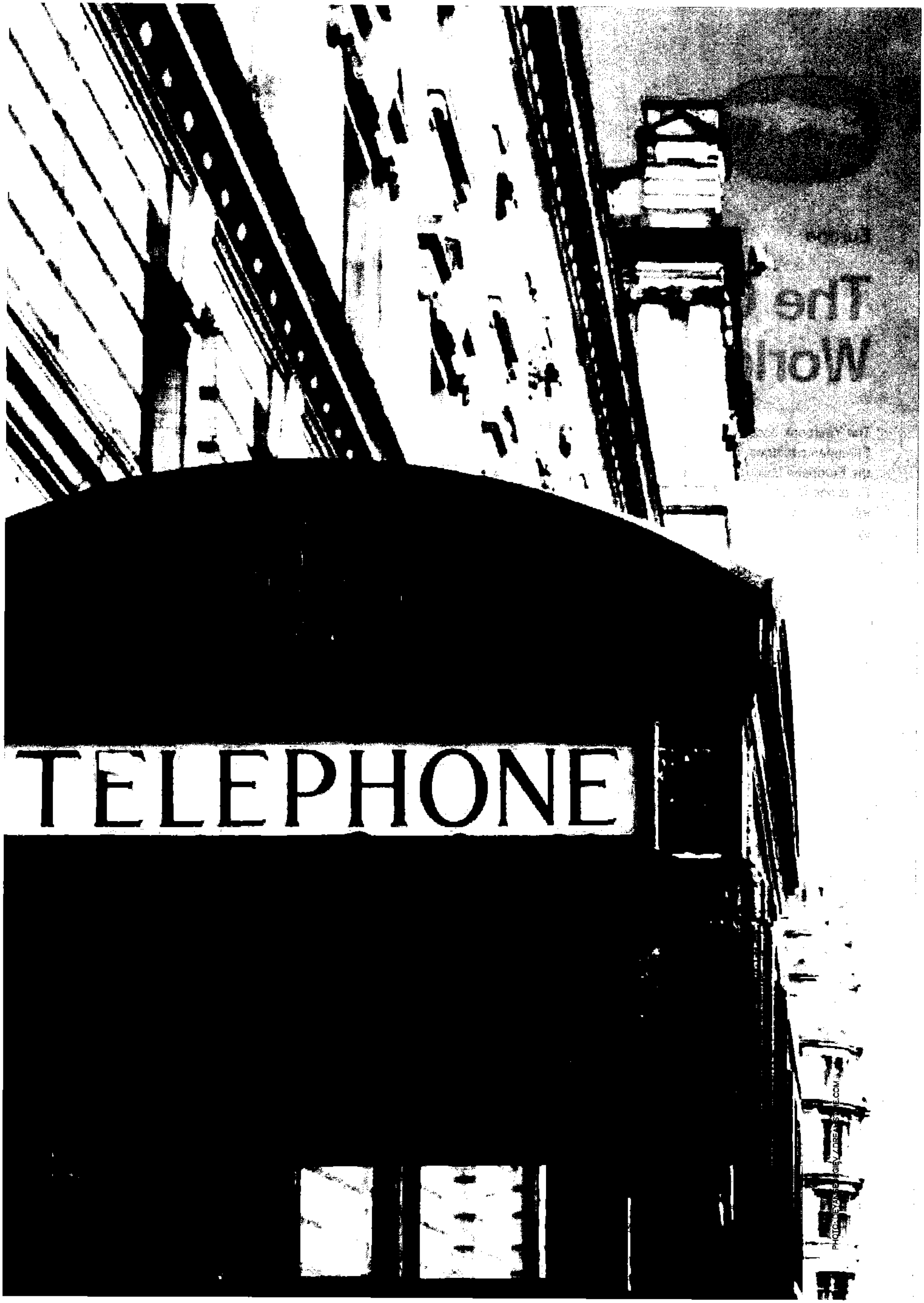


Figure 3
Annualized equity, bond, and currency premia (%)



Note: EP Bonds denotes the equity premium relative to long-term US government bonds; EP Bills denotes the equity premium relative to US Treasury bills; and Mat Prem denotes the maturity premium for USgovernment bond returns relative to US bill returns.

Source: Elroy Dimson, Paul Marsh and Mike Staunton, Credit Suisse Global Investment Returns Sourcebook 2015.


The
TELEPHONE

References

Admati, Anat, and Paul Pfleiderer, 2009, The "Wall Street Walk" and shareholder activism: exit as a form of voice, Review of Financial Studies 22, 2445–2485.

Ahrens, Dan, 2004, Investing in Vice: The Recession Proof Portfolio of Booze, Bets, Bombs and Butts, NY, New York: St Martin's Press.*

Ang, Andrew, 2014, Asset Management: A Systematic Approach to Factor Investing. Oxford: Oxford University Press.

Bebchuk, Lucian, Alma Cohen and Charles Wang, 2013, Learning and the disappearing association between governance and returns, Journal of Financial Economics 108, 323–348.

Becht, Marco, Julian Franks, Jeremy Grant and Hannes Wagner, 2014, The returns to hedge fund activism: an international study, ECGI Finance Working Paper N° 402/2014.

Borgers, Arian, Jeroen Derwall, Kees Koedijk and Jenke ter Horst, 2013, Stakeholder relations and stock returns: on errors in investors' expectations and learning, Journal of Empirical Finance 22, 159–175.

Camejo, Peter, 2003, The SRI Advantage: Why Socially Responsible Investing Has Outperformed Financially, Gabriola Island, BC: New Society Publishers.*

Cowles, Alfred, and Associates, 1938, Common Stock Indexes, first edition, Bloomington, Indiana: Principia Press, Inc.

Dimson, Elroy, Xi Li and Oguzhan Karakas, 2015, Active ownership, Review of Financial Studies, Forthcoming.

Dimson, Elroy, Idar Kreutzer, Rob Lake, Hege Sjo and Laura Starks, 2013, Responsible Investment and the Norwegian Government Pension Fund Global, Oslo: Norwegian Ministry of Finance.

Dimson, Elroy, and Paul Marsh, 2015, Numis Smaller Companies Index Annual Review, 2015, London: Numis Securities.

Dimson, Elroy, Paul Marsh and Mike Staunton, 2002. Triumph of the Optimists: 101 Years of Global Investment Returns, Princeton, NJ: Princeton University Press.

Fabozzi, Frank, KC Ma and Becky Oliphant, 2008, Sin stock returns, Journal of Portfolio Management Fall, 82 – 94.

Fama, Eugene and Kenneth French, 1997, Industry costs of equity, Journal of Financial Economics 43, 153–93.

Gollier, Christian, and Sébastien Pouget, 2014, The "washing machine": investment strategies and corporate behavior with socially responsible investors, Working Paper No 14-457, Toulouse School of Economics.

Gompers, Paul, Joy Ishii and Andrew Metrick, 2003, Corporate governance and equity prices, Quarterly Journal of Economics 118, 107–155.

Graver, Hans Peter (Chairman), 2003. The Report from the Graver Committee, Oslo: Norwegian Ministry of Finance.

Greenwood, Robin, and Michael Schor, 2009, Investor activism and takeovers, Journal of Financial Economics 92, 362–375.

Gregory, Alan, Cherif Guermat and Fawaz Al Shawawreh, 2010, UK IPOs: Long run returns, behavioural timing and pseudo timing, Journal of Business Finance and Accounting 37, 612–47.

Grundy, Bruce, and Spencer Martin, 2001, Understanding the nature of the risks and the source of rewards to momentum investing, Review of Financial Studies 14, 29–78.

Harrington, John, 1992, Investing with Your Conscience: How to Achieve High Returns Using Socially Responsible Investing, Hoboken, NJ: John Wiley & Sons.*

Heston, Steven, and Geert Rouwenhorst, 1994, Does industrial structure explain the benefits of international diversification? Journal of Financial Economics 36, 3–27.

Hirschman, Albert, 1970. Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States, Cambridge, MA: Harvard University Press.

Holland, David, and Matthews Bryant. Market-implied Returns: Past and Present, Credit Suisse Global Investment Returns Yearbook 2011, January 2011, pages 25–29.

Hong, Harrison, and Marcin Kacperczyk, 2009, The price of sin: the effects of social norms on markets, Journal of Financial Economics 93, 15–36.

Humphrey, Jacquelyn, and David Tan, 2014, Does it really hurt to be responsible?, Journal of Business Ethics 122 (3), 375–386.

Huppé, Gabriel, and Tessa Hebb, 2011, The virtue of CalPERS' Emerging Equity Markets Principles, Journal of Sustainable Investment and Finance 1(1), 62–76.

Ibbotson, Roger, 1975, Price performance of common stock new issues, Journal of Financial Economics 2, 235–72.

Ilmanen, Antti, Expected Returns: An Investor's Guide to Harvesting Market Rewards, Chichester: John Wiley & Sons.

Kaufmann, Daniel, Aart Kraay and Massimo Mastruzzi, 2010. The Worldwide Governance Indicators: Methodology and Analytical Issues, World Bank Policy Research Working Paper No. 5430.

Lessard, Don, 1974, World, national and industry factors in equity returns, Journal of Finance 29, 379–91.

Loughran, Timothy and Jay Ritter, 1995, The new issues puzzle, Journal of Finance 50, 23–51.

Luo, Hao (Arthur), and Ronald Balvers, 2014, Social screens and systematic boycott risk, Working paper, McMaster University.

Menchero, Jose, and Andrei Morozov, 2012, The relative strength of industries versus countries in global equity markets, Journal of Investment Management 10, 75–87.

Monks, Robert, and Nell Minow, 1995 (fifth edition, 2011). Watching the Watchers: Corporate Governance in the 21st Century, Chichester: John Wiley & Sons.

Moskowitz, Tobias, and Mark Grinblatt, 1999, Do industries explain momentum, Journal of Finance 54, 1249–90.

Nairn, Alasdair, 2002, Engines that Move Markets: Technology Investing from Railroads to the Internet and Beyond, New York, NY: John Wiley & Sons, Inc.

Nystuen, Gro, Andreas Follesdal and Ola Mestad (Eds.), 2011, Human Rights, Corporate Complicity and Disinvestment, New York, NY: Cambridge University Press.

Richardson, Benjamin, 2013. Socially responsible investing for sustainability: overcoming its incomplete and conflicting rationales, Transnational Environmental Law 2(2), 311–338.

Ritter, Jay, 2014, Initial public offerings: Updated statistics on long-run performance, October 7, Jay Ritter's data website.

Rostow, Walt, and Anna Schwartz, 1953, The Growth and Fluctuation of the British Economy, 1780–1850 (2 vols), Oxford: Oxford University Press.

Scowcroft, Alan, and James Sefton, 2005, Understanding momentum, Financial Analysts Journal 61, 64–82.

Shubb, Gary, 2014, Global Asset Management 2014: Steering the Course to Growth, Boston, MA: Boston Consulting Group.

Siegel, Jeremy, 2005, The Future for Investors: Why the Tried and the True Triumph over the Bold and the New, New York, NY: Crown Publishing Group.

Stierli, Markus, 2014, Sustainable Investments, Research Alert, Zurich: Credit Suisse.

Waxler, Caroline, 2004, Stocking Up on Sin: How to Crush the Market with Vice-Based Investing, Hoboken, NJ: John Wiley & Sons.*

* The dust jackets for these four books, depicted in Figure 2 of the second article above, are reproduced by kind permission of the publishers.

About the authors

Elroy Dimson is Chairman of the Centre for Endowment Asset Management at Cambridge Judge Business School, Emeritus Professor of Finance at London Business School, Chairman of the Strategy Council for the Norwegian Government Pension Fund, and Chairman of FTSE's Advisory Board and Policy Committee. He previously served London Business School in a variety of senior positions. With Paul Marsh and Mike Staunton, he wrote Triumph of the Optimists, and he has published in Journal of Finance, Journal of Financial Economics, Review of Financial Studies, Journal of Business, Journal of Portfolio Management, Financial Analysts Journal, and other journals. His PhD in Finance is from London Business School.

Paul Marsh is Emeritus Professor of Finance at London Business School. Within London Business School he has been Chair of the Finance area, Deputy Principal, Faculty Dean, an elected Governor and Dean of the Finance Programmes, including the Masters in Finance. He has advised on several public enquiries; was previously Chairman of Aberforth Smaller Companies Trust, and a non-executive director of M&G Group and Majedie Investments; and has acted as a consultant to a wide range of financial institutions and companies. Dr Marsh has published articles in Journal of Business, Journal of Finance, Journal of Financial Economics, Journal of Portfolio Management, Harvard Business Review, and other journals. With Elroy Dimson, he co-designed the FTSE 100-Share Index and the Numis Smaller Companies Index, produced since 1987 at London Business School. His PhD in Finance is from London Business School.

Mike Staunton is Director of the London Share Price Database, a research resource of London Business School, where he produces the London Business School Risk Measurement Service. He has taught at universities in the United Kingdom, Hong Kong and Switzerland. Dr Staunton is co-author with Mary Jackson of Advanced Modelling in Finance Using Excel and VBA, published by Wiley and writes a regular column for Wilmott magazine. He has had articles published in Journal of Banking & Finance, Financial Analysts Journal, Journal of the Operations Research Society, and Quantitative Finance. With Elroy Dimson and Paul Marsh, he co-authored the influential investment book Triumph of the Optimists, published by Princeton University Press, which underpins this Yearbook and the accompanying Credit Suisse Global Investment Returns Sourcebook 2015. His PhD in Finance is from London Business School.

David Holland is a Senior Advisor to Credit Suisse. He was previously co-Head of HOLT Valuation & Analytics, the research and development arm of HOLT. He was responsible for improvements to the HOLT CFROI® framework, the development of custom solutions and investment products, and the generation of HOLT content. Mr Holland joined CSFB in May 2002 from Fractal Value Advisors, a corporate advisory business he owned in South Africa. He has worked with the world's leading fund managers designing investment decision processes, advised corporations and private equity firms on strategy and valuation, and regularly writes about equity valuation issues. He co-authored the book Beyond Earnings: A User's Guide to Excess Return Models and the HOLT CFROI® Framework published by Wiley Finance. Mr Holland holds degrees in Chemical Engineering from the University of Illinois (BS) and Stanford University (MS), and an MBA from the University of Cape Town.

Bryant Matthews is the Director of Research at Credit Suisse HOLT Division. He joined Credit Suisse in January 2002 when the firm acquired HOLT Value Associates LP. He has acted as consultant for various banks and professional bodies in the field of finance, and designed numerous valuation frameworks for Institutional Investors. Research conducted by Mr Matthews and his team has resulted in many significant insights into understanding the link between corporate performance and stock price. Areas of recent investigation include empirical research on corporate fade dynamics, estimating discount rates through a market-derived approach, and the development of risk models useful in understanding how risk evolves over time. Mr Matthews is an MBA graduate of Northwestern Kellogg Graduate School of Management.

Pratyasha Rath is a Vice President at Credit Suisse HOLT Division. In her current role of Global Macro Specialist, she is responsible for developing strategic macro investment insights and country/sector allocation ideas for clients. She joined Credit Suisse in 2007 as part of the Research team responsible for continuous improvement of the HOLT framework. Research conducted by Pratyasha has resulted in an improved assessment of risk for government-owned financial firms and real estate firms within HOLT. She has been instrumental in developing and marketing the risk analytics and macro offering at HOLT. Pratyasha holds a BA in Economics from the University of Delhi, India and an MS Finance from Illinois Institute of Technology.

Imprint

PUBLISHER
CREDIT SUISSE AG
Research Institute
Paradeplatz 8
CH-8070 Zurich
Switzerland

cs.researchinstitute@credit-suisse.com

PRODUCTION MANAGEMENT
INVESTMENT STRATEGY & RESEARCH
INVESTMENT PUBLISHING
Markus Kleeb (Head)
Ross Hewitt
Katharina Schlatter

AUTHORS
Elroy Dimson, London Business School, edimson@london.edu
Paul Marsh, London Business School, pmarsh@london.edu
Mike Staunton, London Business School, mstaunton@london.edu
Author of pages 29-33
David A. Holland, Credit Suisse, david.a.holland@credit-suisse.com
Bryant A. Matthews, Credit Suisse, bryant.matthews@credit-suisse.com
Pratyasha Rath, Credit Suisse, pratyasha.rath@credit-suisse.com

Editorial deadline
30 January 2015

Additional copies
Additional copies of this publication can be ordered via the Credit Suisse publication shop www.credit-suisse.com/publications or via your customer advisor.
Copies of the Credit Suisse Global Investment Returns Sourcebook 2015 can be ordered via your customer advisor (Credit Suisse clients only). Non-clients please see below for ordering information.

ISBN 978-3-9524302-2-4

To receive a copy of the Yearbook or Sourcebook (non-clients): This Yearbook is distributed to Credit Suisse clients by the publisher, and to non-clients by London Business School. The accompanying volume, which contains detailed tables, charts, listings, background, sources, and bibliographic data, is called the Credit Suisse Global Investment Returns Sourcebook 2015. The Sourcebook includes detailed tables, charts, listings, background, sources, and bibliographic data. Format: A4 color, perfect bound, 224 pages, 30 chapters, 152 charts, 85 tables, 145 references. ISBN 978-3-9524302-3-1. Please address requests to Patricia Rowham, London Business School, Regents Park, London NW1 4SA, United Kingdom, tel +44 20 7000 7000, fax +44 20 7000 7001, e-mail prowham@london.edu. E-mail is preferred.

To gain access to the underlying data: The Dimson-Marsh-Staunton dataset is distributed by Morningstar Inc. Please ask for the DMS data module. Further guidelines on subscribing to the data are available at www.tinyurl.com/DMSdata.

To quote from this publication: To obtain permission, contact the authors with details of the required extracts, data, charts, tables or exhibits. In addition to citing this publication, documents that incorporate reproduced or derived materials must include a reference to Elroy Dimson, Paul Marsh and Mike Staunton, Triumph of the Optimists: 101 Years of Global Investment Returns, Princeton University Press, 2002. Apart from the article on pages 29-33, each chart and table must carry the acknowledgement Copyright © 2015 Elroy Dimson, Paul Marsh and Mike Staunton. If granted permission, a copy of published materials must be sent to the authors at London Business School, Regents Park, London NW1 4SA, United Kingdom.

General disclaimer / Important information

Also published by the Research Institute

Family businesses: Sustaining performance
September 2012

Global Wealth Report 2012
October 2012

The shale revolution
December 2012

Emerging Consumer Survey 2013
January 2013

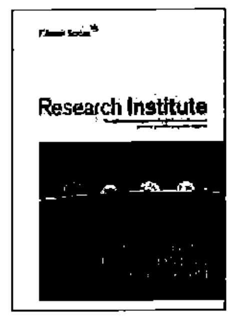
Credit Suisse Global Investment Returns Yearbook 2013
February 2013

Sugar Consumption at a crossroads
September 2013

Global Wealth Report 2013
October 2013

Latin America: The long road
February 2014

Emerging Consumer Survey 2014
February 2014



Credit Suisse Global Investment Returns Yearbook 2014
February 2014

Emerging capital markets: The road to 2030
July 2014

The Success of Small Countries
July 2014

The CS Gender 3000: Women in Senior Management
September 2014

Global Wealth Report 2014
October 2014

Emerging Consumer Survey 2015
January 2015

CREDIT SUISSE

CREDIT SUISSE AG
Research Institute
Paradeplatz 8
CH-8070 Zurich
Switzerland

cs.researchinstitute@credit-suisse.com
www.credit-suisse.com/researchinstitute

FSC
www.fsc.org
MIX
Paper from responsible sources
FSC® C009035

PERFORMANCE
myclimate
neutral
Printed Matter
No. 01-14-637273 – www.myclimate.org
© myclimate – The Climate Protection Partnership

RCE 1546014 ISBN 978-3-9524302-2-4 02/2015

Estimating the Ex Ante Equity Premium

R. Glen Donaldson
Sauder School of Business
University of British Columbia
glen.donaldson AT sauder.ubc.ca

Mark J. Kamstra
Schulich School of Business
York University
mkamstra AT schulich.yorku.ca

Lisa A. Kramer
Rotman School of Management
University of Toronto
Lkramer AT rotman.utoronto.ca

October 2007

Keywords: equity risk premium; simulated method of moments; SMM

JEL classifications: G12, C13, C15, C22

We have benefited from the suggestions of Wayne Ferson, Ian Garrett, Mark Fisher, Joel Hasbrouck, Raymond Kan, Patrick Kelly, Alan Kraus, Tom McCurdy, Federico Nardari, Cesare Robotti, Jacob Sagi, Tan Wang, participants of the Western Finance Association Meetings, the Northern Finance Association Meetings, the Canadian Econometrics Study Group, the European Econometric Society Meetings, the investment conference of the University of Colorado at Boulder's Burridge Center for Securities Analysis and Valuation, and seminar participants at the Board of Governors of the Federal Reserve System, Emory University, the Federal Reserve Bank of Atlanta, Queen's University, the US Securities & Exchange Commission, and the University of British Columbia. We thank the Social Sciences and Humanities Research Council of Canada for financial support. Any remaining errors are our own. A previous version of this paper was titled "Stare Down the Barrel and Center the Crosshairs: Targeting the Ex Ante Equity Premium." The most recent version of this paper is available from: http://ssrn.com/abstract=945192.

Estimating the Ex Ante Equity Premium

Abstract

We find that the true ex ante equity premium very likely lies within 50 basis points of 3.5%. This estimate is similar to values obtained in some recent studies but is considerably more precise. In addition to narrowing the range of plausible ex ante equity premia, we also find that equity premium models that allow for time-variation, breaks, and/or trends are the models that best match the experience of US markets and are the only models not rejected by our specification tests. This suggests that time-variation, breaks, and/or trends are critical features of the equity premium process. Our approach involves simulating the distribution from which interest rates, dividend growth rates, and equity premia are drawn and determining the prices and returns consistent with these distributions. We achieve the narrower range of ex ante equity premium values and the narrower set of plausible models by comparing statistics that arise from our simulations with key financial characteristics of the US economy, including the mean dividend yield, return volatility, and mean return. Our findings are achieved in part with the imposition of more structure than is typically exploited in the literature. In order to mitigate the potential for misspecification with this additional structure, we consider a broad collection of models that variously do or do not incorporate features such as an adjustment in dividend growth rates to account for recently increased share repurchase activity, sampling uncertainty in generating model parameters, and cross-correlation between interest rates, dividend growth rates, and equity premia.

Estimating the Ex Ante Equity Premium

Financial economic theory is often concerned with the premium that investors demand ex ante, when they first decide whether to purchase risky stocks instead of risk-free debt. In contrast, empirical tests of the equity premium often focus on the return investors received ex post.[1] It is well known that estimates of the ex ante equity premium based on ex post data can be very imprecise; such estimates have very wide margins of error, as wide as 1000 basis points in typical studies and 320 basis points in some recent studies. This fact makes it challenging to employ the equity premium estimates for common practical purposes, including evaluating the equity premium puzzle, performing valuation, and conducting capital budgeting. The imprecision of traditional equity premium estimates also makes it difficult to determine if the equity premium has changed over time. Our goals, therefore, are to develop a more precise estimate of the ex ante equity premium and to determine what kind of equity premium model can be supported by the experience of US markets. We accomplish these goals by employing simulation techniques that identify a range of models of the equity premium and the values of the ex ante equity premium that are consistent with values of several key financial statistics that are observed in US market data, including dividend growth rates, interest rates, Sharpe ratios, price-dividend ratios, volatilities, and of course the ex post equity premium.

Our results suggest that the mean ex ante equity premium lies within 50 basis points of 3.5%. These results stand even when we allow for investors' uncertainty about the true state of the world. The tightened bounds are achieved in part with the imposition of more structure than has been commonly employed in the equity premium literature. In order to mitigate the potential for misspecification with this additional structure, we consider a broad collection of models that variously do or do not incorporate features such as a conditionally time-varying equity premium, a downward trend in the equity premium, a structural break in the equity premium, an adjustment in dividend growth rates to account for increased share repurchase activity in the last 25 years, sampling uncertainty in generating model parameters, a range of time series models, and cross-correlation between interest rates, dividend growth rates, and equity premia. We also find that

[1] The equity premium literature is large, continuously growing, and much too vast to fully cite here. For recent work, see Bansal and Yaron (2004), Graham and Harvey (2005), and Jain (2005). For excellent surveys see Kocherlakota (1996), Siegel and Thaler (1997), Mehra and Prescott (2003), and Mehra (2003).

equity premium models that allow for time-variation, breaks, and/or trends in the equity premium process are the models that best match the experience of US markets and are the only models not rejected by our specification tests. This suggests that time-variation, breaks, and/or trends are critical features of the equity premium process, itself an important finding.

We draw on two relatively new techniques in order to provide a more precise estimate of the equity premium than is currently available. The first technique builds on the fundamental valuation dividend discounting method of Donaldson and Kamstra (1996). This technique permits the simulation of fundamental prices, returns, and return volatility for a given ex ante equity premium. Donaldson and Kamstra find that if we allow dividend growth rates and discount rates to be time-varying and dependent, as well as cross-correlated, the fundamental prices and returns that come out of dividend discounting match observed prices and returns, even during extreme events like stock market crashes. The second technique is simulated method of moments (SMM).[2] An attractive feature of SMM is that the estimation of parameters requires only that the model, with a given set of parameters, can generate data. SMM forms estimates of model parameters by using a given model with a given set of parameter values to simulate moments of the data (for instance means or volatilities), measuring the distance between the simulated moments and the actual data moments, and repeating with new parameter values until the parameter values that minimize the (weighted) distance are found.[3] The parameter estimates that minimize this distance are consistent for the true values, are asymptotically normally distributed, and display the attractive feature of permitting tests that can reject misspecified models. The SMM technique has been described as "estimating on one group of moments, testing on another." See Cochrane (2001, Section 11.6). We use SMM rather than GMM because, as we show below, the economic model we use is nonlinear in the parameters and cannot be solved without the use of SMM.

We exploit the dividend discounting method of Donaldson and Kamstra to generate simulated fundamental prices, dividends, returns, and derivative moments such as the mean ex post equity

[2]Simulated method of moments was developed by McFadden (1989) and Pakes and Pollard (1989), and a helpful introduction to the technique is provided in Carrasco and Florens (2002). Examples of papers that employ SMM in an asset pricing context are Duffie and Singleton (1993) and Corradi and Swanson (2005).

[3]The typical implementation of SMM is to weight the moments inversely to their estimated precision; that is minimize the product of the moments weighted by the inverse of the covariance matrix of the moments. This is the approach we adopt.

premium, mean dividend yield, and return volatility for a given ex ante equity premium. We minimize (by choice of the ex ante equity premium) the distance between the simulated moments that the model produces and the moments observed in US stock markets over the past half century. That is, given various characteristics of the US economic experience (such as low interest rates and a high ex post equity premium, high Sharpe ratios and low dividend yields, *etc.*), we determine the range of values of the ex ante equity premium and the set of equity premium models that are most likely to have generated the observed collection of sample moments.

To undertake our study, we consider a broad collection of models, including models with and without conditional time-variation in the equity premium process, with and without trends in the equity premium, with and without breaks in the equity premium, with and without breaks in the dividend growth rate, as well as various autoregressive specifications for dividend growth rates, interest rates, and the equity premium. Virtually every model we consider achieves a minimum distance between the simulated moments and the actual data moments by setting the ex ante equity premium between 3% and 4%, typically very close to 3.5%. That is, the equity premium estimate is very close to 3.5% across our models. Further, the range of ex ante equity premium values that can be supported by the US data for a given model is typically within plus or minus 50 basis points of 3.5%. Our models of fundamentals, which capture the dynamics of actual US dividend and interest rate data, imply that the true ex ante equity premium is 3.5% plus or minus 50 basis points. Simpler models of fundamental valuation, such as the Gordon (1962) constant dividend growth model, are overwhelmingly rejected by the data. Models of the equity premium which do not allow time-variation, trends, or breaks are also rejected by the SMM model specification tests. While we restrict our attention to a stock market index in this study, the technique we employ is more broadly applicable to estimating the equity premium of an individual firm.

In the literature to date, empirical work investigating the equity premium has largely consisted of a series of innovations around a common theme: producing a better estimate of the mean ex ante equity premium. Recent work in the area has included insights such as exploiting dividend yields or earnings yields to provide new, more precise estimates of the return to holding stocks (see Fama and French, 2002, and Jagannathan, McGrattan, and Scherbina, 2000), looking across many countries to account for survivorship issues (see Jorion and Goetzmann, 1999), looking across many countries to decompose the equity premium into dividend growth, price-dividend ratio, dividend yield, and real exchange rate components (see Dimson, Marsh, and Staunton, 2007), modeling equity premium structural breaks in a Bayesian econometric framework (see Pástor and Stambaugh, 2001), or computing out-of-sample forecasts of the distribution of excess returns, allowing for structural breaks which are identified in real time (see Maheu and McCurdy, 2007). Most of this work estimates the ex ante equity premium by considering one moment of the data at a time, typically the mean difference between an estimate of the return to holding equity and a risk free rate, though Maheu and McCurdy (2007) consider higher-order moments of the excess return distribution and Pástor and Stambaugh (2001) incorporate return volatility and direction of price movements through their use of priors.

Unfortunately, the equity premium is still estimated without much precision. Pástor and Stambaugh (2001), exploiting extra information from return volatility and prices, narrow a two standard deviation confidence interval around the value of the ex ante equity premium to plus or minus roughly 280 basis points around a mean premium estimate of roughly 4.8% (a range that spans 2% to 7.6%) and determine that the data strongly support at least one break in the equity premium in the last half century. Fama and French (2002), based on data from 1951 to 2000, provide point estimates of the ex post equity premium of 4.32% (based on earnings growth rate fundamentals) plus or minus roughly 400 basis points (again, two standard deviations) and of 2.55% (based on dividend growth rate fundamentals) plus or minus roughly 160 basis points: a range of approximately 0.95% to 4.15%. That is, the plausible range of equity premia that emerge from Fama and French's study occupy a confidence bound with a width of anywhere from 320 to 800 basis points. Claus and Thomas (2001), like Fama and French (2002), make use of fundamental information to form lower estimates of the ex post equity premium, but their study covers a shorter time period relative to the Fama and French study – 14 years versus 50 years – yielding point estimates that are subject to at least as much variability as the Fama and French estimates.

Not only are the point estimates from the existing literature imprecisely estimated in terms of their standard error, there is also less of an emerging consensus than one would hope. Fama and French (2002) produce point estimates of 2.55% (using dividend yields) and 4.78% (using earnings yields), Pástor and Stambaugh (2001) estimate the equity premium at the end of the 1990s to be 4.8%, and Claus and Thomas (2001) estimate the equity premium to be *no more* than 3%. Welch (2000), surveying academic financial economists, estimates the consensus equity premium to be between 6% and 7% (depending on the horizon). Based on a survey of US CFOs, Graham and Harvey (2005) estimate the ten-year equity premium to be 3.66%. We believe that the lack of consensus across the literature is intimately tied to the imprecision of techniques typically used to estimate the equity premium, such as the simple average excess return. That is, the various estimates cited above all fall within two standard errors of the sample mean estimate of the equity premium, based on US data. Further, the studies that provide these estimates do not explicitly consider which models of the equity premium process can be rejected by actual data, though Pástor and Stambaugh's analysis strongly supports a model that incorporates breaks in the equity premium process.

The remainder of our paper proceeds as follows. The basic methodology of our simulation approach to estimating equity premia is presented in Section 1, along with important details on estimating the equity premium. (Appendices to the paper provide detailed explanations of the technical aspects of our simulations, including calibration of key model parameters.) In Section 2 we compare univariate financial statistics that arise in our simulations with US market data, including dividend yields, Sharpe ratios, and conditional moments including ARCH coefficients. Our results confirm that the simulations generate data broadly consistent with the US market data and, taken one-at-a-time, these financial statistics imply that the ex ante equity premium lies in a range much narrower than between 2% and 8%. We determine how much narrower in Section 3 by exploiting the full power of the simulation methodology. We compare joint multivariate distributions of our simulated data with observed US data, yielding a very precise estimate of the ex ante equity premium and providing strong rejections of models of the equity premium process that fail to incorporate time variation, breaks, and/or trends. We find the range of ex ante equity premium values is very narrow: 3.5% plus or minus 50 basis points. Our consideration of a broad collection of possible data generating processes and models lends confidence to the findings. Section 4 concludes.

I Methodology

Consider a stock for which the price P_t is set at the beginning of each period t and which pays a dividend D_{t+1} at the end of period t. The return to holding this stock (denoted R_t) is defined as

$$R_t = \frac{D_{t+1} + P_{t+1} - P_t}{P_t}.$$

The risk-free rate, set at the beginning of each period, is denoted $r_{t,f}$. The ex ante equity premium, π, is defined as the difference between the expected return on risky assets, $E\{R_t\}$, and the expected risk-free rate, $E\{r_{t,f}\}$:[4]

$$\pi \equiv E\{R_t\} - E\{r_{t,f}\}. \tag{1}$$

We do not observe this ex ante equity premium. Empirically, we only observe the returns that investors actually receive ex post, after they have purchased the stock and held it over some period of time during which random economic shocks impact prices. Hence, the ex post equity premium is typically estimated using historical equity returns and risk-free rates. Define $\overline{R}$ as the average historical annual return on the S&P 500 and $\overline{r}_f$ as the average historical return on US T-bills. Then we can calculate the estimated ex post equity premium, $\hat{\pi}$, as follows:

$$\hat{\pi} \equiv \overline{R} - \overline{r}_f. \tag{2}$$

Given that the world almost never unfolds exactly as one expects, there is no reason to believe that the stock return we estimate ex post is exactly the same as the return investors anticipated ex ante. It is therefore difficult to argue that just because we observe a 6% ex post equity premium in the US data, the premium that investors demand ex ante is also 6% and thus a puzzling challenge to economic theory. So we ask the following question: If investors' true ex ante premium is π, what is the probability that the US economy could randomly produce an ex post premium of at least 6%? The answer to this question has implications for whether or not the 6% ex post premium

[4]See, for instance, Mehra and Prescott (1985), Equation (14). We will consider time-varying equity premium models below.

observed in the US data is consistent with various ex ante premium values, π, with which standard economic theory may be more compatible. We also ask a deeper question: If investors' true ex ante premium is π, what is the probability that we would observe the various *combinations* of key financial statistics and yields that have been realized in the US, such as high Sharpe ratios and low dividend yields, high return volatility and a high ex post equity premium, and so on? The analysis of multivariate distributions of these statistics allows us to narrow substantially the range of equity premia consistent with the US market data, especially relative to previous studies that have considered univariate distributions.

Because the *empirical* joint distribution of the financial statistics we wish to consider is difficult or impossible to estimate accurately, in particular the joint distribution *conditional* on various ex ante equity premium values, we use simulation techniques to estimate this distribution. The simulated joint distribution allows us to conduct formal statistical tests that a given ex ante equity premium could have produced the US experience. Most of our models employ a time-varying ex ante equity premium, so that a simulation described as having an ex ante equity premium of 2.75% actually has a mean ex ante equity premium of 2.75%, while period-by-period the ex ante equity premium can vary somewhat from this mean value. In what follows we refer to the ex ante equity premium and the mean ex ante equity premium interchangeably.

A Matching Moments

Consider the valuation of a stock. Define $1+r_t$ as the gross rate investors use to discount payments received during period t. The price of the stock is then given by Equation (3),

$$P_t = E_t\left\{\frac{D_{t+1}+P_{t+1}}{1+r_t}\right\}, \tag{3}$$

where E_t is the conditional expectations operator incorporating information available to the market when P_t is formed, up to but not including the beginning of period t (*i.e.*, information from the end of period $t-1$ and earlier).

Assuming the usual transversality conditions, we can derive Equation (4) by recursively substituting out for future prices in Equation (3):

$$P_t = E_t \left\{ \sum_{j=0}^{\infty} \left(\Pi_{i=0}^{j} \frac{1}{1+r_{t+i}} \right) D_{t+j+1} \right\}. \tag{4}$$

Defining the growth rate of dividends over the period t as $g_t \equiv (D_{t+1} - D_t)/D_t$, we can re-write Equation (4) as

$$P_t = D_t E_t \left\{ \sum_{j=0}^{\infty} \left(\Pi_{i=0}^{j} \left[\frac{1+g_{t+i}}{1+r_{t+i}} \right] \right) \right\}. \tag{5}$$

Hence we can re-write Equation (1) as

$$\pi \equiv E \left\{ \frac{D_{t+1} + D_{t+1} E_{t+1} \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+1+i}}{1+r_{t+1+i}} \right\} - D_t E_t \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+i}}{1+r_{t+i}} \right\}}{D_t E_t \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+i}}{1+r_{t+i}} \right\}} - r_{t,f} \right\} \tag{6}$$

or

$$\pi \equiv E \left\{ \frac{(1+g_t)\left(1 + E_{t+1} \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+1+i}}{1+r_{t+1+i}} \right\}\right) - E_t \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+i}}{1+r_{t+i}} \right\}}{E_t \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+i}}{1+r_{t+i}} \right\}} - r_{t,f} \right\}. \tag{7}$$

In the case of a constant equity premium π and a possibly time-varying risk-free interest rate we can re-write Equation (7) as

$$\pi \equiv E \left\{ \frac{(1+g_t)\left(1 + E_{t+1} \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+1+i}}{1+\pi+r_{t+1+i,f}} \right\}\right) - E_t \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+i}}{1+\pi+r_{t+i,f}} \right\}}{E_t \left\{ \sum_{j=0}^{\infty} \Pi_{i=0}^{j} \frac{1+g_{t+i}}{1+\pi+r_{t+i,f}} \right\}} - r_{t,f} \right\}. \tag{8}$$

Under interesting conditions, such as risk-free rates and dividend growth rates that conditionally time-vary and covary (we consider, for instance, ARMA models and correlated errors for dividend growth rates and interest rates), the individual conditional expectations in Equation (8) are analytically intractable. The difference between the sample mean return and the sample mean risk-free interest rate provides a consistent estimate of π, as shown by Mehra and Prescott (1985), but unfortunately the sample mean difference is very imprecisely estimated, even based on more than 100 years of data.

We note that another consistent estimator of π is one that directly exploits the method of Donaldson and Kamstra (1996), hereafter referred to as the DK method. The DK method uses (ARMA) models for dividend growth rates and interest rates to simulate the conditional expectations $E_t\left\{\sum_{j=0}^{\infty}\Pi_{i=0}^{j}\frac{1+g_{t+i}}{1+\pi+r_{t+i,f}}\right\}$ and $E_{t+1}\left\{\sum_{j=0}^{\infty}\Pi_{i=0}^{j}\frac{1+g_{t+1+i}}{1+\pi+r_{t+1+i,f}}\right\}$. The DK method allows us, for a given *ex ante* equity premium (or time-varying equity premium process), to simulate the conditional expectations in Equation (8) as well as related (unconditional) moments, including the expected dividend yield, return volatility, ex post equity premium, and Sharpe ratio. Our estimate of π is produced by finding the value of π that minimizes the distance between the collection of simulated moments (produced by the DK procedure) and the analogous sample moments (from the US experience over the last half century). The estimation of these expectations relies on the exact form of the conditional models for dividend growth rates and interest rates, that is, the parameters that characterize these models. A joint estimation of these models' parameters and π (*i.e.* minimizing the distance between simulated and sample moments by varying all the model's parameters and π at once) would be computationally very difficult. We utilize a two-step procedure in which first, for a given ex ante equity premium, we jointly estimate the parameters that characterize the evolution of dividend growth rates and interest rates. We use these models to simulate data to compare with realized S&P 500 data. Second, we do a grid search over values of the ex ante equity premium to find our SMM estimate of π.

It is helpful to consider some examples of estimators based on our simulation technique. The simplest estimator would have us considering only the ex ante equity premium moment, $\pi = E\left[R_t\right] - E\left[r_{f,t}\right]$, ignoring other potentially informative moments of the data, such as the dividend yield and return volatility. Exploiting the DK procedure, we would find that the π in Equation (8) which matches the ex post equity premium (the sample moment analogue of Equation (8)) is the sample estimate of the ex post equity premium, roughly 6%. That is, in this simplest case, when we minimize the distance between the sample moment and the simulated moment and find that the estimate of the ex ante equity premium is the ex post equity premium, we do so by construction. If the DK method is internally consistent, and if we are fitting only the ex post equity premium sample moment, then the difference must be zero at the value of π equal to the ex post equity premium. This DK estimator of π, considering only one moment of the data, would offer no advantage over the ex post equity premium, which is the traditional estimate of the ex ante equity premium. Adding a second moment to our estimation procedure, say the dividend yield, and minimizing the distance between the simulated and sample moments for the ex post equity premium and the dividend yield *jointly*, would likely lead to a somewhat different ex ante equity premium estimate. Furthermore, the estimate would be more precisely estimated (*i.e.,* with a smaller standard error) since two moments are exploited to estimate the ex ante equity premium, not just one moment, at least if the extra moment of the data provided some unique information about the value of the parameter π.

The DK method provides simulated dividend yields, ex post equity premia, and any other statistic that is derivative to returns and prices, such as return volatility, resulting in a broad collection of simulated moments with which to compare moments of the actual US data in order to derive an estimator. The large collection of available moments makes it likely that our analysis can provide a tighter bound on the value of the ex ante equity premium than has been achieved previously.

B The Simulation

To estimate the joint distribution of the financial quantities of interest, we consider models calibrated to the US economy. (We calibrate to US data over 1952 through 2004, with the starting year of 1952 motivated by the US Federal Reserve Board's adoption of a modern monetary policy regime in 1951.) We provide specific details on the nature of the models we consider and how we conduct our simulations in Appendices 1 and 2. Our entire procedure can be generally summarized in the following five steps:

Step 1: Specify assumptions about the ex ante equity premium demanded by investors. Is the premium constant or time-varying? If constant, what value does it take? If time-varying, how does the value change over time? Are there any structural breaks in the equity premium process over time? Pástor and Stambaugh (2001), among others, provide evidence that the equity premium has been trending downward over the sample period we study, finding a modest downward trend of roughly 0.80% in total since the early 1950s. Pástor and Stambaugh (2001) also find fairly strong support for there having been a structural break over the 1990s which led to a 0.5% drop in the equity premium.[5]

Once the process driving the ex ante equity premium is defined, we can specify the discount rate (which equals the risk-free rate plus the equity premium) that an investor would rationally apply to a forecasted dividend stream in order to calculate the present value of a dividend-paying stock. Note that if the equity premium varies over time, then the models generated in the next step are calibrated to mimic the degree of covariation between interest rates, dividend growth rates, and equity premia observed in the US data.

Step 2: Estimate econometric models for the time-series processes driving actual dividends and interest rates in the US economy, allowing for autocorrelation and covariation as observed in the US data. These models will later be used to Monte-Carlo simulate a variety of potential paths for US dividends and interest rates. The simulated dividend and interest rate paths are of course different in each of these simulated economies because different sequences of random innovations are applied to the common stochastic processes in each case. However, the key drivers of the simulated economies themselves are all still identical to those of the US economy since all economies share common stochastic processes fitted to US data.

Some of the models we consider assume that all cashflows received by investors come in the form of dividends (the standard assumption). Another set of models we consider embed higher cashflows and cashflow growth rates than observed in the US S&P 500 dividend data, to account for the observation of Bagwell and Shoven (1989), Fama and French (2002), and others, that dividends under-report total cashflows to shareholders. As reported by these authors, firms have been increasingly distributing cash to shareholders via share repurchases instead of via dividends, a phenomenon commonly known as disappearing dividends, a practice adopted widely beginning in the late 1970s. Fama and French find evidence that the disappearance of dividends is in part due to an increase in the inflow of new listing to US stock exchanges, representing mostly young companies

[5]A falling equity premium is thought to come from several sources, including the declining cost of diversifying through mutual funds over the last half century, the infeasibility before the advent of mutual funds to hold fully diversified portfolios (hence higher returns required by investors to hold relatively undiversified positions), and the broader pool of investors now participating in equity ownership, sharing in the market risk and presumably lowering the required rate of return to risky assets. See Siegel (1999) and Diamond (2000).

with the characteristics of firms that would not be expected to pay dividends, and in part due to a decline in the propensity of firms to pay dividends.

Thus, for some models in our simulations, we adopt higher cashflows than would be indicated by considering US dividend data alone. On a broad set of data, Grullon and Michaely (2002) find that total payouts to shareholders have remained fairly flat, not growing over the period we consider. To the extent that this is true of the S&P 500 data, the models we consider with upward-trending dividend growth are overly aggressive, but as we show below, the higher dividend growth rate only widens the range of plausible ex ante equity premia, meaning our estimate of the precision of our approach is conservative.

Step 3: Allow for the possibility of estimation error in the parameter values for the dividend growth rate, interest rate, and equity premium time-series models. That is, incorporate into the simulations uncertainty about the true parameter values. This allows for some models with more autocorrelation in the dividend growth, interest rate, and equity premium series, some with less, some with more correlation between the processes, some with less, some with a higher variance or mean of dividend growth and interest rates, some with less, and so on. This uncertainty is measured using the estimated covariance of the parameter estimates from our models generated in Steps 1 and 2, and the procedure to randomly select parameters from the estimated joint distribution of the parameters is detailed in Appendix 1. We also account for investor uncertainty about the true fundamental processes underlying prices and returns by performing tests insensitive to this uncertainty and its impact on prices and returns, as we describe below.

Further details about Steps 1 through 3 are contained in Appendix 1. Before continuing with summarizing Steps 4 and 5 of our methodology, it is worth identifying some models that emerge from various combinations of the assumptions embedded in Steps 1 through 3. The key models we consider in this paper are shown in Table I. The first column of Table I indicates numbering that we assign to the models. The second column specifies the time-series process used to generate the interest rate and dividend growth rate series, corresponding to Step 2. The next three columns relate to Step 1 above, indicating whether or not the ex ante equity premium process incorporates a downward trend over time (and if so, how much the mean ex ante equity premium in 1952 differs from the value in 2004), whether or not there is a structural break (consisting of a 50 basis point drop) in the equity premium consistent with the findings of Pástor and Stambaugh (2001), and whether or not there is a break in the dividend growth rate process, consistent with the Bagwell and Shoven (1989) and Fama and French (2002) finding of an increase in share repurchases from the late 1970s onward.[6] The last column corresponds to Step 3, showing which models incorporate uncertainty in generating parameters. We consider a selection of 12 representative models, ranging from a simple model with no breaks or trends in the equity premium process (Model 1) to very complex models.[7] Each model is fully explored in the sections that follow. We now continue describing the two final steps of our basic methodology.

Table I goes about here.

Step 4: Calculate the fundamental stock returns (and hence ex post equity premia) that arise in each simulated economy, using a discounted-dividend-growth-rate model and based on assumptions about the ex ante equity premium from Step 1, the dividend growth rate and interest rate processes specified in Step 2, and the possible parameter uncertainty specified in Step 3. The model is rolled out to produce 53 annual observations of returns, prices, dividends, interest rates, and so on, mimicking the 53 years of annual US data available to us for comparison. Keep in mind the fact that the assumptions made in Steps 1 through 3 are the same for all simulated economies in a given experiment. That is, all economies in a given experiment have the same ex ante equity premium model (for instance a constant ex ante equity premium, or perhaps an ex ante equity premium that time-varies between a starting and ending value) and yet all economies in the set of simulations have different ex post equity premia. Given the returns and ex post equity premia for each economy, as well as the means of the interest rates and dividend growth rates produced for each economy, we are able to calculate various other important characteristics, including return volatility,

[6]In each case where we consider model specifications intended to capture real-world features like breaks and trends in rates and premia, we adopt parameterizations that bias our results to be more conservative (*i.e.* to produce a wider confidence interval for the ex ante equity premium). This allows us to avoid over-stating the gains in precision possible with our technique. For example, while Pástor and Stambaugh (2001) find evidence that there was a break in the equity premium process across several years in the 1990s, we concentrate the entire break into one year (1990). Allowing the break to be spread across several years would lead to a narrower bound on the ex ante equity premium than we find. See Appendix 1 for more details.

[7]For the sake of brevity, the Gordon (1962) constant dividend growth model is excluded from the set of models we explore in this paper. We did analyze the Gordon model and found it to perform very poorly. The model itself is rejected at every value of the ex ante equity premium, even more strongly than any other simple model considered in this paper is rejected.

dividend yields, and Sharpe ratios. There is nothing in our experimental design to exclude (rational) market crashes and dramatic price reversals. Indeed our simulations do produce such movements on occasion. The details of Step 4 are provided in Appendix 2.

Step 5: Examine the distributions of variables of interest, including ex post equity premia, Sharpe ratios, dividend yields, and regression coefficients (from estimating AR(1) and ARCH models for returns) that arise conditional on various mean values and various time-series characteristics of the ex ante equity premia. Comparing the performance of the US economy with various univariate and multivariate distributions of these quantities and conducting joint hypothesis tests allows us to determine a narrow range of equity premia consistent with the US market data. That is, only a small range of mean ex ante equity premia and time-varying equity premium models could have yielded the outcome of the past half century of high mean return and return standard deviation, low dividend yield, high ex post equity premium, *etc.*

A large literature makes use of similar techniques in many asset pricing applications, directly or indirectly simulating stock prices and dividends under various assumptions to investigate price and dividend behavior.[8] However, these studies typically employ restrictions on the dividend and discount rate processes in order to obtain prices from some variant of the Gordon (1962) model and/or some log-linear approximating framework. For instance, the present value (price, defined as P_0) of an infinite stream of expected discounted future dividends can be simplified under the Gordon model as

$$P_0 = D_1/(r - g), \tag{9}$$

where D_1 is the coming dividend, r is the constant discount rate, and g is the constant dividend growth rate. That is, by assuming constant r and g, one can analytically solve for the price. If, however, discount rates or dividend growth rates are in fact conditionally time-varying, then the infinite stream of expected discounted future dividends in Equation (5) cannot be simplified into Equation (9), and it is difficult or impossible to solve prices analytically without imposing other simplifying assumptions.

[8]See, for example, Scott (1985), Kleidon (1986), West (1988a,b), Campbell (1991), Gregory and Smith (1991), Mankiw, Romer, and Shapiro (1991), Hodrick (1992), Timmermann (1993, 1995), and Campbell and Shiller (1998).

Rather than employ approximations to solve our price calculations analytically, we instead simulate the dividend growth and discount rate processes directly, and evaluate the expectation through Monte Carlo integration techniques, adopting the DK method.[9] In the setting of time-varying dividend growth rates and interest rates which conditionally covary, this technique allows us to evaluate prices, returns, and other financial quantities without approximation error.[10] We also take extra care to calibrate our models to the time-series properties of actual market data. For example, annual dividend growth is strongly autocorrelated in the S&P 500 stock market data, counter to the assumption of a logarithmic random walk for dividends sometimes employed for tractability in other applications. Furthermore, interest rates are autocorrelated and cross-correlated with dividend growth rates. Thus we incorporate these properties in our 12 models (shown in Table I), which we use to produce our simulated dividend growth rates, interest rates, and, ultimately, our estimate of the ex ante equity premium.

We estimated each of the 12 models over a grid of discrete values of the ex ante equity premium, with the grid as fine as an eighth of a percent in the vicinity of a 3.5% equity premium, and no coarser than 100 basis points for equity premium values exceeding 5%. The entire exercise was conducted using distributed computing across a grid of 30 high-end, modern-generation computers over the course of a month. On a modern stand-alone computer, estimation of a single model for a single assumed value of the ex ante equity premium would take roughly one week to estimate (and, as stated above, we consider many values of the ex ante equity premium for each of our models).

II Univariate Conditional Distributions For Model 1

All of the results in this section of the paper are based on Model 1, as defined in Table I. Model 1 incorporates interest rates that follow an AR(1) process and dividend growth rates that follow a MA(1) process. The ex ante equity premium in Model 1 follows an AR(1) process (that emerges from Merton's (1980) conditional CAPM, as detailed in Appendix 1), with no trends or breaks in either the equity premium process or dividend growth rate process. We start with this "plain

[9] The Dondaldson and Kamstra (1996) method nests other fundamental dividend-discounting valuation methods as special cases. For instance, in a Gordon (1962) world of constant dividend growth rates and interest rates, the DK method produces the Gordon model price, albeit through numerical integration rather than analytically.

[10] There is still Monte Carlo simulation error, but that is random, unlike most types of approximation error, and it can also be measured explicitly and controlled to be very small, which we do, as explained in Appendix 2.

vanilla" model because it provides a good illustration of how well dividend-discounting models that incorporate time-varying autocorrelated dividend growth and discount rate processes can produce prices and returns that fit the experience of the last half century in the US. This model also provides a good starting point to contrast with models employing breaks and trends in equity premium and dividend growth processes. We consider more complex and arguably more realistic models incorporating trends and breaks later in the paper.

It is well known that the ex ante equity premium is estimated with error. See, for instance, Merton (1980), Gregory and Smith (1991), and Fama and French (1997). Any particular realization of the equity premium is drawn from a distribution, implying that given key information about the distribution (such as its mean and standard deviation), one can construct a confidence interval of statistically similar values and determine whether a particular estimate is outside the confidence interval. As mentioned above, an implication of this estimation error is that most studies have produced imprecise estimates of the mean equity premium. For instance, a typical study might yield an 800 basis point 95% confidence interval around the ex ante equity premium.[11] Studies including Fama and French (2002) have introduced innovations that make it possible to narrow the range. One of our goals is to further sharpen the estimate of the mean ex ante equity premium.

We first consider what we can learn by looking at the univariate statistics that emerge from our simulations. We can use the univariate distributions to place loose bounds on plausible values of the mean ex ante equity premium. While the analysis in this section based on univariate empirical distributions is somewhat casual, in Section III we conduct formal analysis based on χ^2 statistics and the joint distributions of the data, yielding very tight bounds on plausible values of the mean ex ante equity premium and identifying plausible models of the equity premium process, representing our main contributions.

Consider the following: conditional on a particular value of the ex ante equity premium, how unusual is an observed realization of the ex post equity premium? How unusual is an observed realization of the mean dividend yield? Each simulated economy produces a set of financial statistics based on the simulated *annual* time-series observations, and these financial statistics can be

[11] This particular range is based on the simple difference between mean realized equity returns and the average riskfree rate based on the last 130 years of data, as summarized in Table I of Fama and French (2002).

compared and contrasted with the US experience of the last half century. By considering not only the mean of a financial statistic across simulated economies, such as the mean ex post equity premium, but also conditional moments and higher moments including the standard deviation of excess returns produced in our simulations, we can determine with high refinement the ability of our simulated data to match characteristics of the US economy. For instance, market returns, to be discussed below, are volatile. Thus it is interesting to examine the degree to which our simulations are able to produce volatile returns and to look at the distribution of return variance as we vary the mean ex ante equity premium in our simulated economies.

We can compare any financial statistic from the last half century to our simulated economies provided the statistic is based on returns or dividends or prices, as these are data that the simulation produces. We could also consider moments based on interest rates or dividend growth rates, but since we calibrate our models to interest rates and dividend growth rates, all our simulations should (and do) fit these moments well by construction. We choose moments based on two considerations. First, the moments should be familiar and the significance of the moments to economic theory should be obvious. Second, the moments should be precisely estimated; if the moments are too "noisy," they will not help us narrow the range of ex ante equity premia. For instance, return skew and kurtosis are very imprecisely estimated with even 50 years of data, so that these moments are largely uninformative. The moments must also be well-defined; moments must be finite, for instance. The expected value of the price of equity is undefined, but we can use prices in concert with a cointegrated variable like lagged price (to form returns) or dividends (to form dividend yields).

Rather than presenting copious volumes of tabled results, we summarize the simulation results with concise plots of probability distributions of the simulated data for various interesting financial statistics. This permits us to determine if a particular ex ante equity premium produces financial statistics similar to what has been seen over the last half century in the US.

Figure 1 contains four panels, and in each panel we present the probability distribution function for one of various financial statistics (ex post equity premia, dividend yield, Sharpe ratio, and return volatility) based on each of four different ex ante equity premium settings. We also indicate the realized value for the actual US data. Comparison of the simulated distribution with realized values in these plots permits a very quick, if casual, first assessment of how well the realized US data agree with the simulated data, and which assumed values of the ex ante equity premium appear inconsistent with the experience of the last half century of US data.

Panels A through D of Figure 1 contain probability distribution functions (PDFs) corresponding to the mean ex post equity premium, the mean dividend yield, the Sharpe ratio, and return volatility respectively, based on assumed mean ex ante equity premia of 2.75%, 3.75%, 5%, and 8%. For the sake of clarity, the dotted lines depicting the PDFs in Figure 1 are thinnest for the 2.75% case and become progressively thicker for the 3.75%, 5%, and 8% cases. The actual US realized data is denoted in each panel with a solid vertical line.

The actual US mean equity premium, displayed in Panel A, is furthest in the right tail of the distribution corresponding to a 2.75% ex ante equity premium, and furthest in the left tail for the ex ante premium of 8%. The wide range of the distribution of the mean ex post equity premia for each assumed value of the ex ante equity premium is consistent with the experience of the last half century in the US, in which the mean ex post equity premium has a 95% confidence interval spanning plus or minus roughly 4% or 5%. The actual dividend yield of 3.4%, displayed in Panel B, is unusually low for the 5% and 8% ex ante equity premium cases, but it is near the center of the distribution for the ex ante premium values of 2.75% and 3.75%. In Panel C, only the Sharpe ratios generated with an ex ante equity premium of 8% appear inconsistent with the US experience of the last half century. The return volatility, displayed in Panel D, clearly indicates that the experience of the US over the last half century is somewhat unusual for all ex ante equity premia considered, though least unusual for the lowest ex ante equity premium. Casual observation, based on only the evidence in these univariate plots, implies that the ex ante equity premium which could have generated the actual high ex post equity premium and low dividend yield of the last half century of the US experience likely lies above 2.75% and below 5%.

Figure 1 goes about here.

We constructed similar plots for the mean return and for conditional moments, including the return first order autocorrelation coefficient estimate (the OLS parameter estimate from regressing returns on lagged returns and a constant, *i.e.,* the AR(1) coefficient), the return first order autoregressive conditional heteroskedasticity coefficient estimate (the OLS parameter estimate from regressing squared residuals on lagged squared residuals and a constant, *i.e.,* the ARCH(1) coefficient), and the price-dividend ratio's first order autocorrelation coefficient estimate (the OLS parameter estimate from regressing the price-dividend ratio on the lagged price-dividend ratio and a constant). The mean return distributions are similar to the ex post equity premium distributions shown in Figure 1, and all choices of the ex ante equity premium produce returns and price-dividend ratios that have conditional time-series properties matching the US data, so these results are not presented here.

Figure 1 has two central implications of interest to us. First, the financial variable statistics produced in our simulations are broadly consistent with what has been observed in the US economy over the past five decades. Most simulated statistics match the magnitudes of financial quantities from the actual US data, even though we do not calibrate to prices or returns.[12] Second, the results suggest that the 2.75% through 8% interval we present here likely contains the ex ante equity premium consistent with the US economy. Univariate results for Models 2 through 10 are qualitatively very similar to those presented for Model 1. Univariate results for Models 11 and 12, in contrast, are grossly rejected by the experience of the US economy. Detailed univariate results for Models 2 through 12 are omitted for the sake of brevity, but the poor performance of Models 11 and 12 will be evident in multivariate results reported below.

To narrow further the range of plausible ex ante equity premium values, we need to exploit the full power of our simulation procedure by considering the *joint distributions* of statistics that arise in our simulations and comparing them to empirical moments of the observed data. We consider the multivariate distributions of several moments of the data, including ex post equity premia, dividend yields, and return volatility. This exercise allows for inference that is not feasible with the univariate analysis conducted above, and it leads to a very precise estimate of the ex ante equity premium. We turn to this task in the next section, where we also broaden the class of models we consider.

[12]This in itself is noteworthy, as analytically tractable models, such as the Gordon (1962) growth model, typically imply constant or near-constant dividend yields and very little return volatility. In contrast, dividend yields observed in practice vary considerably over time and are strongly autocorrelated, and returns exhibit considerable volatility.

III Model Extensions, Multivariate Analysis, and Tests

The central focus in this section is on *joint* distributions of the financial statistics that emerge from our simulations: combinations of the returns, ex post equity premia, Sharpe ratios, dividend yields, *etc.*, and tests on the value of the ex ante equity premium using these joint distributions. We focus primarily on three moments of the data: the mean ex post equity premium, the excess return volatility, and the mean dividend yield. These three moments have the advantage of being the most precisely estimated and hence most informative for the value of the ex ante equity premium. Other moments that we could have considered are either largely redundant (such as the Sharpe Ratio which is a direct function of excess returns and the excess return standard deviation), or are so imprecisely estimated (for example, the ARCH(1) or AR(1) coefficients) that they would not help sharpen our estimates of the ex ante equity premium. Of course, we also do not consider the distributions of financial variables to which we calibrate our simulations (interest rates and dividend growth rates), as the simulated mean, variance, and covariance of these variables are, by construction, identical to the corresponding moments of the actual data to which we calibrate.

Our purpose in considering joint distributions is two-fold. First, multivariate tests are used to form a tight confidence bound on the true value of the ex ante equity premium. These tests strongly reject our models if the ex ante equity premium is outside of a narrow range around 3.5%. This range is not sensitive to even fairly substantial changes in the model specification, which suggests that the 3.5% finding is robust. Second, this analysis leads us to reject model specifications that fail to incorporate certain features, such as trends and breaks in the equity premium. Interestingly, even when a model specification is rejected, we find the most plausible ex ante equity premium still lies in the same range as the rest of our models, very near 3.5%.

Up to this point we have considered detailed results for Model 1 exclusively. The Model 1 simulation incorporates some appealing basic features, such as parameter uncertainty and calibrated time-series models for equity premia, interest rates, and dividend growth rates. It does not, however, incorporate some features of the equity premium process that have been indicated by other researchers. One omitted feature is a gradual downward trend in the equity premium, as documented in many studies, including Jagannathan, McGrattan, and Scherbina (2000), Pástor and

Stambaugh (2001), Bansal and Lundblad (2002), and Fama and French (2002). Another is a structural break in the equity premium process over the early 1990s, as shown by Pástor and Stambaugh (2001). An increase in the growth rate of cashflows (but not dividends) to investors starting in the late 1970s, as documented by Bagwell and Shoven (1989), Fama and French (2001) and others, is also a feature that Model 1 fails to incorporate. Therefore, in this section we consider models which incorporate one, two, or all three of these features, as well as different time-series models for interest rates and equity premia. We also consider stripped-down models to assess the marginal contribution of model features such as parameter uncertainty and the specification of the time-series process used to model dividend growth rates and interest rates.

In Figures 2 through 8 (to be fully discussed below), we present χ^2 test statistics for the null hypothesis that the US experience during 1952 through 2004 could have been a random draw from the simulated distribution of the mean ex post equity premium, the excess return volatility, and the mean dividend yield.[13]

A significant test statistic, in this context, suggests that the combination of financial statistics observed for the US economy is significantly unusual compared to the collection of simulated data, leading us to reject the null hypothesis that the given model and assumed ex ante equity premium value could have generated the US data of the last half century. It is possible to reject every ex ante equity premium value if we use models of the equity premium that are misspecified (the rejection of the null hypothesis can be interpreted as a rejection of the model). It is also possible that a very wide range of ex ante equity premium values are not rejected for a collection of models, thwarting our efforts to provide a precise estimate of the ex ante equity premium or a small range of allowable equity premium models.

As it happens, models that ignore breaks and trends in the equity premium are rejected for

[13]The χ^2 tests are based on joint normality of sample estimates of moments of the simulated data, which follow an asymptotic normal distribution based on a law of large numbers (see White, 1984, for details). In the case of the excess return volatility, we consider the *cube root* of the return variance, which is approximately normally distributed (see page 399 of Kendall and Stuart, 1977, for further details). We also estimate the probability of rejection using bootstrapped p-values, to guard against deviations from normality. These bootstrapped values are qualitatively identical to the asymptotic distribution p-values. Finally, when performing tests that include the dividend yield moment, if the simulation includes a break in dividends corresponding to an increase in cash payouts starting in 1978 in the US data (again, see Fama and French, 2001), we also adjust the US data to reflect the increase in mean payout levels. This makes for a small difference in the mean US payout ratio and no qualitative change to our results if ignored.

virtually every value of the ex ante equity premium we consider. But for a group of sophisticated models that incorporate trends and breaks in the equity premium, we cannot reject a narrow range of ex ante equity premia, roughly between 3% and 4%. We also find that models tend to be rejected if the impact on cashflows to shareholders from share repurchases are ignored. We begin with some simple models, then consider models that are arguably more realistic as they incorporate equity premium and cashflow trends and breaks, and finish by considering a host of related issues, including the impact of parameter estimation error and, separately, *investor* uncertainty about the fundamental value of equities.

A Simple (One-at-a-Time) Model Extensions

We now consider extensions to Model 1, each extension adding a single feature to the base model. Recall that the features of each model are summarized in Table I. For Model 2, an 80 basis point downward trend is incorporated in the equity premium process. For Model 3, a 50 basis point drop in year 39 of the simulation (corresponding to 1990 for the S&P 500 data) is incorporated in the equity premium process. For Model 4, the dividend growth rate process is shifted gradually upward a total of 100 basis points, starting in year 27 of the simulation (corresponding to 1978 for the S&P 500 data) and continuing for 20 years at a rate of 5 basis points per year. These one-at-a-time feature additions help us evaluate if one or another feature documented in the literature can markedly improve model performance over the simple base model.

Panel A of Figure 2 and Panel A of Figure 3 display plots of the value of joint χ^2 tests on three moments of the data, the mean ex post equity premium, the excess return volatility, and the mean dividend yield, for Models 1 though 4, and shows how the test statistic varies as the ex ante equity premium varies from 2.25% to 8% in increments as small as an eighth of a percent toward the lower end of that range. Panels B through D of Figures 2 and 3 display the univariate Student t-test statistics for each of these three moments of the data, again showing how the test statistic varies with the assumed value of the ex ante equity premium. The values of the ex ante equity premia indicated on the horizontal axis represent the *ending* values of the ex ante equity premium in each set of simulations. For models which incorporate a downward trend or a structural break in the equity premium, the ending value of the ex ante equity premium differs from the starting value.

So, for instance, Model 2 has a starting ex ante equity premium that is 80 basis points higher than that displayed in Figure 2, as Model 2 has an 80 basis point trend downward in the ex ante equity premium. For Model 1 the value of the ex ante equity premium is the same at the *end* of the 53-year simulation period as it is at the *start* of the 53-year period, as Model 1 does not incorporate a downward trend or structural break in the equity premium process. Critical values of the test statistics corresponding to statistical significance at the 10%, 5%, and 1% levels are indicated by thin dotted horizontal lines in each panel, with the lowest line indicating significance at the 10% level and the highest line the 1% significance level.

Figures 2 and 3 go about here.

Consider now specifically Panel A of Figures 2 and 3. (Note that we use a log scale for the vertical axis of the plots in Panel A of Figures 2 through 8 for clarity of presentation. Note as well that we postpone further discussion of Panels B through D until after we have introduced results for all the models, 1 through 12.) On the basis of Panel A of Figures 2 and 3, we see that only in the case of Model 4 do we observe χ^2 test statistics lower than the cutoff value implied by a 10% significance level (again, indicated by the lowest horizontal dotted line in the plot). The test statistics dip (barely) below the 10% cutoff line only for values of the ex ante equity premium within about 25 basis points of 4%. Models 1-3, in contrast, are rejected at the 10% level for *every* ex ante equity premium value. If we allow fairly substantial departures of the S&P 500 data from the expected distribution, say test statistics that are unusual at the 1% level of significance (the upper horizontal dotted line in the plot), then all the models indicate ranges of equity premia that are not rejected, in each case centered roughly between 3.5% and 4%. Recall that the equity premium plotted is the *ending* value, so if the model has a downward trend or decline because of a break in the equity premium, its ending value is below its average ex ante equity premium.

One conclusion to draw from the relative performance of these four competing models is that each additional feature over the base model, the dividend growth acceleration in the late 1970s and the trends and breaks in the equity premium, lead to better performance relative to the base model, but each in isolation is still inadequate. The model most easily rejected is clearly that which does not account for trends and breaks in the equity premium and cashflow processes.

B Further Model Extensions (Two or More at a Time)

We turn now to joint tests based on Models 5 though 10. These models incorporate the basic features of Model 1, including time-varying and dependent dividend growth and interest rates, parameter uncertainty, and, with the exception of Model 10, an equity premium process derived from the Merton (1980) conditional CAPM (detailed in Appendix 1). These models also permit trends and/or breaks in the equity premium and dividend growth rate processes two or more at-a-time and incorporate alternative time-series models for the interest rate and the equity premium processes. Models 1 through 4 demonstrate that it is not sufficient to model the equity premium as an autoregressive time-varying process, and that one-at-a-time augmentation with trends or breaks in the equity premium process is also not sufficient, though the augmentations do lead to improvements over the base model in our ability to match sample moments from the US experience of the last half century. Models 5 through 10 allow us to explore questions like: do we need a conditionally time-varying equity premium model built on the Merton conditional CAPM model, or is it sufficient to have an equity premium that simply trends downward with a break? If we have a break, a trend, and time-variation in the equity premium process, is it still essential to account for the disappearing dividends of the last 25 years? Are our results sensitive to the time-series model specifications we employ in our base model?

Model 5 is the base model, Model 1, augmented to include an 80 basis point gradual downward trend in the equity premium and a 100 basis point gradual upward trend in the dividend growth rate. Model 6 is the base model adjusted to incorporate a 30 basis point gradual downward trend in the equity premium, a 50 basis point abrupt decline in the equity premium, and a 100 basis point gradual upward trend in the dividend growth rate. Model 7 is the best model as indicated by the Bayesian Information Criterion (BIC),[14] augmenting the equity premium process with a 30 basis point gradual downward trend and a 50 basis point abrupt decline and adding a 100 basis point gradual upward trend in the dividend growth rate. Model 8 takes the second-best BIC model

[14]For Models 7 and 8 we employ the BIC to select the order of the ARMA model driving each of the interest rate, equity premium, and dividend growth rate processes. The order of each AR process and each MA process for each series is chosen over a (0, 1, 2) grid. The BIC has been shown by Hannan (1980) to provide consistent estimation of the order of linear ARMA models. We employ the BIC instead of alternative criteria because it delivers relatively parsimonious specifications and because it is widely used in the literature (*e.g.*, Nelson, 1991, uses the BIC to select EGARCH models).

and incorporates a 30 basis point gradual downward trend in the equity premium, a 50 basis point abrupt decline in the equity premium, and a 100 basis point gradual upward trend in the dividend growth rate. Model 9 is the base model adjusted to incorporate a 30 basis point gradual downward trend in the equity premium and a 50 basis point abrupt decline in the equity premium. Model 10 has the equity premium model following a deterministic downward trend with a 50 basis point structural break, interest rates following an AR(1), and dividend growth rates following an MA(1).

Given the existing evidence in support of a gradual downward trend in the equity premium, a structural break in the equity premium process over the early 1990s, and an increase in the growth rate of non-dividend cashflows to investors (such as share repurchases) starting in the late 1970s, we believe Models 6, 7, and 8 to be the best calibrated and therefore perhaps the most plausible among all the models we consider, and Model 5 to be a close alternative.

In Panel A of Figures 4, 5, and 6 we present plots of the χ^2 test statistics on three moments of the data, the mean ex post equity premium, the excess return volatility, and the mean dividend yield. Again, we consider Panels B through D later. We see in Panel A of Figures 4 and 5 that for Models 5 through 8 we cannot reject a range of ex ante equity premium values at the 5% level. These models produce test statistics that drop well below even the 10% critical value (recall that Panel A's scale is logarithmic, and thus compressed). These models all embed the increased cashflow feature and either an eighty basis point downward trend in the equity premium, or both a break and a trend in the equity premium, adding to an eighty basis point decline over the last half century. The range of ex ante equity premia supported (not rejected) is narrowest for Model 7 (the best model indicated by BIC) and Model 8 (the second best model indicated by BIC) with a range less than 75 basis points at the 10% level. The range is slightly wider for Models 5 and 6, roughly 75 to 100 basis points. In each case, the ex ante equity premium that yields the minimum joint test statistic, corresponding to our estimate of π, is centered between 3.25% and 3.75%.

For the models which exclude the cashflow increase, Models 9 and 10, displayed in Figure 6, we see that we can reject at the 10% level all ex ante equity premium values. Model 9 is best compared to Model 6, as it is equivalent to Model 6 with the sole difference of excluding the cashflow increase. We see from Panel A of Figures 4 and 6 that excluding the cashflow increase flattens the trough of the plot of χ^2 statistics, and approximately doubles the test statistic value, from a little over 3 for

Model 6 in Figure 4 to a little over 6 for Model 9 in Figure 6 (recall that the scale is compressed in Panel A as we use a log scale). Model 10 is identical to Model 9 apart from the sole difference that Model 10 excludes the Merton CAPM conditionally-varying equity premium process. Exclusion of this conditional time variation (modeled as a first order autoregressive process) worsens the ability of the model to match moments to the US experience at every value of the ex ante equity premium. The difference in performance leads us to reject a model excluding a conditionally-varying equity premium.

Figures 4, 5, and 6 go about here.

On the basis of our most plausible models, Models 6, 7, and 8, we can conservatively conclude that the ex ante equity premium is within 50 basis points of 3.5%. We can also conclude that models that allow for breaks and/or trends in the equity premium process are the only models that are not rejected by the data. Simple equity premium processes, those that rule out any one of a downward break and/or trend or a Merton (1980) CAPM conditionally-varying equity premium process, cannot easily account for the observed low dividend yields, high returns, and high return volatility. Ignoring the impact of share repurchases on cashflows to investors over the last 25 years also compromises our ability to match the experience of US prices and returns of the last half century.

C Is Sampling Variability (Uncertainty) in Generating Parameters Important?

All of the models we have considered so far, Models 1-10, incorporate parameter value uncertainty. This uncertainty is measured using the estimated covariance of the parameter estimates from our models. We generate model parameters by randomly drawing values from the joint distribution of the parameters, exploiting the asymptotic result that our full information maximum likelihood procedure produces parameter estimates that are jointly normally distributed, with an easily computed variance-covariance structure.

Now we consider two models that have no parameter sampling variability built into them, Models 11 and 12. In these models the point estimates from our ARMA estimation on the S&P 500 data are used for each and every simulation. Ignoring uncertainty about the true values for the parameters of the ARMA processes for interest rates, dividend growth rates, and the equity premium should dampen the variability of the generated financial statistics from these simulations, and potentially understate the range of ex ante equity premia supported by the last half century of US data. Model 11 is the base model augmented to incorporate a 30 basis point gradual downward trend in the equity premium, a 50 basis point abrupt decline in the equity premium, and a 100 basis point gradual upward trend in the dividend growth rate, with no parameter uncertainty. (Model 11 is identical to Model 6 apart from ignoring parameter uncertainty.) Model 12 is the base model, Model 1, with no parameter uncertainty.

Figure 7 goes about here.

In Panel A of Figure 7 we present plots of the χ^2 test statistics on three moments of the data, the mean ex post equity premium, the excess return volatility, and the mean dividend yield. Again, we consider Panels B through D later. We see in Panel A that both Models 11 and 12 are rejected for all values of the ex ante equity premium, though Model 11, which allows for trends and breaks, performs better than Model 12. The log scale for the vertical axis compresses the values, but the minimum χ^2 statistic for Model 12 is close to 30, indicating very strong rejection of the model, while the minimum χ^2 statistic for Model 11 is roughly 10. In each case, the ex ante equity premium that yields the minimum joint test statistic, corresponding to our estimate of π, is centered around 3%. It is apparent that parameter uncertainty is an important model feature. Ignoring parameter uncertainty leads to model rejection, even at the ex ante equity premium setting that corresponds to the minimum test statistic.

D The Moments That Matter

An interesting question that arises with regard to the joint tests is, where does the test power come from? That is, which variables give us the power to reject certain ranges of the ex ante equity premium in our joint χ^2 tests? An examination of the ranges of the ex ante equity premium consistent with the *individual* moments can shed some light on the source of the power of the joint tests. Panels B, C, and D of Figures 2 through 7 display plots of the univariate t-test statistics based on each of the variables we consider in the joint tests plotted in Panel A of these figures. Panel B of each figure plots t-test statistics on the ex post equity premium, Panel C of each figure plots t-test statistics on the excess return volatility, and Panel D of each figure plots t-test statistics on the price-dividend ratio.

Consider first Panel B of Figures 2 through 7. Virtually all of the models have a minimum t-test statistic at a point that is associated with an ex ante equity premium close to 6%.[15] Because our method involves minimizing the distance between the ex post equity premium based on the actual S&P 500 value (which is a little over 6%) and the ex post equity premium estimate based on the simulated data, it is not surprising that the minimum distance is achieved for models when they are set to have an ex ante equity premium close to 6%. The t-test on the mean ex post equity premium rises linearly as the ex ante equity premium setting departs from 6% for each model, but does not typically reject ex ante equity premium values at the 10% level until they deviate quite far from the ex ante value at which the minimum t-test is observed. For example, in Panel B of Figure 4 the ending ex ante equity premium must be as low as 2.25% or as high as 7% before we see a rejection at the 10% level. This wide range reflects the imprecision of the estimate of the ex post equity premium which is also evident in the actual S&P 500 data.

The t-tests on the excess return volatility, presented in Panel C of Figures 2 through 7, indicate that lower ex ante equity premium values lead to models that are better able to match the S&P 500 experience of volatile returns.[16] Note that as the ex ante equity premium decreases, the volatility of returns *increases*, so high ex ante equity premia lead to simulated return volatilities that are much lower than the actual S&P 500 return volatility we have witnessed over the last half century. The test statistic, however, rises slowly as the ex ante equity premium grows larger, in contrast to the joint test statistics plotted in Panel A of Figures 2 through 7, in which the χ^2 test statistic

[15] Recall that the ex ante equity premium values shown on the horizontal axes are *ending* values, so if the model has a downward trend or break in the equity premium process, its ending value is below the mean equity premium. For instance, Model 11 has a data generating process that incorporates trends and breaks that lead to an ending equity premium lower than the starting value. Accordingly, for this model we observe (in Panel B of Figure 7) a minimum t-test at an *ending* value of the ex ante equity premium which is below the 6% *average* equity premium. The coarseness of the grid of ex ante equity premium values around 6% prevents this feature from being more obvious for some of the other models.

[16] The intuition behind this result is easiest to see by making reference to the Gordon (1962) constant dividend growth model, shown above in Equation 9. As the discount rate, r, declines in magnitude, the Gordon price increases. The variable r equals the risk-free rate plus the equity premium in our simulations, so low values of the equity premium lead to values of the discount rate that are closer to the dividend growth rate, resulting in higher prices. When the value of the equity premium is low, small increases in the dividend growth rate or small decreases in the risk-free rate lead to large changes in the Gordon price. In our simulations (where the conditional mean dividend growth rate and conditional mean risk-free rate change over time), when the value of the equity premium is low, small changes in the conditional means of dividend growth rates or risk-free rates also lead to large prices changes, *i.e.* volatility.

rises *sharply* as the ex ante equity premium grows larger (recall that the Panel A vertical axis has a compressed log scale in Figures 2 through 7). Given these contrasting patterns, the return volatility moment is unlikely, *by itself,* to be causing the sharply rising joint test statistic.

Consider now the t-test statistics on the price-dividend ratio, plotted in Panel D of Figures 2 through 7. Notice that in all cases the t-test on the price-dividend ratio jumps up sharply as the ex ante equity premium rises above 3%. Thus the sharply increasing χ^2 statistics we saw in Panel A of the three figures are likely due in large part to information contained in the price-dividend ratio. However, return volatility reinforces and amplifies the sharp rejection of premia above 4% that the dividend yield also leads us to. In terms of the three moments we have considered in the joint χ^2 and univariate t-test statistics, it is evident that the upper range of ex ante equity premia consistent with the experience of the last half century in the US is limited by the high average S&P 500 price-dividend ratio (or equivalently, the low average S&P 500 dividend yield) together with the high volatility of returns. This result is invariant to the way we model dividend growth, interest rates, or the equity premium process. Even an ex ante equity premium of 5% produces economies with price-dividend ratios and return volatilities so low that they are greatly at odds with the high return volatility and high average price-dividend ratio observed over the past half century in the US.

D.1 Sensitivity to Declining Dividends Through Use of the Price-Dividend Ratio

To ensure that our results are not driven by a single moment of the data, in particular a moment of the data possibly impacted by declining dividend payments in the US, we perform two checks. First, in Models 4 through 8 we incorporate higher dividends and dividend growth rates than observed in US corporate dividends. This is to adjust for the practice, adopted widely beginning in the late 1970s, of US firms delivering cashflows to investors in ways (such as share repurchases) which are not recorded as corporate dividends. As we previously reported, Models 4 through 8 (the models that incorporate higher cashflows to investors than recorded by S&P 500 dividend payments, *i.e.*, the models that use cashflows *including* share repurchases) are best able to account for the observed US data. Reassuringly, the estimate of the equity premium emerging from Models 4 through 8 is virtually identical to that produced by the models that exclude share repurchases.

Our second check is to perform joint tests excluding the price-dividend ratio. Any sensitivity to mismeasurement of the price-dividend ratio should be mitigated if we consider joint test statistics that are based only the ex post equity premium and return volatility, excluding the price-dividend ratio. These (unreported) joint tests confirm two facts. First, when the joint tests exclude the price-dividend ratios, the value of the χ^2 statistic rises less sharply for values of the ex ante equity premium above 4%. Essentially, this indicates that using two moments of the data (excluding the price-dividend ratio) rather than all three makes it more difficult to identify the minimum test statistic value and thus more difficult to identify our estimate of the ex ante equity premium. This confirms our earlier intuition that the price-dividend ratio is instrumental in determining the steep rise of the joint test statistic in Panel A of Figures 2 through 7. Second, and most importantly, the minimum test statistic is still typically achieved for models with an ex ante equity premium value between 3% and 4%. For some of the models, the minimum test statistic is 25 or 50 basis points lower than that found when basing joint tests on the full set of three moments. For a few models, the minimum test statistic is 25 or 50 basis points higher. Again Models 1 through 3 are rejected for every value of the ex ante equity premium, and again for Models 4 through 8 the range of ex ante equity premia that are not rejected is narrow.

E Investors' Model Uncertainty

We have been careful to explore the impact of estimation uncertainty by simulating from the sampling distribution of our model parameters, and to explore the impact of model specification choice (and implicitly model *mis*specification) by looking at a variety of models for interest rates, dividend growth rates, and equity premium, ranging from constant rate models to various ARMA specifications, with and without trends and breaks in the equity premium and dividend growth rates. Comparing distributions of financial statistics emerging from this range of models to the outcome observed in the US over the last half century leads us to the conclusion that the range of true ex ante equity premia that could have generated the US experience is fairly narrow, under 100 basis points, centered roughly on 3.5%. We have not yet addressed, however, the impact of *investor* uncertainty regarding the true fundamental value of the assets being priced. Up to this point, all simulated prices and returns have been generated with knowledge of the (fundamental) processes generating interest rates and dividends.

It is impossible to be definitive in resolving the impact of investor uncertainty on prices and returns. To do so we would have to know what (incorrect) model of fundamental valuation investors are actually using. We can nonetheless focus our attention on procedures likely to be less affected by investor uncertainty than others. Up to this point, the joint tests we have used to identify the plausible range of ex ante equity premia have employed the observed return volatility over the last half century in the US and the volatility of returns produced in our simulated economies. However, investor uncertainty could cause market prices to over- and under-shoot fundamental prices, impacting return volatility, perhaps significantly. A joint test statistic based on only the mean equity premium and the mean price-dividend ratio, however, should be relatively immune to the impact of investor uncertainty. (In the absence of extended price bubbles, *mean* yields should not be impacted greatly by temporary pricing errors.) Thus we now consider the joint χ^2 test statistic based on only the mean return and the mean price-dividend ratio. Figure 8, Panel A plots the test statistics for Models 1, 2, and 3, Panel B plots the test statistics for Models 4, 5, and 6, Panel C plots the test statistics for Models 7, 8, and 9, and Panel D plots the test statistics for Models 10, 11, and 12, with a log scale for the vertical axis in all cases.

Figure 8 goes about here.

First consider results for Models 1 through 4, shown in Panels A and B of Figure 8. These are the base model with no trends or breaks, and models which incorporate only one feature (trend or break in the equity premium or dividend growth rate) at a time. We see again that Model 1 is rejected outright for every value of the ex ante equity premium, at the 10% level of significance, and we see again that adding trends or breaks, even one-at-a-time, improves performance. Now Model 2 (incorporating an 80 basis point downward trend in the equity premium) and Model 4 (incorporating the increased cashflow growth rate) are not rejected over narrow ranges at the 10% significance level. We find that Models 5, 6, 7, and 8, all incorporating trends and breaks in the equity premium and dividend growth rate processes and shown in Panels B and C of Figure 8, deliver a wide range of ex ante equity premia which cannot be rejected at any conventional level of statistical significance. We also see that Model 9 in Panel C, incorporating a trend (of 30 basis points) and a break (of 50 basis points) in the equity premium, performs similarly to Model 2, which has only a trend of 80 basis points (neither model incorporates a cashflow change). In Panel D we see Model 10 which has a deterministic equity premium with trends and breaks. This model's performance is also similar to Model 2, but slightly worse, rejected at the 10% level at every ex ante equity premium. Also in Panel D we see that Models 11 and 12, which do not incorporate parameter estimation uncertainty, are almost everywhere rejected. (In contrast to the joint test shown in Panel A of Figure 7, based on all three moments, we find that Model 11 is not rejected only for the 3% value of the ex ante equity premium.)

Overall, the value of the ex ante equity premium at which the joint test statistic is minimized (*i.e.*, our estimate of the ex ante equity premium) is not particularly affected by our having based the joint tests on two moments of the data rather than the original three, nor is our selection of plausible models for the equity premium process. Across the models, the highest estimate of the ex ante equity premium is roughly 4% (for Model 4) and the lowest is 3% (for Models 11 and 12). With the joint tests based on two moments, all models support (*i.e.*, do not reject) broader ranges of the ex ante equity premium, with the range widest for Models 4 through 8 (now spanning roughly 200 basis points for any given model, from ex ante equity premium values as low as 2.25% for Model 7 to values as high as 4.5% for Model 4). This widening of the range of plausible ex ante equity premia is consistent with a decline in the power of our joint test, presumably from omitting an important moment of the data, the return volatility. The widening of the range of plausible ex ante equity premia is also consistent with investors being uncertain about the true fundamental value of the assets being priced. The last half century of data from the US will be less informative as investor uncertainty about the processes governing fundamentals exaggerates the volatility of returns and hence reduces the precision of estimates of the ex ante equity premium.

To the extent that market prices are set in an efficient market dominated by participants with models of dividend growth rates and interest rates that reflect reality, these ranges of plausible ex ante equity premia based on only the two-moment joint test are overly wide. Still these ranges are useful for putting a loose bound on the likely range of the ex ante equity premium.

F Bootstrapped Test Statistics

Up to this point, all of our test statistics have relied on asymptotic distribution theory for critical values. The asymptotic distributions should be reliable both because we are looking at averages over independent events (our simulations are by construction independent) and because we have many simulations over which to average (2,000). Nonetheless, it is straightforward to use our simulated test statistics to bootstrap the distribution of the test statistics, thus we do so. While use of the bootstrap produces small quantitative changes to our results, our main findings remain unchanged. The best estimate of the mean ex ante equity premium and the range of plausible ex ante equity premia and equity premium models do not budge.

IV Conclusions

The equity premium of interest in theoretical models is the extra return investors anticipate when purchasing risky stock instead of risk-free debt. Unfortunately, we do not observe this ex ante equity premium in the data. We only observe the returns that investors actually receive ex post, after they purchase the stock and hold it over some period of time during which random economic shocks impact prices. US stocks have historically returned roughly 6% more than risk-free debt. Ex post estimates provided by recent papers suggest the US equity premium may be falling in recent years. However, all of these estimates are imprecise, and there is little consensus emerging about the true value of the ex ante equity premium. The imprecision and lack of consensus both hamper efforts to use equity premium estimates in practice, for instance to conduct valuation or to perform capital budgeting. The imprecision of equity premium estimates also complicates resolution of the equity premium puzzle and makes it difficult to determine if the equity premium changes over time.

In order to determine the most plausible value of the ex ante equity premium and the most plausible restrictions on how the equity premium evolves over time, we have exploited information not just on the ex post equity premium and the precision of this estimate, but also on related financial statistics that define the era in which this ex post equity premium was estimated. The idea of looking at related fundamental information in order to improve the estimate of the mean ex ante equity premium follows recent work on the equity premium which has also sought improvements through the use fundamental information like the dividend and earnings yields (Fama and French, 2002, and Jagannathan, McGrattan, and Scherbina, 2000), higher-order moments of the excess return distribution (Maheu and McCurdy, 2007) and return volatility and price movement directions (Pástor and Stambaugh, 2001).

Our central insight is that the knowledge that a low dividend yield, high ex post equity premium, high return volatility, and high Sharpe ratio all occurred together over the last five decades tells us something about the mean ex ante equity premium and the likelihood that the equity premium is time-varying with trends and breaks. Certainly, if sets of these financial statistics are considered together, we should be able to estimate the equity premium more accurately than if we were to look only at the ex post equity premium. This insight relies on the imposition of some structure from economic models, but our result is quite robust to a wide range of model structures, lending confidence to our conclusions.

We employ the simulated method of moments technique and build on the dividend discounting method of fundamental valuation of Donaldson and Kamstra (1996) to estimate the ex ante equity premium. We reject as inconsistent with the US experience all but a narrow range of values of the mean ex ante equity premium and all but a small number equity premium time-series models. We do so while incorporating model estimation uncertainty and allowing for investor uncertainty about the true state of the world. The range of ex ante equity premia that is most plausible is centered very close to 3.5% for virtually every model we consider. The models of the equity premium not rejected by our model specification tests – that is, consistent with the experience of the US over the last half century – incorporate substantial autocorrelation, a structural break, and/or a gradual downward trend in the equity premium process. For these models, the range of ex ante equity premia supported by our tests is very narrow, plus or minus 50 basis points around 3.5%. All together, our tests strongly support the notion that the equity premium process over the last half century in the US was very unlikely to have been constant, was likely to have demonstrated at least one sharp downward break, and was likely to have demonstrated a gradual downward trend.

References

Bagwell, L., Shoven, J., 1989, Cash distributions to shareholders, *Journal of Economic Perspectives* 3, 129-149.

Bansal, Ravi and Amir Yaron, 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, *Journal of Finance* 59(4), 1481-1509.

Bansal, Ravi and Christian Lundblad, 2002, Market efficiency, asset returns, and the size of the risk premium in global equity markets, *Journal of Econometrics* 109, 195-237.

Bekaert, Geert and Campbell R. Harvey, 1995, Time-varying world market integration, *Journal of Finance* 50, 403-444.

Campbell, John Y., 1991, A Variance Decomposition for Stock Returns, *The Economic Journal* 101, 157-179.

Campbell, John Y. and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195-228.

Campbell, John. Y., and Robert. J. Shiller, 1998, Valuation Ratios and the Long-Run Stock Market Outlook, *Journal of Portfolio Management* 24(6), 11-26.

Carrasco, Marine and Jean-Pierre Florens, 2002, Simulation-based method of moments and efficiency, *Journal of Business and Economic Statistics* 20(4), 482-492.

Claus, James and Jacob Thomas, 2001, Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets, *Journal of Finance* 56(5), 1629-1666.

Cochrane, John H., 2001 *Asset Pricing,* Princeton NJ, Princeton University Press.

Corradi, Valentina and Norman R. Swanson, 2005, Bootstrap specification tests for diffusion processes, *Journal of Econometrics* 124(1), 117-148.

Diamond, Peter A., 2000, What stock market returns to expect for the future? *Social Security Bulletin* 63(2), 38-52.

Dimson, Elroy, Paul Marsh, and Mike Staunton, 2007, The Worldwide Equity Premium: A Smaller Puzzle, in Rajnish Mehra, ed.: *Handbook of Investments: Equity Risk Premium*, Amsterdam: Elsevier.

Donaldson, R. Glen and Mark J. Kamstra, 1996, A New Dividend Forecasting Procedure That Rejects Bubbles in Asset Prices, *Review of Financial Studies* 9, 333-383.

Duffie, Darrell and Kenneth J. Singleton, 1993, Simulated Moments Estimation of Markov Models of Asset Prices, *Econometrica* 61(4), 929-952.

Fama, Eugene F. and Kenneth R. French, 1997, Industry Costs of Equity, *Journal of Financial Economics* 43, 153-193.

Fama, Eugene F. and Kenneth R. French, 2001, Disappearing dividends: Changing firm characteristics or lower propensity to pay, *Journal of Financial Economics* 60, 3-43.

Fama, Eugene F. and Kenneth R. French, 2002, The Equity Premium, *Journal of Finance* 57 (2), 637-659.

Graham, John, and Campbell Harvey, 2005, The Long-Run Equity Risk Premium, *Finance Research Letters* 2, 185-194.

Gregory, Allan W. and Gregor W. Smith, 1991, Calibration as Testing: Inference in Simulated Macroeconomic Models. *Journal of Business and Economic Statistics* 9(3), 297-303.

Gordon, Myron, 1962, *The Investment, Financing and Valuation of the Corporation*, Homewood, IL: Irwin.

Grullon, Gustavo and Roni Michaely, 2002, Dividends, Share Repurchases, and the Substitution Hypothesis, *Journal of Finance* 57(4), 1649-1684.

Hannan, E.J., 1980, The Estimation of the Order of an ARMA Process, *The Annals of Statistics* 8, 1071-1081.

Hodrick, Robert J., 1992, Dividend Yields and Expected Stock Returns: Alternative Procedures for inference and Measurement, *Review of Financial Studies* 5(3), 357-386.

Hull, John, 1993, *Options, Futures and Other Derivative Securities,* Englewood Cliffs, NJ: Prentice-Hall.

Jagannathan, Ravi, Ellen R. McGrattan, and Anna Scherbina, 2000, The Declining US Equity Premium, NBER Working Paper W8172, *Federal Reserve Bank of Minneapolis Quarterly Review* 24, 3-19.

Jain, Pankaj K. 2005, Financial Market Design and the Equity Premium: Electronic versus Floor Trading, *Journal of Finance* 60(6), p2955-2985

Jorion, Philippe and William N. Goetzmann, 1999, Global Stock Markets in the Twentieth Century, *Journal of Finance* 54(3), 953-980.

Kendall, Maurice and Alan Stuart, 1977, *The Advanced Theory of Statistics,* Volume 1, 4th edition, New York: MacMillan Publishing.

Kleidon, Allan W., 1986, Variance Bound Tests and Stock Price Valuation Models, *Journal of Political Economy* 94, 953-1001.

Kocherlakota, Narayana R., 1996, The Equity Premium: It's Still a Puzzle, *Journal of Economic Literature* 34(1), 42-71.

Maheu, John M. and Thomas H. McCurdy, 2007, How useful are historical data for forecasting the long-run equity return distribution? forthcoming, *Journal of Business and Economic Statistics*.

Mankiw, N. Gregory, David H. Romer, and Matthew D. Shapiro, 1991, Stock Market Forecastability and Volatility: A Statistical Appraisal, *Review of Economic Studies* 58(3), 455-477.

McFadden, Daniel, 1989, A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration, *Econometrica* 47, 995-1026.

Mehra, Rajnish, 2003, The Equity Premium: Why Is It A Puzzle? *Financial Analysts Journal* 59(1) 54-69.

Mehra, Rajnish and Edward C. Prescott, 1985, The Equity Premium: A Puzzle, *Journal of Monetary Economics* 15, 145-161.

Mehra, Rajnish and Edward C. Prescott, 2003, The Equity Premium in Retrospect, in Constantinides, G.M., M. Harris, and R. Stulz, eds.: *Handbook of the Economics of Finance* (Elsevier).

Merton, Robert C., 1980, On Estimating the Expected Return on the Market: An Exploratory Investigation, *Journal of Financial Economics* 8, 323-361.

Nelson, Daniel B., 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* 59, 347-370.

Newey, Whitney K. and Kenneth D. West, 1987, A Simple, Positive, Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.

Pakes, Ariel and David Pollard, 1989, Simulation and the Asymptotics of Optimization Estimators, *Econometrica* 57, 1027-1057.

Pástor, Ľuboš and Robert Stambaugh, 2001, The Equity Premium and Structural Breaks, *Journal of Finance* 56(1), 1207-1239.

Rozeff, Michael S., 1984, Dividend yields are equity risk premiums, *Journal of Portfolio Management* 11(1), 68-75.

Scott, Louis O., 1985, The Present Value Model of Stock Prices: Regression Tests and Monte Carlo Results, *The Review of Economics and Statistics* 67(4), 599-605.

Siegel, Jeremy J., 1999, The shrinking equity premium, *Journal of Portfolio Management*, 26(1), 10-17.

Shiller, Robert J., 1984, Stock Prices and Social Dynamics, *Brookings Papers on Economic Activity*, 457-498.

Siegel, Jeremy J. and Richard H. Thaler, 1997, Anomalies: The Equity Premium Puzzle, *Journal of Economic Perspectives* 11(1), 191-200.

Timmermann, Allen G., 1993, How Learning in Financial Markets Generates Excess Volatility and Predictability in Stock Prices, *The Quarterly Journal of Economics* 108(4), 1135-1145.

Timmermann, Allen G., 1995, Tests of Present Value Models with a Time-Varying Discount Factor, *Journal of Applied Econometrics* 10(1), 17-31.

Welch, I., 2000, Views of Financial Economists on the Equity Premium and on Professional Controversies, *Journal of Business* 73(4), 501-537.

West, Kenneth D., 1988a, Bubbles, Fads and Stock Price Volatility Tests: A Partial Evaluation, *Journal of Finance* 43, 639-656.

West, Kenneth D., 1988b, Dividend Innovations and Stock Price Volatility, *Econometrica* 56, 37-61.

White, Halbert, 1984, *Asymptotic Theory for Econometricians,* Orlando: Academic Press.

Appendices

Appendix 1: Models for Generating Data

In creating distributions of financial variables modeled on the US economy, we must generate the fundamental factors that drive asset prices: dividends and discount rates (where the discount rate is defined as the risk-free rate plus a possibly time-varying equity premium). Thus we must specify time-series models for dividend growth, interest rates, and ex ante equity premia so that our Monte Carlo simulations will generate dividends and discount rates that share key features with observed S&P 500 dividends and US discount rates. We consider a range of models to generate data in our simulations, as outlined in Table I. Each model incorporates specific characteristics that define the way we generate interest rates and dividend growth rates, and each model makes specific assumptions about the way the ex ante equity premium evolves over time, if indeed it does evolve over time. In providing further information about these defining aspects of our models, we consider each model feature from Table I in turn, starting with the time-series processes for interest rates, dividend growth rates, and the ex ante equity premium.

A1.1 Processes for the Interest Rate, Dividend Growth Rate and the Ex Ante Equity Premium

The interest rate and dividend growth rate series we generate are calibrated to the time-series properties of data observed in the US over the period 1952 to 2004. We considered the ability of various time-series models to eliminate residual autocorrelation and ARCH (evaluated with LM tests for residual autocorrelation and for ARCH, both using 5 lags), and we evaluated the log likelihood function and Bayesian Information Criterion (BIC) across models. Although we will describe the process of model selection one variable at-a-time, our final models were chosen using a Full Information Maximum Likelihood (FIML) systems equation estimation and a joint-system BIC optimization.

Economic theory admits a wide range of possible processes for the risk-free interest rate, from constant to autoregressive and highly non-linear heteroskedastic forms. We find that in practice, both AR(1) and ARMA(1,1) models of the logarithm of interest rates, based on the model of Hull (1993, page 408), perform well in capturing the time-series properties of observed interest rates. We also find the AR(1) and ARMA(1,1) specifications perform comparably to one another, markedly dominating the performance of other specifications including higher order models like ARMA(2,2). An attractive feature of modeling the log of interest rates is that doing so restricts nominal interest rates to be positive. Finally, we find standard tests for normality of the error term (and hence conditional log-normality of interest rates) do not reject the null of normality.

Since dividend growth rates have a minimum value of -100% and no theoretical maximum, a natural choice for their distribution is the log-normal. Thus we model the log of 1 plus the dividend growth rate, and we find that both a MA(1) and an AR(1) specification fit the data well, removing evidence of residual autocorrelation and ARCH at five lags. These specifications are preferred on the basis of the same criteria used to choose the specification for modeling interest rates. As with the interest rate data, we find standard tests for normality of the error term (and hence conditional log-normality of dividend growth rates) do not reject the null of normality.

Most of our models incorporate an ex ante equity premium that follows an ARMA process emerging from Merton's (1980) conditional CAPM. Merton's conditional CAPM is expressed in terms of returns in excess of the risk-free rate, or, in other words, the period-by-period equity premium. For the i^{th} asset,

$$E_t(r_{i,t}) = \lambda \; cov_{t-1}(r_{i,t} r_{m,t}), \tag{10}$$

where $r_{i,t}$ are excess returns on the asset, $r_{m,t}$ are excess returns on the market portfolio, cov_{t-1} is the time-varying conditional covariance between excess returns on the asset and on the market portfolio, and E_t is the conditional-expectations operator incorporating information available to the market up to but not including the beginning of period t. λ is a parameter of the model, described below.

For the expected excess market return, (10) becomes

$$E_t(r_{m,t}) = \lambda \; var_{t-1}(r_{m,t}) \tag{11}$$

where var_{t-1} is the market time-varying conditional variance. Merton (1980) argues that λ in (11) is the weighted sum of the reciprocal of each investor's coefficient of relative risk aversion, with the weight being related to the distribution of wealth among individuals.

Equation (11) defines a time-varying equity premium but has the equity premium varying only as a function of time-varying conditional variance. Following Bekaert and Harvey (1995), it is possible to allow λ in Equation (11) to vary over time by making it a parametric function of conditioning variables (indicated below as Z_{t-1}). The functional form Bekaert and Harvey employ (in Equation (12) of their paper) is exponential, restricting the price of risk to be positive:

$$\lambda_{t-1} = exp\left(\delta' Z_{t-1}\right). \tag{12}$$

Shiller (1984), Rozeff (1984), Campbell and Shiller (1988), Hodrick (1992), and Bekaert and Harvey (1995) all document the usefulness of dividend yields to predict returns, so we use lagged dividend yields as our conditioning variable. We make use of a simple ARCH specification to model $var_{t-1}(r_{m,t})$. Once again we calibrate to the S&P 500 over 1952 to 2004, estimating the following model:

$$r_{m,t} = \lambda_{t-1}\ var_{t-1}(r_{m,t}) + e_{m,t} \tag{13}$$

$$var_{t-1}(r_{m,t}) = \omega + \alpha e^2_{m,t-1} \tag{14}$$

$$\lambda_{t-1} = exp\left(\delta_0 + \delta_1 \frac{D_{t-1}}{P_{t-1}}\right). \tag{15}$$

The values of estimated parameters are $\delta_0 = -3.93$, $\delta_1 = 0.277$, $\omega = 0.0194$, and $\alpha = 0.542$. The R^2 of this model is 2.8%.

For our simulations, we model the time-series process of the ex ante time-varying equity premium (denoted π_t) by using the excess return as a proxy for the equity premium:

$$\hat{\pi}_t = \hat{\lambda}_{t-1} v\hat{a}r_{t-1}(r_{m,t}), \tag{16}$$

where $\hat{\lambda}_{t-1} = exp\left(-3.93 + 0.277\frac{D_{t-1}}{P_{t-1}}\right)$, $v\hat{a}r_{t-1}(r_{m,t}) = 0.0194 + 0.542\hat{e}^2_{m,t-1}$, and $\hat{e}_{m,t-1} = r_{m,t-1} - \hat{\pi}_{t-1}$. The time-varying equity premium we estimate here, $\hat{\pi}_t$, follows a strong AR(1) time-series process, similar to that of the risk-free interest rate,[17] so that when the equity premium is perturbated it reverts to its mean slowly. This permits slightly more volatile returns in our simulations than would otherwise be the case. The best way to see the impact of this slow mean reversion of the equity premium on our simulations is to compare Models 9 and 10. Model 9 has a conditionally time-varying equity premium (together with a trend and break in the premium) while Model 10 is identical except the equity premium does not conditionally vary. We find standard tests for normality of the error term (and hence conditional log-normality of the equity premium) show some evidence of non-normality when estimated as a single equation, but less or no evidence if estimated in a system of equations with the interest rate and dividend growth rate equations.

Hence we generate the ex ante equity premia, interest rate, and dividend growth rate series as autocorrelated series with jointly normal error terms, calibrated to the degree of autocorrelation observed in the US data. The processes we simulate also mimic the covariance structure between the residuals from the time-series models of equity premia, interest rates, and dividend growth rates as estimated using US data. We adjust the mean and the standard deviation of these log-normal processes to generate the desired level and variability for each when they are transformed back into levels. The coefficients and error covariance structure are estimated with FIML (very similar results are obtained using iterative GMM and Newey and West, 1987, heteroskedasticity and autocorrelation consistent covariance estimation).

To give a sense for what our estimated models for interest rates, dividend growth rates, and the equity premium look like, we present in Table A.I the estimated parameters of Model 1, which incorporates an AR(1) model for interest rates (r), a MA(1) model for dividend growth rates (g), and an AR(1) model for the ex ante equity premium (π).

[17] The mean of the estimated equity premium from this model is 5.8% and its standard deviation is 2.2%. An AR(1) model of the natural logarithm of the equity premium has a coefficient of 0.79 on the lagged equity premium, with a standard error of 0.050 and an R^2 of 0.83.

Table A.I
Estimated Parameters of Model 1

$log(r_t)$	$=$	-0.214	$+0.929\ log(r_{t-1})$	$+\epsilon_{r,t}$
		(0.262)	(0.086)	
$log(1+g_t)$	$=$	0.0516	$+0.454\ \epsilon_{g,t-1}$	$+\epsilon_{g,t}$
		(0.0063)	(0.084)	
$log(\hat{\pi}_t)$	$=$	-0.562	$+0.851\ log(\hat{\pi}_{t-1})$	$+\epsilon_{\pi,t}$
		(0.230)	(0.070)	

In Table A.I, standard errors of the estimated coefficients are shown in parentheses. The covariance of $\epsilon_{r,t}$ and $\epsilon_{g,t}$ equals 0.00240, the covariance of $\epsilon_{r,t}$ and $\epsilon_{\pi,t}$ equals -0.0117, and the covariance of $\epsilon_{g,t}$ and $\epsilon_{\pi,t}$ equals 0.0018. The variance of $\epsilon_{r,t}$ equals 0.0890, the variance of $\epsilon_{g,t}$ equals 0.000986, and the variance of $\epsilon_{\pi,t}$ equals 0.0648. The adjusted R^2 for the interest rate equation is 72.9%, the adjusted R^2 for the dividend growth rate equation is 30.0%, and the adjusted R^2 for the equity premium equation is 79.5%.

A1.2 Allowing a Downward Trend in the Ex Ante Equity Premium Process

Pástor and Stambaugh (2001), among others, provide evidence that the equity premium has been trending downward over the sample period we study, finding a modest downward trend of roughly 0.80% in total since the early 1950s, with much of the difference coming from a steep decline in the 1990s. Their study of the equity premium has the premium fluctuating between about 4% and 6% since 1834. Given this evidence and the fact that we calibrate to data starting in the 1950s, we investigate a 0.80% trend in the equity premium, and when modeling a trend with a break we limit ourselves to a 0.30% trend with an additional 50 basis point break, as discussed below. This is accomplished in conjunction with setting the ex ante equity premium to follow an AR(1) process.

A1.3 Allowing a Structural Break in the Equity Premium Process

Pástor and Stambaugh (2001) estimate the probability of a structural break in the equity premium over the last two centuries. They find fairly strong support for there having been a structural break over the 1990s which led to a 0.5% drop in the equity premium. An aggressive interpretation of their results would have the majority of the drop in the equity premium over the 1990s occurring at once. We decide to adopt a one-time-drop specification because doing so makes our results more conservative (*i.e.* produces a wider confidence interval for the ex ante equity premium). Spreading the drop in the premium across several years serves only to narrow the range of ex ante equity premium consistent with the US returns data over the last 50 years, which would only bolster our claims to provide a much tighter confidence interval about the estimate of the ex ante equity premium. Thus we incorporate an abrupt 50 basis point drop in the equity premium in some of the models we consider. We time the drop to coincide with 1990, 39 years into our simulation period. This feature of the equity premium process can be accomplished with or without incorporating other features discussed above.

A1.4 Allowing for Sampling Variability in Generating Parameters

Our experiments are motivated by the large sampling variability of the ex post equity premium, but when we produce our simulations we have to first estimate the parameter values for the time-series models of dividend growth rates, interest rates, and ex ante equity premia. These estimates themselves incorporate sampling variability. Fortunately, estimates of the sampling variability are available to us through the covariance matrix of our parameters, so we can incorporate uncertainty about the true values of these parameters into our simulations. We estimate our system of equations (the dividend growth rate, interest rate, and the ex ante equity premium equation) jointly with FIML, and generate for *each* simulation an independent set of parameters drawn randomly from the joint limiting normal distribution of these parameter estimates (including the variance and covariance of the equation residuals) subject to some technical considerations[18] and data consistency checks.[19] This process accounts for possible variability in the true state of the world that generates dividends, interest rates, and ex ante equity premia.

To illustrate, for Model 1 reported in Table A.I,

[18] The time-series models must exhibit stationarity, the growth rate of dividends must be strictly less than the discount rate, and the residual variances must be greater than zero.

[19] The parameters must generate mean interest rates, dividend growth rates, and ex post equity premia that lie within three standard deviations of the US data sample mean. Also, the limiting price-dividend ratio must be within 50 standard deviations of the mean US price-dividend ratio. This last consistency check rules out some extreme simulations generated when the random draw of parameters leads to near unit root behavior. The vast majority of simulations do not exhibit price-dividend ratios that are more than a few standard deviations from the mean of the US data.

$$
\begin{aligned}
log(r_t) &= \alpha_r &&+\rho_r \; log(r_{t-1}) &&+\epsilon_{r,t} \\
log(1+g_t) &= \alpha_g &&+\theta_g \; \epsilon_{g,t-1} &&+\epsilon_{g,t} \\
log(\hat{\pi}_t) &= \alpha_\pi &&+\rho_\pi \; log(\hat{\pi}_{t-1}) &&+\epsilon_{\pi,t},
\end{aligned}
$$

the estimated covariance matrix of the parameter estimates is shown in Table A.II.

Table A.II
Estimated Covariance Matrix for Model 1 Parameters

	α_r	ρ_r	α_g	θ_g	α_π	ρ_π
α_r	0.068705	0.022307	-.000051933	.000226443	-0.012165	-0.003511
ρ_r	0.022307	0.007436	-.000040346	.000114831	-0.004730	-0.001401
α_g	-0.000052	-0.000040	0.000039674	.000025651	0.000153	0.000031
θ_g	0.000226	0.000115	0.000025651	.007086714	0.001699	0.000454
α_π	-0.012165	-0.004730	0.000153376	.001699151	0.052664	0.015791
ρ_π	-0.003511	-0.001401	0.000031495	.000453874	0.015791	0.004844

The top-left element of Table A.II, equal to 0.068705, is the variance of the parameter estimate of α_r. The entry below the top-left element, equal to 0.022307, is the covariance between the estimate of α_r and ρ_r, and so on. The *estimated covariance matrix* of the equation residual variances is shown in Table A.III. (The variances themselves are reported in Section A1.1, as are the parameter estimates of the mean.)

Table A.III
Estimated Covariance Matrix of Model 1 Residual Variances

	ϵ_r^2	$\epsilon_r\epsilon_g$	$\epsilon_r\epsilon_\pi$	ϵ_g^2	$\epsilon_g\epsilon_\pi$	ϵ_π^2
ϵ_r^2	0.0000944	$1.9729{\cdot}10^{-6}$	$-8.351{\cdot}10^{-7}$	$-1.902{\cdot}10^{-7}$	$-1.564{\cdot}10^{-6}$	$-1.69{\cdot}10^{6}$
$\epsilon_r\epsilon_g$	$1.9729{\cdot}10^{-6}$	$8.5163{\cdot}10^{-7}$	$1.0437{\cdot}10^{-6}$	$4.3066{\cdot}10^{-8}$	$-1.602{\cdot}10^{-7}$	$9.1448{\cdot}10^{-7}$
$\epsilon_r\epsilon_\pi$	$-8.351{\cdot}10^{-7}$	$1.0437{\cdot}10^{-6}$	0.0000797	$1.8827{\cdot}10^{-7}$	$5.001{\cdot}10^{-6}$	-0.000044
ϵ_g^2	$-1.902{\cdot}10^{-7}$	$4.3066{\cdot}10^{-8}$	$1.8827{\cdot}10^{-7}$	$4.8337{\cdot}10^{-8}$	$9.6885{\cdot}10^{-8}$	$1.3458{\cdot}10^{-6}$
$\epsilon_g\epsilon_\pi$	$-1.564{\cdot}10^{-6}$	$-1.602{\cdot}10^{-7}$	$5.001{\cdot}10^{-6}$	$9.6885{\cdot}10^{-8}$	$3.5567{\cdot}10^{-6}$	0.0000203
ϵ_π^2	$-1.69{\cdot}10^{-6}$	$9.1448{\cdot}10^{-7}$	-0.000044	$1.3458{\cdot}10^{-6}$	0.0000203	0.0005009

The top-left element, equal to 0.0000944, is the variance of ϵ_r^2. The entry below the top-left element, equal to $-1.9729{\cdot}10^{-6}$, is the covariance between the estimate of ϵ_r^2 and the product of ϵ_r and ϵ_g, and so on.

Exploiting block diagonality of the parameters of the mean and variance, and asymptotic normality of all the estimated parameters, we generate two sets of normally distributed random variables.

Each set is independent of the other, the first set of six having the covariance matrix from Table A.II with means equal to the parameter estimates listed in Table A.I, and the second set of six having the covariance matrix from Table A.III, with means equal to the equation residual covariances listed in Section A1.1. This set of 12 random variables is then used to simulate interest rates, dividend growth rates, and equity premia, subject to the consistency checks footnoted earlier.

A1.5 Allowing for Disappearing Dividends

An issue with our calibration to dividends is the impact of declining dividend payments in the US. This phenomenon is a result of a practice adopted widely beginning in the late 1970s, whereby US firms have been increasingly delivering cashflows to investors in ways not recorded as corporate dividends, such as share repurchases. Fama and French (2001) document the widespread decline of regular dividend payments starting in 1978, consistent with evidence provided by Bagwell and Shoven (1989) and others. Fama and French find evidence that the disappearance of dividends is in part due to an increase in the inflow of new listing to US stock exchanges, representing mostly young companies with the characteristics of firms that would not be expected to pay dividends, and in part due to a decline in the propensity of firms to pay dividends. Fama and French find only a small decline in the probability to pay dividends among the firms that we calibrate to, those in the S&P 500 index.

Consistent with Fama and French, we find no evidence of a break in our data on dividend growth rates. Though dividend *yields* on the S&P 500 index have dropped dramatically over time, dividend growth rates have not. The decline in yields has been a function of prices rising faster than dividends since 1978, not dividends declining in any absolute sense. From 1952 through 1978, the year Fama and French document as the year of the structural break in dividend payments, dividend growth rates among the S&P 500 firms have averaged 4.9% with an annual standard deviation of 3.9%, and from 1979 to 2000 the dividend growth rates have averaged 5.5% with an annual standard deviation of 3.8%, virtually indistinguishable from the pre-1979 period. Time series properties pre- and post-1978 are also very similar across these two periods. Consistent with this stability of dividend growth pre- and post-1978 and Bagwell and Shoven's documentation of increased share repurchases in the 1980s, *earnings growth rates* of firms in the S&P 500 index have accelerated since the 1952-1978 period, from 6.8% pre-1979 to 7.8% post-1978. Similar to the dividend growth rate data, the time-series properties of the earnings growth rate data did not change.

In order to determine the sensitivity of our experiments to mismeasurement of cashflows to investors, we consider a dividend growth rate process with a structural break 27 years into the time series to correspond to a possible break in our dividend data for the S&P 500 data after 1978. We calibrate to the S&P 500 earnings data mean growth rate increase over 1979-2000, an upward shift of 100 basis points, to proxy for the increase in total cashflows to investors. That is, we increase the growth rate of dividends by 5 basis points a year for 20 years, starting in year 27 of the simulation (corresponding to 1978 for the S&P 500 data), to increase the mean growth rate of our dividend growth series 100 basis points, mimicking the proportional increase in earnings growth rates.

Appendix 2: Further Details on the Simulations

A2.1 Fundamentals

We define P_t as a stock's beginning-of-period-t price and E_t as the expectations operator conditional on information available up to but not including the beginning of period t. The discount rate (r_t, which equals the risk-free rate plus the equity premium) is the rate investors use to discount payments received during period t (*i.e.*, from the beginning of period t to the beginning of period $t+1$). Recall that investor rationality requires that the time t market price of a stock, which will pay a dividend D_{t+1} one period later and then sell for P_{t+1}, satisfy Equation (3):

$$P_t = E_t \left\{ \frac{P_{t+1} + D_{t+1}}{1 + r_t} \right\}. \tag{3}$$

Invoking the standard transversality condition that the expected present value of the stock price P_{t+i} falls to zero as i goes to infinity, and defining the growth rate of dividends during period t as $g_t \equiv (D_{t+1} - D_t)/D_t$, allows us rewrite Equation (3) as:

$$P_t = D_t E_t \left\{ \sum_{i=0}^{\infty} \left(\Pi_{k=0}^{i} \left[\frac{1 + g_{t+k}}{1 + r_{t+k}} \right] \right) \right\}. \tag{5}$$

One attractive feature of expressing the present value stock price as in Equation (5), in terms of dividend growth rates and discount rates, is that this form highlights the irrelevance of inflation, at least to the extent that expected and actual inflation are the same. Notice that working with nominal growth rates and discount rates, as we do, is equivalent to working with deflated nominal rates (*i.e.*, real rates). That is, $\frac{1+([g_t-I_t]/[1+I_t])}{1+([r_t-I_t]/[1+I_t])} = \frac{(1+g_t)}{(1+r_t)}$, where I_t is inflation. Working with nominal values in our simulations removes a potential source of measurement error associated with attempts to estimate inflation.

Properties of prices and returns produced by Equation (5) depend in important ways on the modeling of the dynamics of the dividend growth, interest rate, and equity premium processes. For instance, the stock price would equal a constant multiple of the dividend level and returns would be very smooth over time if dividend growth and interest rates were set equal to constants plus independent innovations. However, using models that capture the serial dependence of dividend growth rates, interest rates, and equity premia observed in the data, as we do, would typically lead to time-varying price-dividend ratios and variable returns of the sort we observe in observed stock market data.

A2.2 Numerical Simulation

We now provide details on the numerical simulation which comprises Step 4 of the 5-step procedure outlined in Section I above. That is, we detail for the n^{th} economy the formation of the prices (P_t^n), returns (R_t^n), ex post equity premia ($\hat{\pi}^n$), *etc.* (where $n = 1, \cdots, N$ and $t = 1, \cdots, T$), given dividends, dividend growth rates, risk-free interest rates, and the equity premium of the n^{th} economy: D_t^n, g_{t-1}^n, and $r_{t-1}^n = r_{f,t-1}^n + \pi$.[20] For simplicity, we illustrate our methodology by assuming fixed parameters (no parameter uncertainty), a constant ex ante equity premium, and an AR(1) model for interest rates. Further, to illustrate the procedure required for a moving average error model, we assume a MA(1) process for dividend growth rates. Relaxing these assumptions (the assumptions to incorporate parameter uncertainty, ARMA(1,1) processes for interest rates and dividend growth rates, and a time-varying equity premium) complicates the procedure outlined below only slightly. Note that in our actual simulations we set the initial dividend growth rate and

[20]We set the number of economies, N, at 2,000. This is a sufficiently large number of replications to produce results with very small simulation error.

interest rate to their unconditional means, innovations to zero, and dividends to \$1, then simulate the economies out for 50 periods. At period 51 we start our calculation of market prices, returns, *etc.* (to avoid contaminating the simulations with the initial conditions). For simplicity, we do not include this detail in the description below but for concreteness we describe a similar prototypical simulation.

In terms of timing and information, recall that P_t^n is the stock's beginning-of-period-t price, r_t^n is the rate used to discount payments received during period t and is known at the beginning of period t, D_t^n is paid at the beginning of period t, g_t^n is defined as $(D_{t+1}^n - D_t^n)/D_t^n$ and is not known at the beginning of period t since it depends on D_{t+1}^n, and $E_t\{\cdot\}$ is the conditional expectation operator, with the conditioning information being the set of information available to investors up to but not including the beginning of period t. Finally, recall Equation (5), rewritten to correspond to the n^{th} economy:

$$P_t^n = D_t^n E_t \left\{ \sum_{i=0}^{\infty} \left(\Pi_{k=0}^{i} \left[\frac{1+g_{t+k}^n}{1+r_{t+k}^n} \right] \right) \right\}. \tag{17}$$

Returns are constructed as $R_t^n = (P_{t+1}^n + D_{t+1}^n - P_t^n)/P_t^n$, and $\hat{\pi}^n = \overline{R}^n - \overline{r}_f^n$ where $\overline{R}^n = \frac{1}{T}\sum_{t=1}^{T} R_t^n$ and $\overline{r}_f^n = \frac{1}{T}\sum_{t=1}^{T} r_{f,t}^n$.

Based on Equation (17), we generate prices by generating a multitude of possible streams of dividends and discount rates, present-value discounting the dividends with the discount rates, and averaging the results, *i.e.*, by conducting a Monte Carlo integration.[21] Hence we produce prices (P_t^n), returns (R_t^n), ex post equity premia ($\hat{\pi}^n$), and a myriad of other financial quantities, utilizing only dividend growth rates and discount rates. The *exact* procedure by which we conduct this numerical simulation is described below and summarized in Figure A.1. (These steps, labeled Steps 4A through 4C, collectively constitute Step 4 of the 5-step procedure outlined in Section I above.)

[21] According to Equation (17), the stream of dividends and discount rates should be infinitely long, however truncating the stream at a sufficiently distant point in time denoted I leads to a very small approximation error. We discuss this point more fully below.

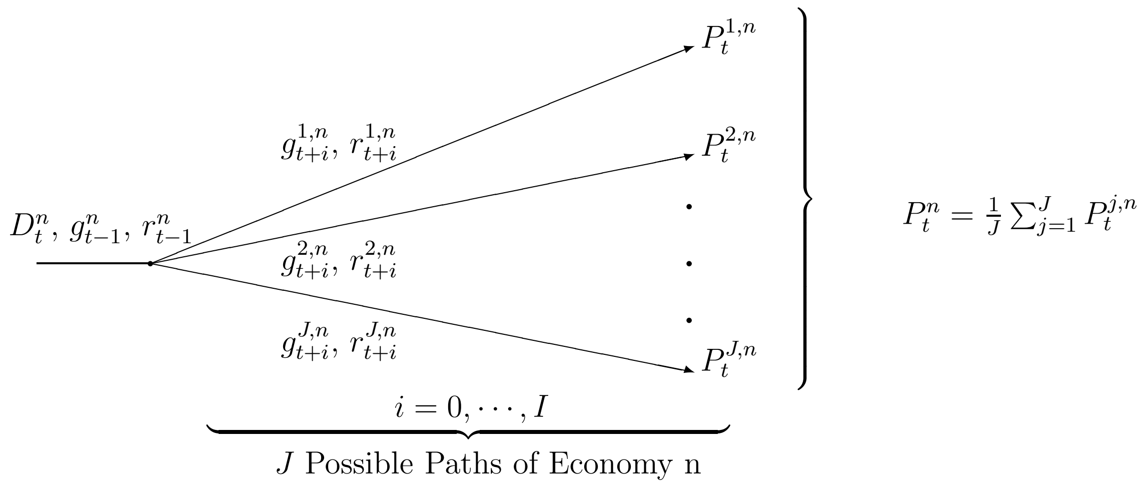


Figure A.1 *Diagram of a Simple Market Price Calculation for the t^{th} Observation of the n^{th} Economy (Steps 4A and 4B)*

Step 4A: In forming P_t^n, the most recent fundamental information available to an investor would be g_{t-1}^n, D_t^n, and r_{t-1}^n. Thus g_{t-1}^n, D_t^n, and r_{t-1}^n must be generated directly in our simulations, whereas P_t^n is calculated based on these g, D, and r. The objective of Steps 4A(i)-(iii) outlined below is to produce dividend growth and interest rates that replicate real-world dividend growth and interest rate data. That is, the simulated dividend growth and interest rates must have the same mean, variance, covariance, and autocorrelation structure as observed S&P 500 dividend growth rates and US interest rates. In terms of Figure A.1, Step 4A forms g_{t-1}^n, D_t^n, and r_{t-1}^n only.

Step 4A(i): Note that since, as described above, the logarithm of one plus the dividend growth rate is modeled as a MA(1) process, $log(1 + g_t^n)$ is a function of only innovations, labeled ϵ_g^n. Note also that since the logarithm of the interest rate is modeled as an AR(1) process, $log(r_{f,t}^n)$ is a function of $log(r_{f,t-1}^n)$ and an innovation labeled ϵ_r^n. Set the initial dividend, D_1^n, equal to the total S&P 500 dividend value for 1951 (observed at the end of 1951), and the lagged innovation of the logarithm of the dividend growth rates $\epsilon_{g,0}^n$ to 0. To match the real-world interest rate data, set $log(r_{f,0}^n) = -2.90$ (the mean value of log interest rates required to produce interest rates matching the mean of observed T-bill rates). Then generate two independent standard normal random numbers, η_1^n and ν_1^n (note that the subscript on these random numbers indicates time, t), and form two correlated random variables, $\epsilon_{r,1}^n = 0.319(0.25\eta_1^n + (1 - .25^2)^{.5}\nu_1^n)$ and $\epsilon_{g,1}^n = 0.0311\eta_1^n$. These are the simulated innovations to the interest rate and dividend growth rate processes, formed to have standard deviations of 0.319 and 0.0311 respectively to match the data, and to be correlated with correlation coefficient 0.25 as we find in the S&P 500 return and T-bill rate data. Next, form

$log(1+g_1^n) = 0.049+0.64\epsilon_{g,0}^n+\epsilon_{g,1}^n$ and $log(r_{f,1}^n) = -0.35+0.88log(r_{f,0}^n)+\epsilon_{r,1}^n$ to match the parameters estimated on the S&P 500 index data 1952-2004 of these models (using Full Information Maximum Likelihood).[22] Also form $D_2^n = D_1^n(1+g_1^n)$.

Step 4A(ii): Produce two correlated normal random variables, $\epsilon_{r,2}^n$ and $\epsilon_{g,2}^n$ as in Step 4A(i) above, and conditioning on $\epsilon_{g,1}^n$ and $log(r_{f,1}^n)$ from Step 4A(i) produce $log(1+g_2^n) = 0.049+0.64\epsilon_{g,1}^n+\epsilon_{g,2}^n$, $log(r_{f,2}^n) = -0.35+0.88log(r_{f,1}^n)+\epsilon_{r,2}^n$, and $D_3^n = D_2^n(1+g_2^n)$.

Step 4A(iii): Repeat Step 4A(ii) to form $log(1+g_t^n)$, $log(r_{f,t}^n)$, and D_t^n for $t = 3,4,5,\cdots,T$ and for each economy $n = 1,2,3,\cdots,N$. Then calculate the dividend growth rate g_t^n and the discount rate r_t^n (which equals $r_{f,t}^n$ plus the ex ante equity premium).

Step 4B: For each time period $t = 1,2,3,\cdots,T$ and economy $n = 1,2,3,\cdots,N$ we calculate prices, P_t^n. In order to do this we must solve for the expectation of the infinite sum of discounted future dividends conditional on time $t-1$ information for economy n. That is, we must produce a set of possible paths of dividends and interest rates that might be observed in periods $t, t+1, t+2, \cdots$ given what is known at period $t-1$ and use these to solve the expectation of Equation (17). We use the superscript j to index the possible paths of future economies that could possibly evolve from the current state of the economy. In Step 4B(iv) below, we describe how we are able to solve for the expectation of an infinite sum using a *finite* stream of future dividends.

Step 4B(i): Set $\epsilon_{g,t-1}^{j,n} = \epsilon_{g,t-1}^n$ and $log(r_{f,t-1}^{j,n}) = log(r_{f,t-1}^n)$ for $j = 1,2,3,\cdots,J$.[23] Generate two independent standard normal random numbers, $\eta_t^{j,n}$ and $\nu_t^{j,n}$, and form two correlated random variables $\epsilon_{r,t}^{j,n} = 0.319(0.25\eta_t^{j,n}+(1-.25^2)^{.5}\nu_t^{j,n})$ and $\epsilon_{g,t}^{j,n} = 0.0311\eta_t^{j,n}$ for $j = 1,2,3,\cdots,J$.[24] These

[22]Note that by construction these parameters do not match those reported for the system reported in Appendix 1 as this system does not incorporate a time-varying equity premium.

[23]We choose J to lie between 1,000 and 100,000, as needed to ensure the Monte Carlo simulation error in calculating prices and returns is controlled to be less than 0.20%. For the typical case the simulation error is far less than 0.20%. To determine the simulation error, we conducted a simulation of the simulations. Unlike some Monte Carlo experiments (such as those estimating the size of a test statistic under the null) the standard error of the simulation error for most of our estimates (returns, prices, *etc.*) are themselves analytically intractable, and must be simulated. In order to estimate the standard error of the simulation error in estimating market prices, we estimated a single market price 2,000 times, each time independent of the other, and from this set of prices computed the mean and variance of the price estimate. If the experiment had no simulation error, each of the price estimates would be identical. With the number of possible paths, J, equal to no less than 1,000 we find that the standard deviation of the simulation error is less than 0.20% of the price, which is sufficiently small as not to be a source of concern for our study. The number of simulations has to be substantially greater than 1,000 for some cases depending on the model specification and the ex ante equity premium.

[24]For our random number generation we made use of a variance reduction technique, stratified sampling. This technique has us drawing pseudo-random numbers ensuring that $q\%$ of these draws come from the q^{th} percentile, so that our sampling does not weight any grouping of random draws too heavily.

are the simulated innovations to the interest rate and dividend growth rate processes, respectively. Form $log(1+g_t^{j,n}) = 0.049 + 0.64\epsilon_{g,t-1}^{j,n} + \epsilon_{g,t}^{j,n}$ and $log(r_{f,t}^{j,n}) = -0.35 + 0.88log(r_{f,t-1}^{j,n}) + \epsilon_{r,t}^{j,n}$.

Step 4B(ii): Produce two correlated normal random variables $\epsilon_{r,t+1}^{j,n}$ and $\epsilon_{g,t+1}^{j,n}$ as in Step 4B(i) above, and conditioning on $\epsilon_{g,t}^{j,n}$ and $log(r_{f,t}^{j,n})$ from Step 4B(i) produce $log(1+g_{t+1}^{j,n}) = 0.049 + 0.64\epsilon_{g,t}^{j,n} + \epsilon_{g,t+1}^{j,n}$ and $log(r_{f,t+1}^{j,n}) = -0.35 + 0.88log(r_{f,t}^{j,n}) + \epsilon_{r,t+1}^{j,n}$ for $j = 1, 2, 3, \cdots, J$.

Step 4B(iii): Repeat Step 4B(ii) to form $log(1+g_{t+i}^{j,n})$ and $log(r_{t+i}^{j,n})$ for $i = 2, 3, 4, \cdots, I$, $j = 1, 2, 3, \cdots, J$, and economies $n = 1, 2, 3, \cdots, N$.

Step 4B(iv): The discounted present value of each of the individual J streams of dividends is now taken in accordance with Equation (17), with the j^{th} present value price noted as $P_t^{j,n}$. Finally, the price for the n^{th} economy in period t is formed: $P_t^n = \frac{1}{J}\sum_{j=1}^{J} P_t^{j,n}$.

In considering these prices, note that according to Equation (17) the stream of discount rates and dividend growth rates should be infinitely long, while in our simulations we extend the stream for only a finite number of periods, I. Since the ratio of gross dividend growth rates to gross discount rates are less than unity in steady state, the individual product elements in the infinite sum in Equation (17) eventually converge to zero as I increases. (Indeed, this convergence to zero is exactly what is required for the standard transversality condition that the expected present value of the stock price P_{t+i} falls to zero as i goes to infinity.) We therefore set I large enough in our simulations so that the truncation does not materially effect our results. We find that setting $I = 1,000$ years is sufficient in all cases we studied. That is, the discounted present value of a dividend payment received 1,000 years in the future is essentially zero. Also note that the steps above are required to produce P_t^n, D_t^n, g_t^n, and r_t^n for $n = 1, \cdots, N$ and $t = 1, \cdots, T$; the intermediate terms superscripted with a j are required only to perform the numerical integration that yields P_t^N. Note that the length of the time series T is chosen to be 53 to imitate the 53 years of annual data we have available for the S&P 500 from 1952 to 2004.

Step 4C: After performing Steps 4A(i)-(iii) and 4B(i)-(iv) for $t = 1, \cdots, T$, rolling out N independent economies for T periods, we construct the market returns for each economy, $R_t^n = (P_{t+1}^n + D_{t+1}^n - P_t^n)/P_t^n$, and the ex post equity premium that agents in the n^{th} economy would observe, $\hat{\pi}^n$, estimated from Equation (1) as the mean difference in market returns and the risk-free rate.

Table I

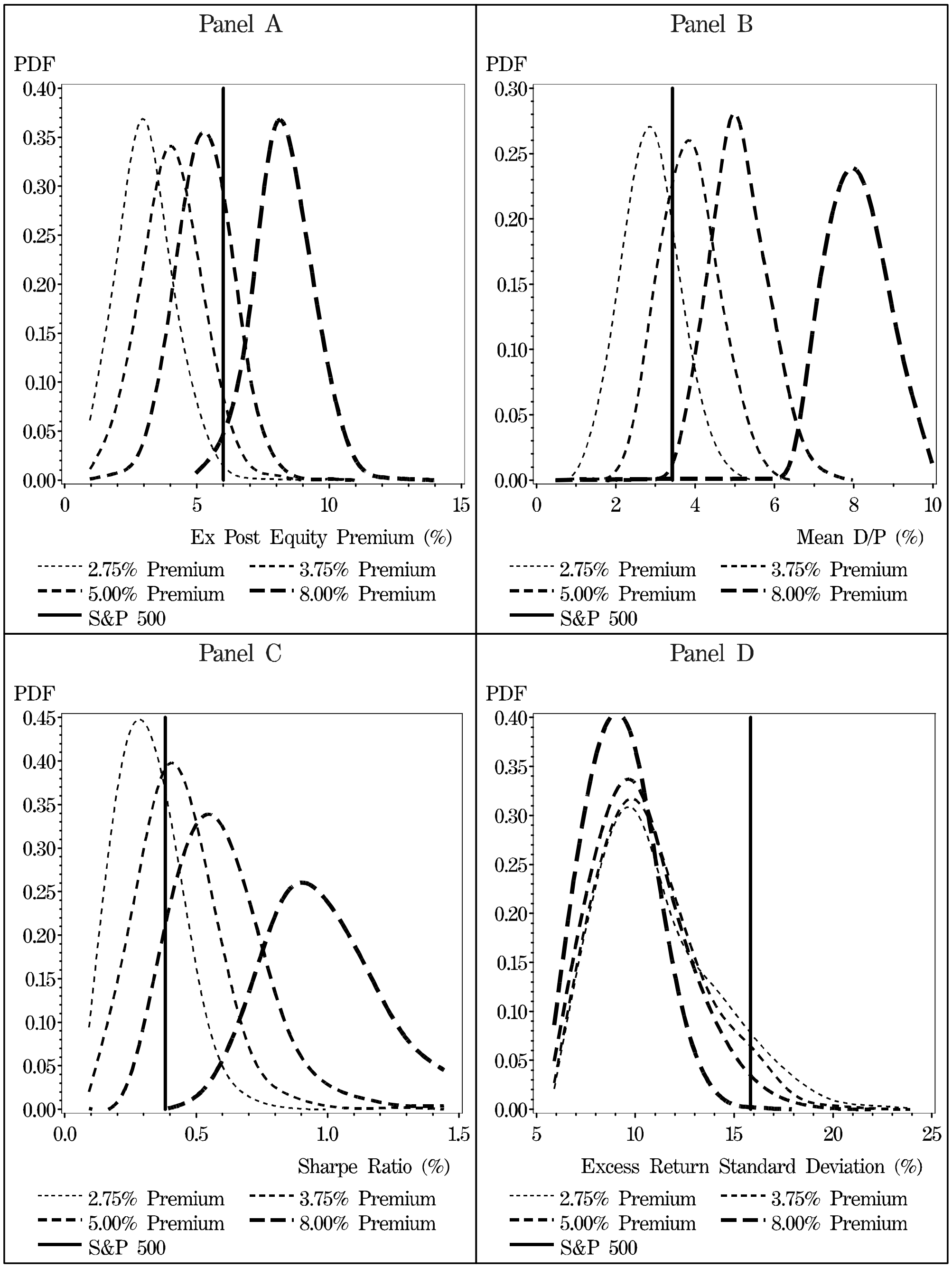
Characteristics of Simulated Models

Here we present the 12 models we consider, identifying the characteristics of their underlying data generating processes. The column titled "Processes for r, g, & π" indicates the nature of the time-series models used to generate the interest rates, dividend growth rates, and equity premium. See Appendix 1 for details on how this set of models was chosen and a description of how the equity premium series is produced. The column titled "Downward Trend in Equity Premium Process," identifies whether the ex ante equity premium trends downward over the course of the 53-year experiment, and if it does, provides the amount of the downward trend. The next column, "Structural Break in Equity Premium Process," indicates whether the model incorporates a sudden 50 basis point (bps) drop in the value of the ex ante equity premium. The column "Structural Break in Dividend Growth Process," indicates whether the model incorporates a gradual 100 basis point increase in the growth rate of the dividend growth rate. The final column indicates that all the models except Models 11 and 12 incorporate sampling variability in generating parameters. Additional model details are as follows. Parsimonious Model: interest rates follow an AR(1), dividend growth rates follow a MA(1), the equity premium follows an AR(1). Deterministic π Model: interest rates follow an AR(1), dividend growth rates follow a MA(1), the equity premium follows a deterministic downward trend with a 50 bps structural break. Best BIC Model:† interest rates follow an ARMA(1,1), dividend growth rates follow a MA(1), the equity premium follows an AR(1). Second-Best BIC Model:† interest rates follow an ARMA(1,1), dividend growth rates follow a MA(1), the equity premium follows an ARMA(1,1). Further details about each model feature are provided in Appendix 1.

Model	Processes for r, g, & π	Downward Trend in Equity Premium Process	Structural Break in Equity Premium Process	Structural Break in Dividend Growth Process	Sampling Variability in Generating Parameters
1	Parsimonious Model	No	No	No	Yes
2	Parsimonious Model with π Trend	Yes (80 bps)	No	No	Yes
3	Parsimonious Model with π Break	No	Yes (50 bps)	No	Yes
4	Parsimonious Model with Dividend Growth Trend	No	No	Yes	Yes
5	Parsimonious Model with π Trend and Dividend Growth Trend	Yes (80 bps)	No	Yes	Yes
6	Parsimonious Model with π Break, π Trend, and Dividend Growth Trend	Yes (30 bps)	Yes (50 bps)	Yes	Yes
7	Best BIC Model† with π Break, π Trend, and Dividend Growth Trend	Yes (30 bps)	Yes (50 bps)	Yes	Yes
8	Second-Best BIC Model† with π Break, π Trend, and Dividend Growth Trend	Yes (30 bps)	Yes (50 bps)	Yes	Yes
9	Parsimonious Model with π Break and π Trend	Yes (30 bps)	Yes (50 bps)	No	Yes
10	Deterministic π Model with π Break and π Trend	Yes (30 bps)	Yes (50 bps)	No	Yes
11	Parsimonious Model with Constant Parameters π Break, π Trend, and Dividend Growth Trend	Yes (30 bps)	Yes (50 bps)	Yes	No
12	Parsimonious Model with Constant Parameters	No	No	No	No

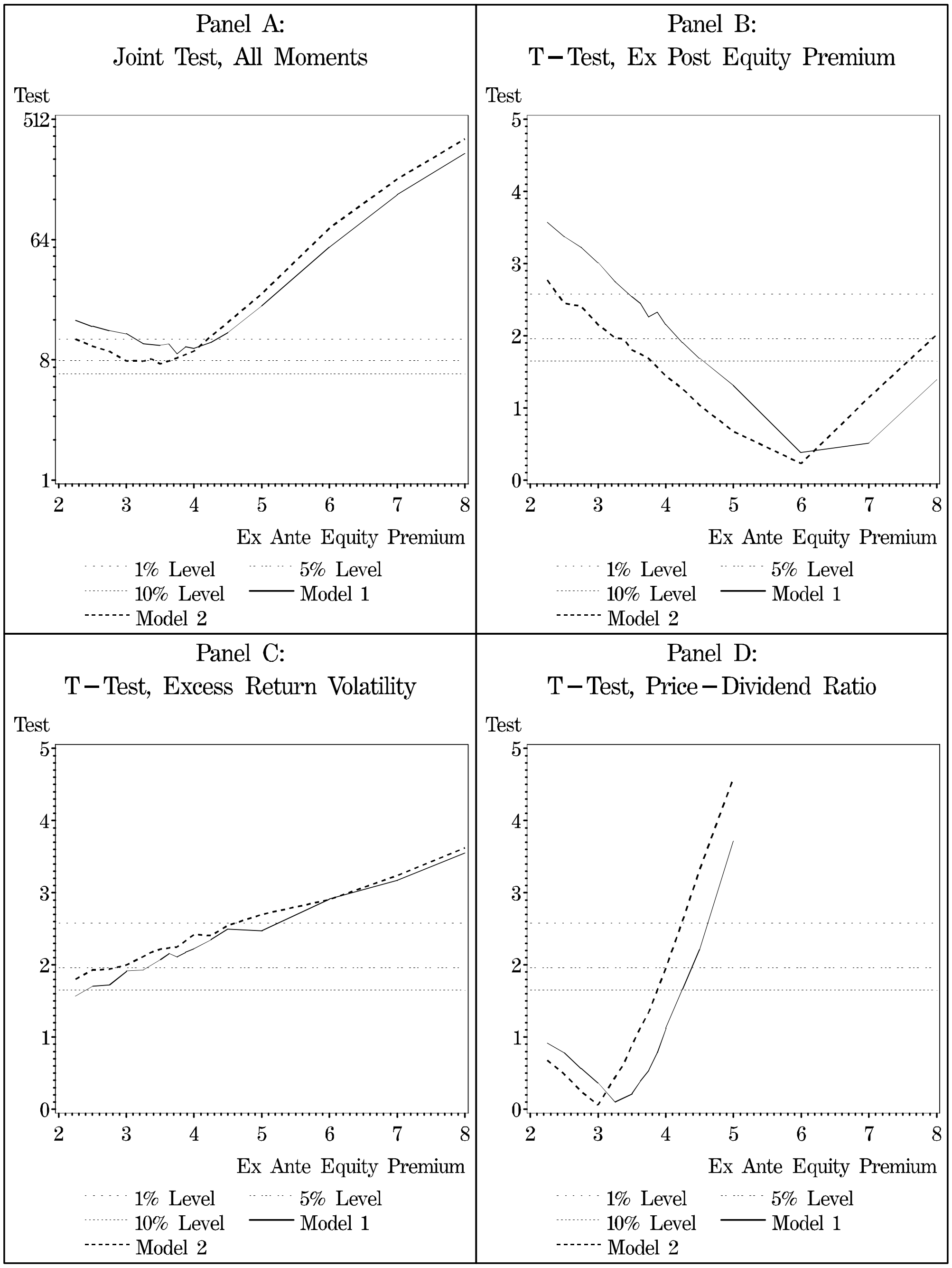
† For Models 7 and 8 we employ the Bayesian Information Criterion (BIC) to select the order of the ARMA model driving each of the interest rate, equity premium, and dividend growth rate processes. The order of each AR process and each MA process for each series is chosen over a (0, 1, 2) grid.

Figure 1: Probability Distribution Functions of Simulated Ex Post Equity Premia, Dividend Yields, Sharpe Ratios, and Return Standard Deviations



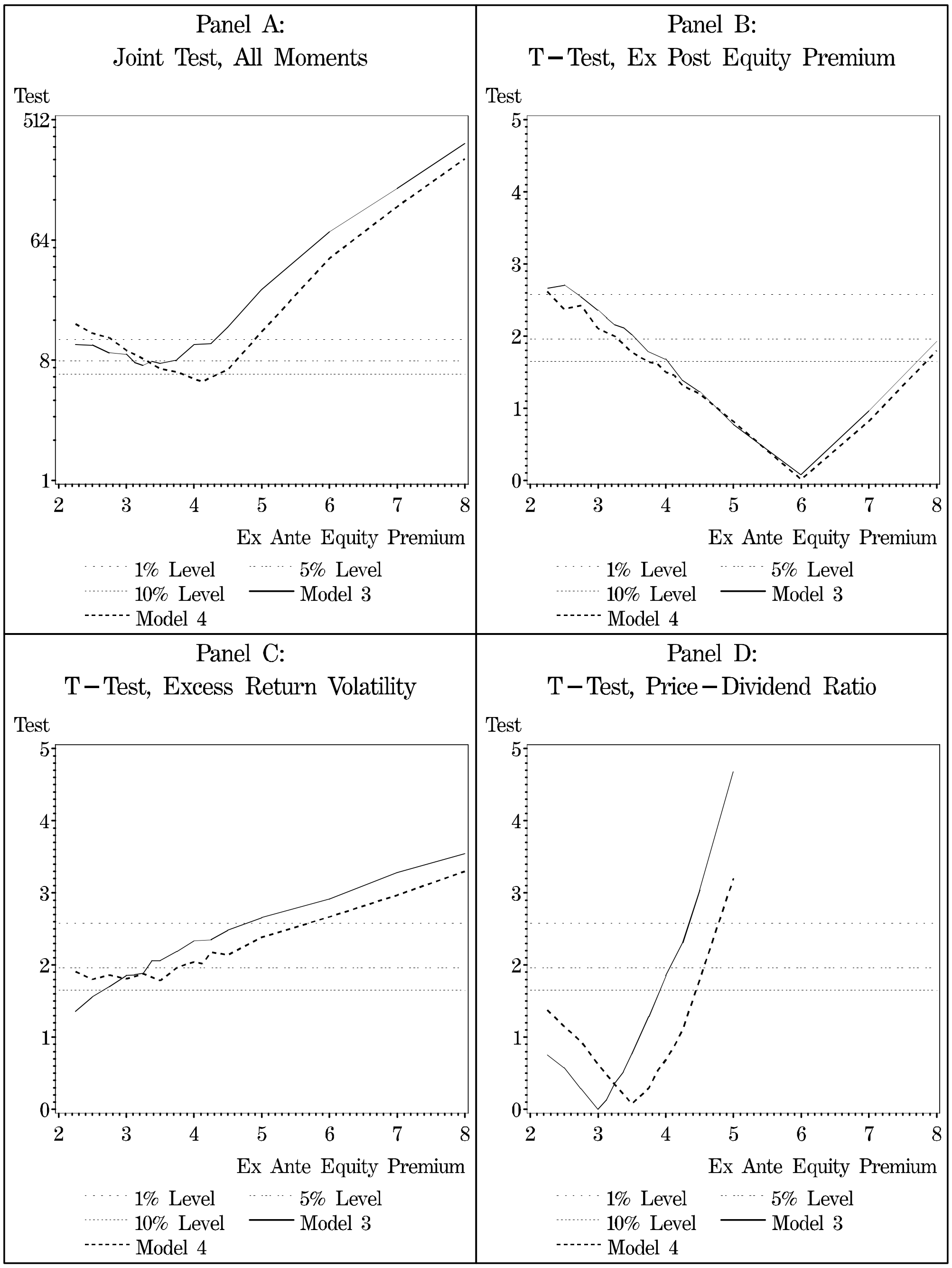
This figure contains probability distribution functions (PDFs) for various financial statistics generated in 2,000 simulated economies based on Model 1 from Table I. Each panel contains a PDF for each of four different assumed values of the ex ante equity premium: 2.75%, 3.75%, 5%, and 8%. Panel A shows the distribution of the ex post equity premium (mean return minus mean interest rate), Panel B shows the mean dividend yield distribution (dividend divided by price), Panel C shows the Sharpe ratio distribution (excess return divided by the standard deviation of the excess return), and Panel D shows the distribution of the standard deviation of excess returns. In each panel, a vertical line indicates the US data realized over 1952-2004, the value of the estimated ex post equity premium, mean dividend yield, mean Sharpe ratio, and excess return standard deviation, respectively. The simulated statistics are estimated on 53 years of generated data for each economy, mimicking the data period we used to estimate the actual US results.

Figure 2: Joint and Individual Tests Statistics for Models 1 and 2



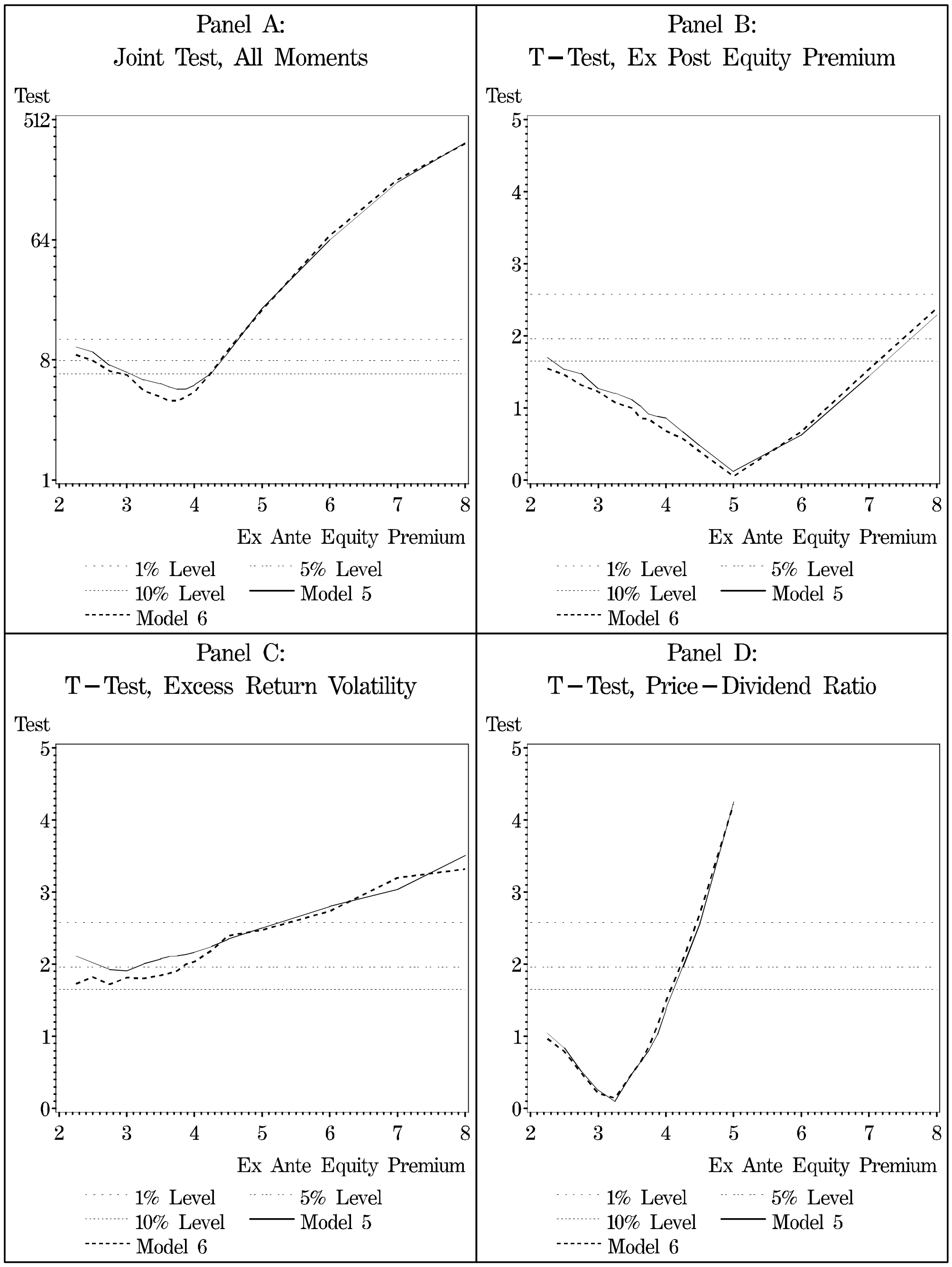
This figure contains plots of test statistics for Models 1 and 2. Panel A plots joint χ^2 tests based on a set of three variables (the ex post equity premium, the mean dividend yield, and the excess return volatility) for various ending values of the ex ante equity premium for each model. In Panel A the vertical axis is plotted on a log scale. The remaining panels contains t-test values corresponding to tests on the individual variables for each of the models: the ex post equity premium in Panel B, the excess return volatility in Panel C, and price-dividend ratio in Panel D. In each panel the critical values of the test statistics corresponding to test significance at the 10%, 5%, and 1% levels are indicated by horizontal lines.

Figure 3: Joint and Individual Tests Statistics for Models 3 and 4



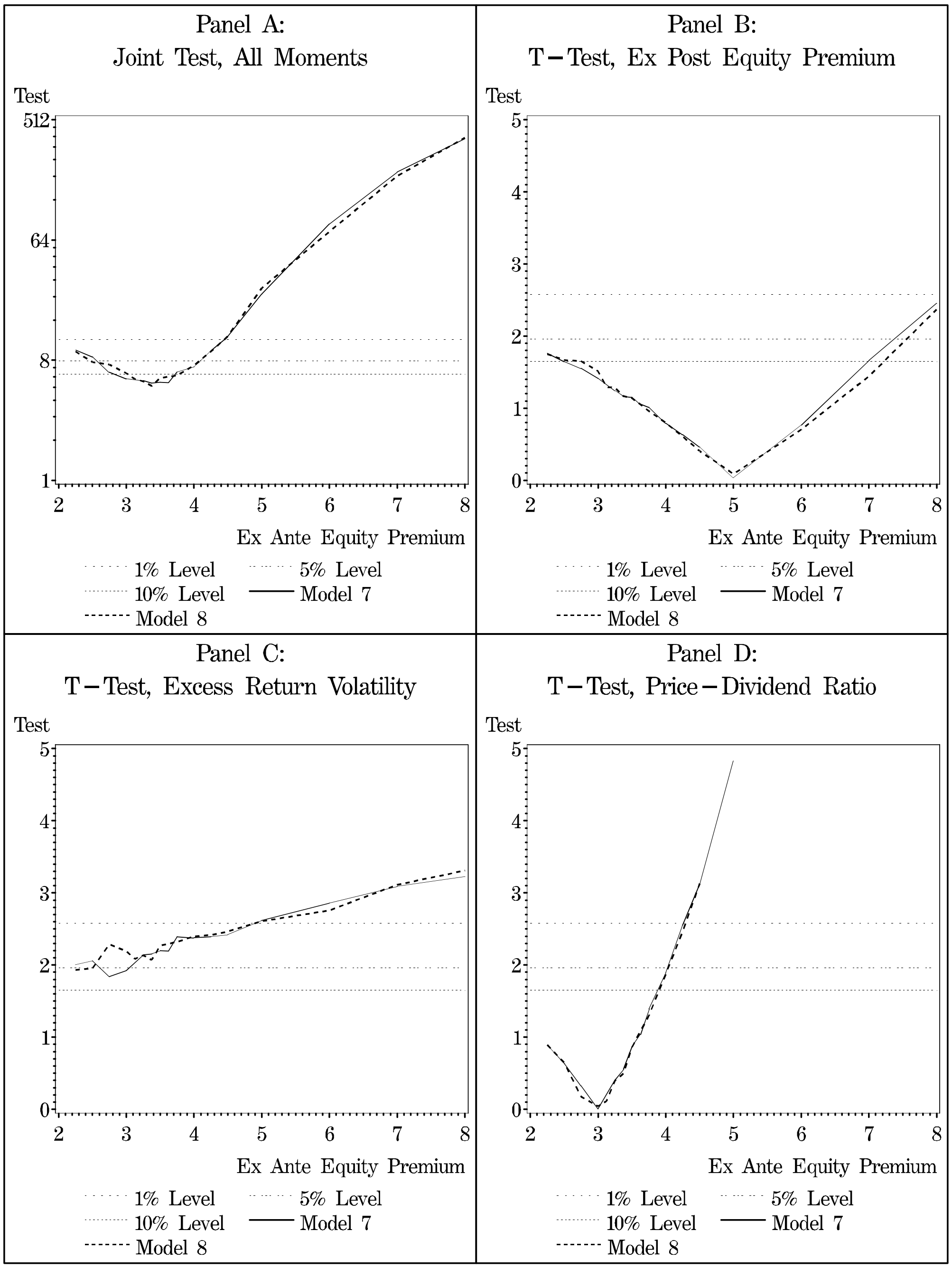
This figure contains plots of test statistics for Models 3 and 4. Panel A plots joint χ^2 tests based on a set of three variables (the ex post equity premium, the mean dividend yield, and the excess return volatility) for various ending values of the ex ante equity premium for each model. In Panel A the vertical axis is plotted on a log scale. The remaining panels contains t-test values corresponding to tests on the individual variables for each of the models: the ex post equity premium in Panel B, the excess return volatility in Panel C, and price-dividend ratio in Panel D. In each panel the critical values of the test statistics corresponding to test significance at the 10%, 5%, and 1% levels are indicated by horizontal lines.

Figure 4: Joint and Individual Tests Statistics for Models 5 and 6



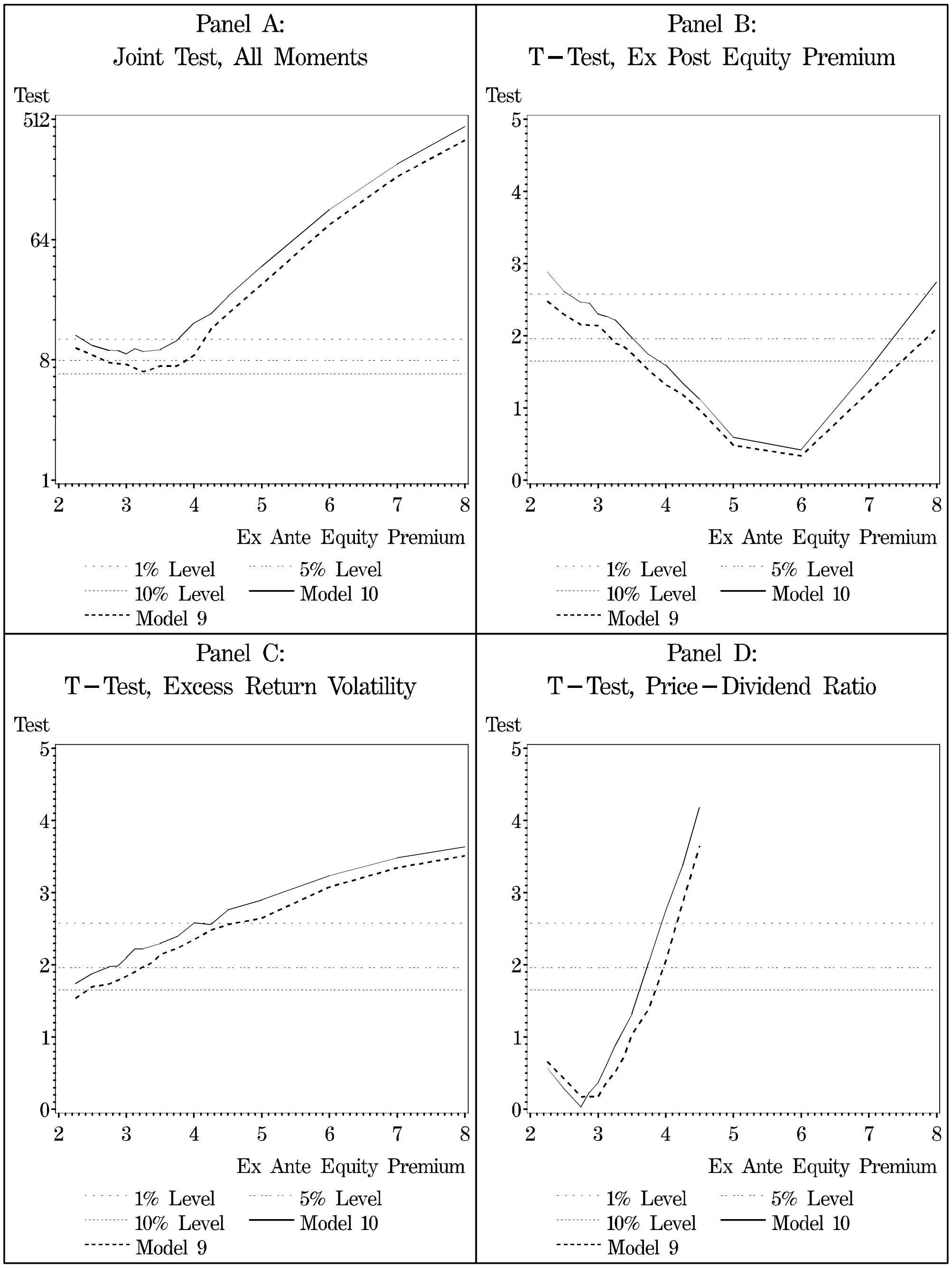
This figure contains plots of test statistics for Models 5 and 6. Panel A plots joint χ^2 tests based on a set of three variables (the ex post equity premium, the mean dividend yield, and the excess return volatility) for various ending values of the ex ante equity premium for each model. In Panel A the vertical axis is plotted on a log scale. The remaining panels contains t-test values corresponding to tests on the individual variables for each of the models: the ex post equity premium in Panel B, the excess return volatility in Panel C, and price-dividend ratio in Panel D. In each panel the critical values of the test statistics corresponding to test significance at the 10%, 5%, and 1% levels are indicated by horizontal lines.

Figure 5: Joint and Individual Tests Statistics for Models 7 and 8



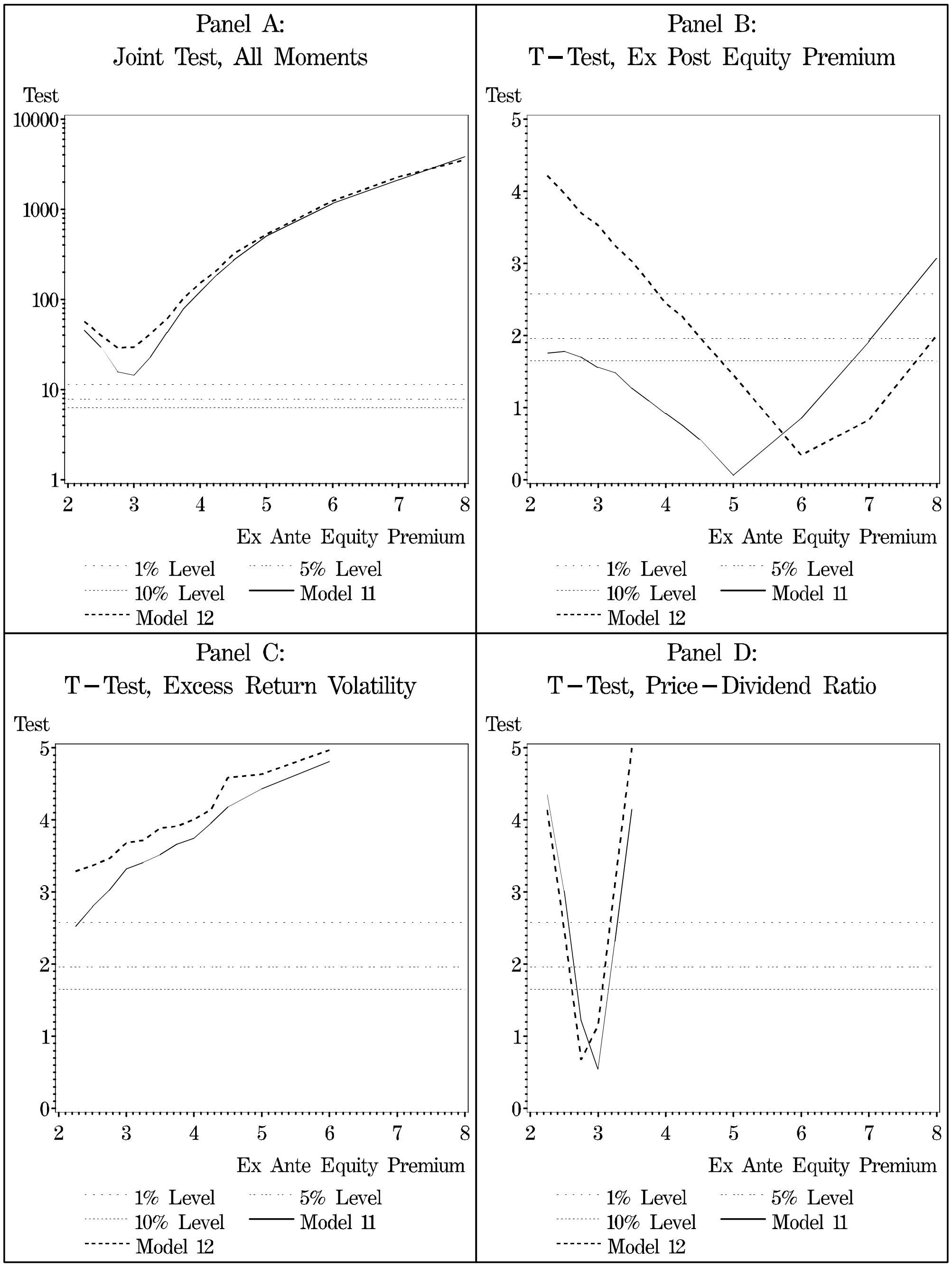
This figure contains plots of test statistics for Models 7 and 8. Panel A plots joint χ^2 tests based on a set of three variables (the ex post equity premium, the mean dividend yield, and the excess return volatility) for various ending values of the ex ante equity premium for each model. In Panel A the vertical axis is plotted on a log scale. The remaining panels contains t-test values corresponding to tests on the individual variables for each of the models: the ex post equity premium in Panel B, the excess return volatility in Panel C, and price-dividend ratio in Panel D. In each panel the critical values of the test statistics corresponding to test significance at the 10%, 5%, and 1% levels are indicated by horizontal lines.

Figure 6: Joint and Individual Tests Statistics for Models 9 and 10



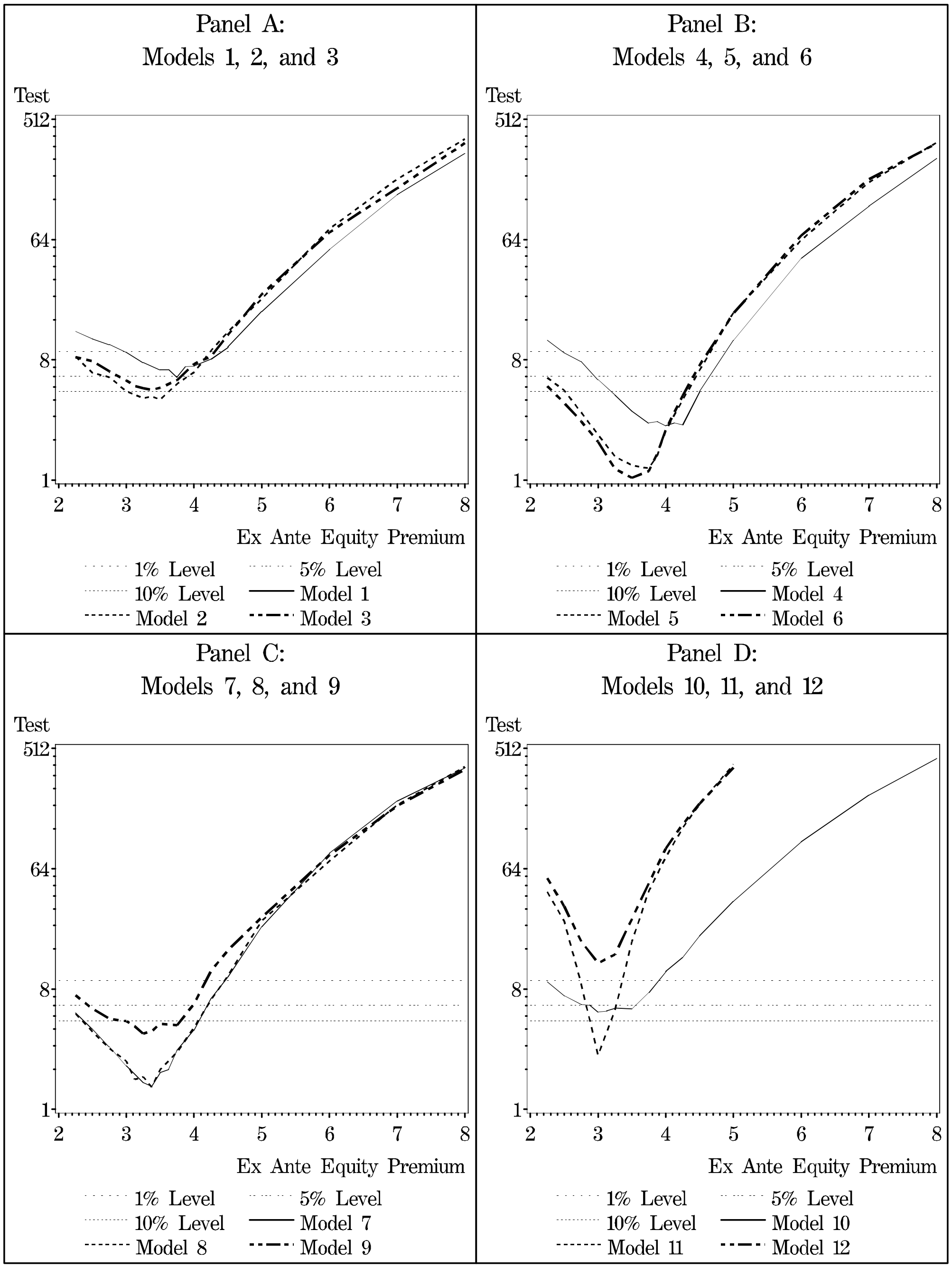
This figure contains plots of test statistics for Models 9 and 10. Panel A plots joint χ^2 tests based on a set of three variables (the ex post equity premium, the mean dividend yield, and the excess return volatility) for various ending values of the ex ante equity premium for each model. In Panel A the vertical axis is plotted on a log scale. The remaining panels contains t-test values corresponding to tests on the individual variables for each of the models: the ex post equity premium in Panel B, the excess return volatility in Panel C, and price-dividend ratio in Panel D. In each panel the critical values of the test statistics corresponding to test significance at the 10%, 5%, and 1% levels are indicated by horizontal lines.

Figure 7: Parameter Estimation Certainty:
Joint and Individual Tests Statistics for Models 11 and 12



This figure contains plots of test statistics for Models 11 and 12. Panel A plots joint χ^2 tests based on a set of three variables (the ex post equity premium, the mean dividend yield, and the excess return volatility) for various ending values of the ex ante equity premium for each model. In Panel A the vertical axis is plotted on a log scale. The remaining panels contains t-test values corresponding to tests on the individual variables for each of the models: the ex post equity premium in Panel B, the excess return volatility in Panel C, and price-dividend ratio in Panel D. In each panel the critical values of the test statistics corresponding to test significance at the 10%, 5%, and 1% levels are indicated by horizontal lines.

Figure 8: Investors' Model Uncertainty
Joint Tests Based on a Subset of Moments for Models 1-12



This figure contains plots of joint χ^2 tests based on a set of two variables, the ex post equity premium and the mean dividend yield, for various ending values of the ex ante equity premium for each model. Panel A presents the test statistics for Models 1, 2, and 3, Panel B presents the test statistics for Models 4, 5, and 6, Panel C presents the test statistics for Models 7, 8, and 9, and Panel D presents the test statistics for Models 10, 11, and 12. The vertical axis of each plot is on a log scale. In each panel the critical values of the test statistics corresponding to test significance at the 10%, 5%, and 1% levels are indicated by horizontal lines.

Are Stocks Cheap? A Review of the Evidence

Fernando Duarte and Carlo Rosa

We surveyed banks, we combed the academic literature, we asked economists at central banks. It turns out that most of their models predict that we will enjoy historically high excess returns for the S&P 500 for the next five years. But how do they reach this conclusion? Why is it that the equity premium is so high? And more importantly: Can we trust their models?

The equity risk premium is the *expected* future return of stocks minus the risk-free rate over some investment horizon. Because we don't directly observe market expectations of future returns, we need a way to figure them out indirectly. That's where the models come in. In this post, we analyze twenty-nine of the most popular and widely used models to compute the equity risk premium over the last fifty years. They include surveys, dividend-discount models, cross-sectional regressions, and time-series regressions, which together use more than thirty different variables as predictors, ranging from price-dividend ratios to inflation. Our calculations rely on real-time information to avoid any look-ahead bias. So, to compute the equity risk premium in, say, January 1970, we only use data that was available in December 1969.

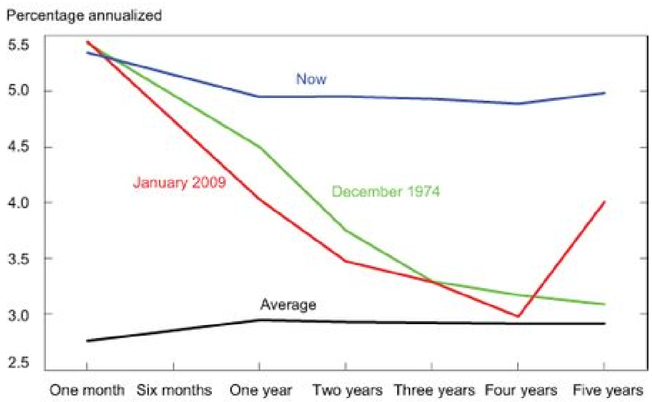
Let's now take a look at the facts. The chart below shows the weighted average of the twenty-nine models for the one-month-ahead equity risk premium, with the weights selected so that this single measure explains as much of the variability across models as possible (for the geeks: it is the first principal component). The value of 5.4 percent for December 2012 is about as high as it's ever been. The previous two peaks correspond to November 1974 and January 2009. Those were dicey times. By the end of 1974, we had just experienced the collapse of the Bretton Woods system and had a terrible case of stagflation. January 2009 is fresher in our memory. Following the collapse of Lehman Brothers and the upheaval in financial markets, the economy had just shed almost 600,000 jobs in one month and was in its deepest recession since the 1930s. It is difficult to argue that we're living in rosy times, but we are surely in better shape now than then.

Today's equity premium has reached a historic high

Sources: Authors' calculations; Barclays; Deutsche Bank; Duke/CFO Business Outlook survey; Federal Reserve Board; Federal Reserve Bank of New York; Goldman Sachs; J.P. Morgan; Nomura; the Center for Research in Security Prices; Federal Reserve Economic Data; Thomson Reuters; the websites of NYU's Aswath Damodaran; Dartmouth's Kenneth French, University of Lausanne's Amit Goyal, University of California at Berkeley's Martin Lettau, Yale's Robert Shiller.

The next chart shows a comparison between those two episodes and today. For 1974 and 2009, the green and red lines show that the equity risk premium was high at the one-month horizon, but was decreasing at longer and longer horizons. Market expectations were that at a four-year horizon the equity risk premium would return to its usual level (the black line displays the average levels over the last fifty years). In contrast, the blue line shows that the equity risk premium today is high irrespective of investment horizon.

The equity premium is elevated at all horizons

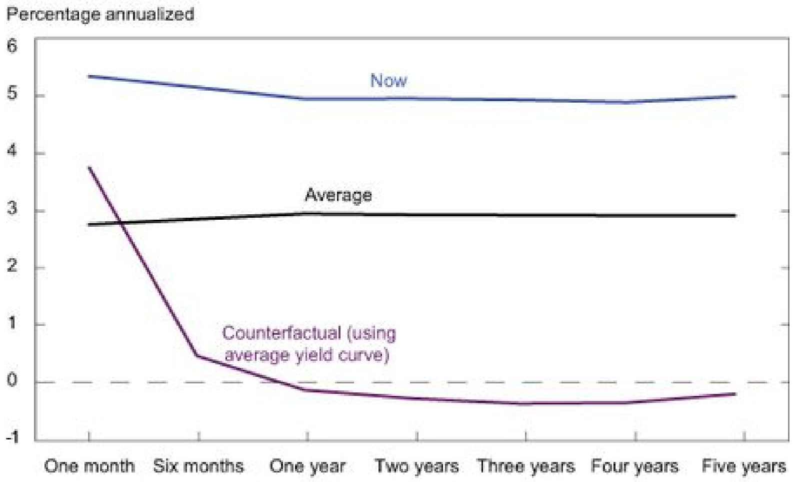


Sources: Authors' calculations; Barclays; Deutsche Bank; Duke/CFO Business Outlook survey; Federal Reserve Board; Federal Reserve Bank of New York; Goldman Sachs; J.P. Morgan; Nomura; the Center for Research in Security Prices; Federal Reserve Economic Data; Thomson Reuters; the websites of NYU's Aswath Damodaran; Dartmouth's Kenneth French, University of Lausanne's Amit Goyal, University of California at Berkeley's Martin Lettau, Yale's Robert Shiller.

Why is the equity premium so high right now? And why is it high at all horizons? There are two possible reasons: low discount rates (that is, low Treasury yields) and/or high current or future expected dividends. We can figure out which factor is more important by comparing the twenty-nine models with one another. This strategy works because some models emphasize changes in dividends, while others emphasize changes in risk-free rates. We find that the equity risk premium is high mainly due to exceptionally low Treasury yields at all foreseeable horizons. In contrast, the current level of dividends is roughly at its historical average and future dividends are expected to grow only modestly above average in the coming years.

In the next chart we show, in an admittedly crude way, the impact that low Treasury yields have on the equity risk premium. The blue and black lines reproduce the lines from the previous chart: the blue is today's equity risk premium at different horizons and the black is the average over the last fifty years. The new purple line is a counterfactual: it shows what the equity premium would be today if nominal Treasury yields were at their average historical levels instead of their current low levels. The figure makes clear that exceptionally low yields are more than enough to justify a risk premium that is highly elevated by historical standards.

The equity premium is high because Treasury yields are low

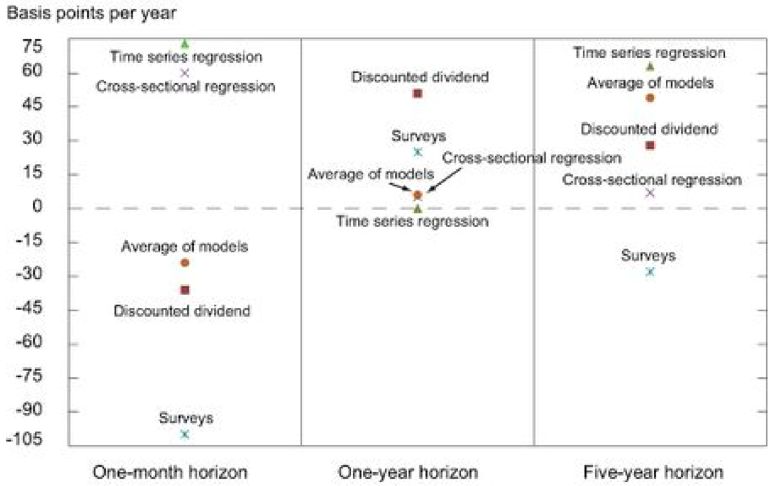


Sources: Authors' calculations; Barclays; Deutsche Bank; Duke/CFO Business Outlook survey; Federal Reserve Board; Federal Reserve Bank of New York; Goldman Sachs; J.P. Morgan; Nomura; the Center for Research in Security Prices; Federal Reserve Economic Data; Thomson Reuters; the websites of NYU's Aswath Damodaran; Dartmouth's Kenneth French, University of Lausanne's Amit Goyal, University of California at Berkeley's Martin Lettau, Yale's Robert Shiller.

But none of this analysis matters if excess returns are unpredictable because the equity risk premium is all about expected returns. So…*are* returns predictable? The jury is still out on this one, and the debate among academics and practitioners is alive and well. The simplest predictive method is to assume that future returns will be equal to the average of all past returns. It turns out that it is remarkably tricky to improve upon this simple method. However, with so many models at hand, we couldn't help but ask if any of them can, in fact, do better.

The table below gives the extra returns that investors could have earned by using the models instead of the historical mean to predict future returns. For investment horizons of one month, one year, and five years, we pick the best model in each of the four classes we consider together with the weighted average of all twenty-nine models. We compute these numbers by assuming that investors can allocate their wealth in stocks or bonds, and that they are not too risk-averse (for the geeks again, we solved a Merton portfolio problem in real time assuming that the coefficient of relative risk aversion is equal to one). The table shows positive extra returns for most of the models, especially at long horizons.

Model performance is varied, but better at longer horizons



Sources: Authors' calculations; Barclays; Deutsche Bank; Duke/CFO Business Outlook survey; Federal Reserve Board; Federal Reserve Bank of New York; Goldman Sachs; J.P. Morgan; Nomura; the Center for Research in Security Prices; Federal Reserve Economic Data; Thomson Reuters; the websites of NYU's Aswath Damodaran; Dartmouth's Kenneth French, University of Lausanne's Amit Goyal, University of California at Berkeley's Martin Lettau, Yale's Robert Shiller.

Notes: We tested twenty-nine models in four classes (surveys, dividend-discount models, cross-sectional regressions, and time series regressions) over three investment horizons. In this chart, we plot the single best-performing model in each category. We also show how the optimal weighted average (the first principal component) of all models performs.

At face value, this result means that the models are actually helpful in forecasting returns. However, we should keep in mind some of the limitations of our analysis. First, we have not shown confidence intervals or error bars. In practice, those are quite large, so even if we could have earned extra returns by using the models, it may have been solely due to luck. Second, we have selected models that have performed well in the past, so there is some selection bias. And of course, past performance is no guarantee of future performance.

Disclaimer

The views expressed in this post are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

The Equity Risk Premium: A Consensus of Models

By [Fernando Duarte] and [Carlo Rosa]

Abstract

We estimate the equity risk premium by combining information from twenty models. Our main finding is that there is broad agreement across models that the equity premium reached historical heights in July 2013 even when the models are substantially different from each other and use more than one hundred different economic variables. Our preferred estimator places the one-year-ahead equity premium in July 2013 at 14.5 percent, the highest level in fifty years and well above the 10.5 percent that was reached during the financial crisis in 2009. The models also show broad agreement that the term structure of equity risk premia is high and flat: expected excess returns at all foreseeable horizons are just as high as at the one-year horizon. A high equity premium that is not expected to mean-revert in the near future is an unprecedented phenomenon. Because expected dividend growth has not been above average in 2013, we conclude the high equity premium is mostly due to unusually low discount rates at all horizons.

1. Introduction

The equity risk premium —the expected return of stocks in excess of the risk-free rate— is a fundamental quantity in all of asset pricing, both for theoretical and practical reasons. It is a key measure of aggregate risk-aversion and an important determinant of the cost of capital for corporations, saving decisions of individuals and budgeting plans for governments. Recently, the equity risk premium (ERP) has also returned to the forefront of policymaking as a leading indicator of the evolution of the economy, a potential explanation for the jobless recovery and a gauge of financial stability. As an indicator of future activity, a high ERP at short horizons tends to be followed by higher GDP growth, higher inflation and lower unemployment, thus informing both fiscal and monetary decisions. Bloom (2009) and new research by Duarte, Kogan and Livdan (2013) point to large effects of the ERP on real aggregate investment, a component that has been lagging in the present recovery compared to policymakers' forecasts in the current cycle and actual performance in past cycles. As a potential explanation of the jobless recovery, Hall (2013) and Kuehn, Petrosky-Nadeau and Zhang (2012) have proposed that increased risk-aversion has prevented firms from hiring as much as would be expected in today's macroeconomic environment. From the perspective of financial stability, the so-called "great rotation" from bonds to stocks could be exacerbated in speed and magnitude if the ERP is persistently high. A sudden flow of money out of the bond market into stocks could spell large capital losses for fixed income investors, including the Federal Reserve. Low returns in other asset classes could provide incentives for investors to engage in potentially unsafe "reach for yield" either through excessive use of leverage or through other forms of risk-taking. The ERP is also important from the perspective of unconventional monetary policy: a high ERP may make the portfolio channel of Large Scale Asset Purchases more effective because it further increases the demand for risky assets.

In this article, we estimate the ERP by combining information from twenty models that are prominently used by practitioners and featured in the academic literature. Our main finding is that there is broad agreement across models that the ERP has reached historical heights even when the models are substantially different from each other and use more than one hundred different economic variables. Our preferred estimator places the one-year-ahead ERP in July 2013 at 14.5 percent, the highest level in fifty years and well above the 10.5 percent that was reached during the financial crisis in 2009. The models also show broad agreement that expected excess returns at all foreseeable horizons are just as high as at the one-year horizon. A high equity premium that is not expected to mean-revert in the near future is an unprecedented phenomenon.

In addition to estimating the level of the ERP, it is useful for policymakers and other economic agents to know why the ERP is high. We conclude the ERP is high at all foreseeable horizons because Treasury yields are unusually low at all maturities. In other words, the term structure of equity premia is high and flat because the term structure of interest rates is low and flat. Current and expected future dividend and earnings growth play only a minor role. A high ERP caused by low bond yields indicates that a stock market correction is likely to occur only when bond yields start to rise. Additionally, a bond-driven ERP makes it more unlikely that irrational exuberance can take hold in equity markets, especially at times of increasing expectations for a steepening of the yield curve. Another implication of a bond-driven ERP is that we should no longer rely on traditional indicators of the ERP like the price-dividend or price-earnings ratios, which all but ignore the term structure of risk-free rates.

As a second contribution, we evaluate the performance of different ERP models. Statements about the implications of a high ERP are valid only to the extent that expected returns predict future realized returns. For the models we consider, predictability is weak but present. We first categorize the twenty models we study into five groups: predictors that use historical mean returns, dividend-discount models, cross-sectional regressions, time-series regressions and surveys. To assess whether models can indeed predict returns, we regress realized excess returns on the corresponding ERP given by the models. We then use the out-of-sample R-squared for these regressions as a measure of success. We find that dividend-discount models perform best at short horizons, while cross-sectional regressions perform best at longer horizons. Combining all models into a single principal component — our preferred measure—reduces noise. A mean-variance investor with unit risk aversion using the principal component as an investment signal would have earned 15 percentage points more over the last fifty years (30 basis points per year) than if she had assumed expected returns are equal to past mean returns.

2. The Equity Risk Premium: Definition

Conceptually, the ERP is the compensation investors require to make them indifferent between holding the risky market portfolio and a risk-free bond. Because this compensation depends on the future performance of stocks, the ERP incorporates expectations of future stock market returns, which are not directly observable. At the end of the day, any model of the ERP is a model of investor expectations. Additionally, it is not clear what truly constitutes the market return and the risk-free rate in the real world. In practice, the most common measures of market returns are given by broad stock market indices, like the S&P 500 or the Dow Jones Industrial Average, but those indices do not include the whole universe of traded stocks and miss several other components of wealth. Even if we included all traded stocks, we still have several choices to make, such as whether to use value or equal-weighted indices, or whether to exclude penny stocks or rarely traded stocks. A similar problem arises with the risk-free rate. While we almost always use Treasury yields as measures of risk-free rates, they are not completely riskless since nominal Treasuries are exposed to inflation and liquidity risks. In this paper, we follow common practice and always use the S&P 500 as a measure of stock market prices and either nominal or real Treasury yields as risk-free rates. The models we consider differ only in how expectations are computed.

While implementing the concept of the ERP has pitfalls, we can precisely define the ERP mathematically. First, we decompose stock returns into an expected component and an unpredictable random component:

$$R_{t+k} = E_t[R_{t+k}] + error_{t+k} \tag{1}$$

In equation (1), R_{t+k} are net *realized* returns between t and $t+k$, $E_t[R_{t+k}]$ are the returns that were expected from t to $t+k$ using information available at time t and $error_{t+k}$ is a mean-zero random variable that is unknown at time t but is realized at $t+k$. The ERP at time t for horizon k is defined as

$$ERP_t(k) = E_t[R_{t+k}] - R^f_{t+k} \tag{2}$$

where R^f_{t+k} is the net risk-free rate for investing from t to $t+k$ (which, being risk-free, is known at time t).

This definition shows three important aspects of the ERP. First, because the unexpected component $error_{t+k}$ is stochastic and orthogonal to expected returns, the ERP is always less volatile than realized excess returns. Therefore, while realized stock returns are very volatile compared to bonds, we expect good ERP estimates to be somewhat smoother. Second, the ERP itself is a random variable, since expectations can change through time when new information arrives. Third, the ERP has an investment horizon k embedded in it, since we can consider expected excess returns over, say, one month, one year or five years from today. If we fix t, and let k vary, we trace the *term structure* of the equity risk premium.

3. Data

In constructing all estimates of the ERP we use over one hundred variables. The sources and definitions are standard. The nominal and real price, earnings and dividends for the S&P 500 are from Shiller. Inflation, the "cyclically adjusted price-earnings ratio" and the ten-year nominal treasury yield are also from Shiller. Expected earnings per share are mean analyst forecasts from Thomson Reuters I/B/E/S. Nominal bond yields for all maturities except 10-years[1] and all TIPS yields are from the Federal Reserve Board. Fama-French and momentum factors and portfolios are from Professor French's website. Corporate bond spreads and the NBER recession indicator are from the St. Louis Federal Reserve (FRED). Book value per share for the S&P 500 is from Compustat. Debt issuance and equity issuance are from Jeffrey Wurgler's website. Consumption to wealth ratio measured by *cay* is from Martin Lettau's website (Ludvigson and Lettau, 2001). ERP estimates from CFOs are from the Duke CFO survey. The sentiment measure of Baker and Wurgler is from Jeffrey Wurgler's website. Professor Damodaran's estimates of the ERP are from his website. All variables are monthly from January 1960 to July 2013, except for *cay* and CFO surveys, which are quarterly, and book value per share and Damodaran's ERP estimates, which are annual. Other variables are constructed using the variables mentioned before. A detailed description is in Appendix A.

4. Models of the Equity Risk Premium

We classify models of the ERP into five categories and discuss their advantages and disadvantages. We also describe in detail the models we use within each category and how to obtain a term-structure of the ERP for each one. Of course, there are many more models of the ERP than the ones we consider. We selected which models to include in our study based on the recent academic literature and widespread use by practitioners. All models are constructed in real time[2], so that an investor who lived through the

[1] Except for the 10-year yield, which, as described above, is from Shiller. We use Shiller's 10-year yield for ease of comparability with the existing literature. Results are virtually unchanged if we use all yields, including the 10-year yield, from the Federal Reserve Board.

[2] The one exception is Adrian, Crump and Moench's (2013) cross-sectional model, which is constructed using full-sample regression estimates. Our out-of-sample predictability results are essentially unchanged if we omit this model from the analysis.

sample would have been able to construct the measures at each point in time using available information only. This helps avoid look-ahead bias and makes the out-of-sample evaluation of the models meaningful.

4.1 Historical mean

The easiest approach to estimating the ERP is to assume it is equal to the historical mean of realized market returns in excess of the contemporaneous risk-free rate. The main choice is how far into the past to go when computing the historical mean. This model is very simple and, as we show in Section 8, quite difficult to improve upon when considering out-of-sample performance measures. The main drawbacks are that it is purely backward looking, and assumes that the future will behave like the past, i.e. it assumes the conditional mean of excess returns is not time-varying, giving very little time-variation in the ERP.

To trace the term structure of the ERP using the historical mean method, we simply use returns computed over different horizons and the corresponding maturity risk-free rate before taking the mean.

Model 1: We compute the historical mean going as far back into the past as the data allows.

Model 2: Same as Model 1 but we compute the mean using the previous 5-years of data only (i.e. we use a backward looking 5-year rolling window).

4.2 Dividend discount models (DDM)

All DDM start with the basic intuition that the value of a stock is determined by no more and no less than the cash flows it produces for its shareholders (Gordon 1962). Today's stock price should then be the sum of all expected future dividends, discounted at an appropriate rate to take into account their riskiness and the time value of money. The formula that reflects this intuition is

$$P_t = E_t \sum_{k=0}^{\infty} \frac{D_{t+k}}{\rho_{t+k}} \tag{3}$$

where E_t is the conditional expectations operator, P_t is the current price of the stock, D_t is the current level of dividends, D_{t+k} is the level of dividends k periods from now, and ρ_{t+k} is the discount rate for time $t + k$. The discount rate can be decomposed into

$$\rho_{t+k} = 1 + R_{t+k}^{f} + ERP_t(k) \tag{4}$$

When using a DDM, we refer to $ERP_t(k)$ as the *implied* ERP, since we plug in observed or estimated values for the price, dividends and the risk-free rates, and figure out what value of $ERP_t(k)$ makes the right-hand side equal to the left-hand side in equation (3). In this framework, the risk-free rate captures the discounting associated with the time value of money and the ERP captures the discounting associated with the riskiness of the dividends.

DDM are forward looking and are consistent with no arbitrage. In fact, equation (3) is an equilibrium condition that must hold in any bubble-free economy with no arbitrage. Another advantage of DDM is that they are easy to implement. A drawback of DDM is that the results are sensitive to how we measure expectations of future dividends. In addition, ignoring the bubble term, i.e. assuming that

$$\lim_{k\to\infty} E_t\left[\frac{D_{t+k}}{\rho_{t+k}}\right] = 0$$

may impute a higher ERP whenever a bubble is present but not considered in the model.

Even though DDM do not require the term structure of the ERP to be flat, in practice all DDM assume that $ERP_t(k) = ERP_t(j)$ for all k and j. With a single ERP measure for all horizons, equation (3) pins down the ERP completely, while if we had different ERP estimates for different horizons, equation (3) would become a single equation in several unknowns and the ERP would not be identified.

Model 3: The simplest DDM assumes a constant growth rate of dividends and a flat yield curve in addition to a flat term structure of the ERP (Gordon 1962). Under these assumptions, equation (3) becomes

$$P_t = \sum_{k=0}^{\infty} \frac{D_t(1+g)^k}{(1+R_t^f + ERP_t)^k} = \frac{D_t}{R_t^f + ERP_t - g}$$

Solving for the ERP gives

$$ERP_t = \frac{D_t}{P_t} - (R_t^f - g) \tag{5}$$

Note that even though the term-structure of the ERP is assumed to be flat, this model does not assume a constant ERP or a constant risk-free rate. In practice, there are several ways to operationalize equation (5). Model 3, called[3] the "Fed Model", uses the nominal ten-year Treasury yield as an estimate of $R_t^f - g$ and current earnings E_t as a proxy for current dividends D_t.

Model 4: The "Shiller model". Same as Model 3 but uses Shiller's cyclically adjusted price-earnings ratio (CAPE) as a proxy for the price-dividend ratio. CAPE is the current price of the S&P 500 divided by a trailing twelve month average of earnings.

Model 5: Same as Model 3, but uses the real ten-year Treasury yield as an estimate of $R_t^f - g$ (computed as the ten-year nominal Treasury rate minus the ten year breakeven inflation implied by TIPS). There are two typical justifications for this choice. First, in the long run, the growth rate of dividends g should be at least approximately equal to breakeven inflation. Second, the dividend-price ratio D_t/P_t, being the ratio of two nominal variables, is a real variable. Thus, it should be compared to the real risk-free rate and not to the nominal one as in Model 3.

Model 6: Same as Model 3, but uses one-year ahead expected earnings as a proxy for dividends. The usual justification is that including future expectations should better capture the forward-looking nature of the DDM.

Model 7: A variation in the assumptions in Models 3, 4 and 5: it uses the nominal ten-year Treasury yield and one-year ahead expected earnings.

---

[3] The name "Fed Model" was coined by Ed Yardeni, at Deutsche Morgan Grenfell, in reference to a report issued in 1997 by the Federal Reserve that used the model. However, the Federal Reserve has never endorsed this model. See Asness (2003) for a critical view of the "Fed Model".

Model 8: A two-stage DMM from Panigirtzoglou and Loeys (2005) where the first stage corresponds to the first five years, and the second stage corresponds to years 6 and onwards. In this case, formula (3) becomes[4]

$$P_t = \frac{D_t[1 + g_{LR} + 5(g_{SR} - g_{LR})]}{R_t^f + ERP_t - g_{LR}}$$

where R_t^f is the ten year nominal Treasury yield; D_t is estimated by the current (observable) level of earnings-per share multiplied by a payout ratio assumed to be 50%; g_{LR} is the long-run estimate for earnings growth and assumed to be 2.2 percent; g_{SR} is the estimated growth rate of earnings over the first five years, which is estimated by using the fitted values in a regression of average realized earnings growth over the last five years on its lag and lagged earnings-price ratio. The main advantage of having two stages instead of a single one (as in Models 3 through 7) is that it allows for changes in the growth rate of dividends, a useful feature when growth rates are far away from their long-run level.

Model 9: A multi-stage DDM constructed by Damodaran (2012). We simplify equation (3) by assuming there are 6 stages. Each of the first five stages corresponds to each of the first five years, while the last stage corresponds to years six and onwards. Dividends are assumed to grow at a rate g_t for each of the first five stages, and then at a rate equal to the ten year nominal Treasury yield for the final stage. The discount rate is assumed to be constant over different horizons, so that $\rho_{t+k} = \rho_t^k$. With these assumptions, equation (3) becomes

$$P_t = \sum_{k=1}^{5} \frac{D_t(1 + g_t)^k}{\rho_t^k} + \frac{D_{t+6}(1 + g_t)^6}{(\rho_t - R_t^f)\rho_t^5} \tag{6}$$

where $\rho_t = 1 + R_t^f + ERP_t$ and R_t^f is the ten year nominal Treasury yield. Given P_t, D_t, R_t^f and g_t, equation (6) determines a unique ERP_t.

Model 10: Is the same as Model 9 –and also proposed by Damodaran (2012)— but includes stock buybacks in cash flows. The idea is that investors care about total cash flows, not just dividends, and that buybacks are significant enough to affect measures of the ERP. In practice, we use free-cash-flow-to-equity as a proxy for dividends plus stock buybacks. Damodaran (2012) estimates that buybacks can increase the ERP by one to four percentage points per year.

4.3 Cross-sectional regressions

This method exploits the variation in returns and exposures to the S&P 500 of different assets to infer the ERP.[5] Intuitively, this method finds the ERP by answering the following question: what is the level of the ERP that makes expected returns of a variety of stocks consistent with their exposure to the S&P 500? Because we need to explain the relationship between returns and exposures for multiple assets with a single value for the ERP (and perhaps a small number of other controls), this model imposes tight restrictions on the estimation of the ERP.

The first step is to find the exposures of assets to the S&P 500 by estimating an equation of the following form:

[4] For a derivation, see Fuller and Hsia (1984).
[5] See Polk, Thompson and Vuolteenaho (2006) and Adrian, Crump and Moench (2012) for a detailed description of this method.

$$R_{t+k}^{i} - R_{t+k}^{f} = \alpha_{t+k}^{i} \times state\ variables_{t+k} + \beta_{t+k}^{i} \times risk\ factors_{t+k} + idiosyncratic\ risk_{t+k}^{i} \quad (7)$$

In equation (7), R_{t+k}^{i} is the realized return on a stock or portfolio from time t to $t+k$. $State\ variables_{t+k}$ are any economic indicators that help identify changes in the investment opportunity set (possibly including a constant). $Risk\ factors_{t+k}$ are any measures of systematic contemporaneous co-variation in returns across all stocks or portfolios. Finally, $idiosyncratic\ risk_{t+k}^{i}$ is the component of returns that is particular to each individual stock or portfolio that is not explained by $State\ variables_{t+k}$ or $Risk\ factors_{t+k}$. Examples of state variables are inflation, unemployment, the term spread, the yield spread between Aaa and Baa bonds and the S&P 500's dividend-to-price ratio. It is crucial that we include the excess return on the S&P 500 as a risk-factor in the estimation so that we can infer the ERP. Other risk-factors usually used are the Fama-French (1992) factors and the momentum factor of Carhart (1997). The value of α_{t+k}^{i} gives the strength of asset-specific return predictability and β_{t+k}^{i} gives the asset-specific risk exposures we are trying to estimate. For the cross-section of assets, we can use the whole universe of traded stocks, a subset of them, or portfolios of stocks grouped, for example, by industry, size, book-to-market or recent performance.

The second step is to find the ERP associated with the S&P 500 by estimating the cross-sectional equation

$$R_{t+k}^{i} - R_{t+k}^{f} = ERP_{t}(k) \times \hat{\beta}_{t+k}^{i} \quad (8)$$

where $\hat{\beta}_{t+k}^{i}$ are the values found when estimating equation (7). Equation (8) attempts to find the single number $ERP_{t}(k)$ (or vector of numbers, if we have more than one risk factor) that makes exposures β_{t+k}^{i} consistent with realized excess returns of all stocks or portfolios considered. The term structure of the ERP is obtained by computing returns over different horizons on the left hand side of equations (7) and (8).

One advantage of the cross-sectional regression method is that it uses more asset prices than other models, which provide more independent information about the ERP. Cross-sectional regressions also have sound theoretical foundations, since they are one way to implement Merton's (1973) Intertemporal Capital Asset Pricing Model (ICAPM). Finally, this method nests many of the other models considered. The two main drawbacks of this method are that results are dependent on what portfolios, state variables and risk factors are used and that it is not easy to implement.

Model 11: The most widely used cross-sectional model is the Fama-French model (Fama and French 1992). The only state variable is a constant, and there are three risk factors: the returns on the market portfolio, a size portfolio and a book-to-market portfolio. Equation (7) is estimated by running rolling OLS regressions over the previous five years, and equation (8) is estimated by OLS without a constant[6].

Model 12: Same as Model 11, but includes momentum as an additional risk factor (Carhart 1997).

Model 13: Same as Model 12, but also includes inflation as a risk factor, which has been shown to account for a substantial part of the equity premium beyond the four factors of Carhart's model (Duarte 2013). Additionally, the time-varying coefficients α_{t+k}^{i} and β_{t+k}^{i} are estimated with the non-parametric kernel estimator of Ang and Kristensen (2012).

Model 14: This model is from Adrian, Crump, and Moench (2012). The state variables are the dividend yield, the default spread, and the risk free rate, which are commonly thought to capture changes in the

[6] Using OLS *with* a constant is an equally valid procedure; whether to include a constant depends on the familiar tradeoff between efficiency and robustness (Cochrane 2001).

investment opportunity set. The inclusion of these state variables allows the model to capture dynamics of the pricing kernel not captured by Models 11 through 13. The risk free rate is the one-month Treasury bill rate; the dividend yield is for the S&P500; and the default spread is calculated as the difference between Moody's seasoned Baa corporate bond yield and the 20-year Treasury bond yield at constant maturity. The market is the single risk factor. The model is estimated using a three step regression approach. First, the market return is orthogonalized with respect to the state variables and the residual of that regression is the considered the risk factor. Then each stock or portfolio's excess return is regressed on the lagged state variables and the risk factor to obtain the coefficients α^{i}_{t+k} and β^{i}_{t+k}. Finally, the ERP is obtained by estimating equation (8) using OLS.

4.4 Time-series regressions

This method uses the relationship between economic variables and stock returns to estimate the ERP. The idea is to run a linear regression of realized excess returns on lagged "fundamentals":

$$R_{t+k} - R^{f}_{t+k} = a + b \times Fundamental_t + error_t \tag{9}$$

Once estimates $\hat{a}$ and $\hat{b}$ for a and b are obtained, the ERP is obtained by ignoring the error term:

$$ERP_t(k) = \hat{a} + \hat{b} \times Fundamental_t \tag{10}$$

In other words, we estimate only the forecastable or expected component of excess returns. This method attempts to implement equations (1) and (2) as directly as possible in equations (9) and (10), with the assumption that "fundamentals" are the right sources of information to look at when computing expected returns and that the conditional expectation is a linear function.

The use of time-series regression requires minimal assumptions; there is no concept of equilibrium and no absence of arbitrage necessary for the method to be valid. In addition, implementation is quite simple, since it involves running univariate OLS regressions. The challenge of this method is to select the variables to include in the right-hand side of equation (9), since results can change substantially depending on what fundamental variables are used. In addition, including more than a single variable gives poor out-of-sample predictions even if economic theory may suggest a role for many variables to be used as predictors[7]. Finally, time-series regressions ignore information in the cross-section of stock returns.

The term structure of the ERP, as equations (9) and (10) suggest, is easily obtained in this method by simply running the predictive regressions with excess returns computed over different horizons.

Model 15: This model uses the dividend-price ratio as the only predictive variable. The key rationale is that the dividend-price ratio is first-order stationary so that it should eventually return to its long-run mean[8]: Values of the dividend-price ratio above its mean should forecast either low returns or high dividends going forward (and vice-versa for low values). Empirically, a high dividend-price ratio forecasts higher returns, not lower dividends, so the price-dividend ratio contains information about the ERP (Cochrane 2011).

Model 16: Same as Model 15, but uses the twelve predictive variables proposed by Goyal and Welch (2008). We use each variable independently and all of them together. At each point in time, we select

[7] Goyal and Welch (2008).
[8] See Lettau and Van Nieuwerburgh (2008) for an argument against first-order stationarity and its implications for predictability of returns.

the specification that performs the best out-of-sample (see Section 7 for a detailed description of how we do this) and use that specification for the next period. In the following period, we repeat the procedure; it is possible that this method uses different predictors depending on which one is performing best at each point in time.

Model 17: Same as Model 16, but imposes two restrictions on the estimation. First, the coefficient b in equation (9) is replaced by zero if it has the "wrong" theoretical sign. For example, if the price-dividend ratio has a negative coefficient, then we replace b by zero. Second, we replace the estimate of the ERP by zero if the estimation otherwise finds a negative ERP. These two restrictions are imposed one at a time and then together, and considered for the same twelve predictive variables considered in Model 16. The best specification at time *t* is used for prediction of *t+k* returns, so specifications can be changing over time. This model is advocated by Campbell and Thompson (2008), who argue that the restrictions, being based on theory, should improve estimation efficiency compared to unrestricted estimation.

Model 18: Uses as predictors the price-dividend ratio adjusted by the growth rate of earnings RE_t, dividends RD_t or stock prices RP_t:

$$RE_t = \frac{D_t}{P_{t-1}} + \left(\frac{E_t}{D_{t-1}}\right)\frac{CPI_{t-1}}{CPI_t} - 1 \tag{11}$$

$$RD_t = \frac{D_t}{P_{t-1}} + \left(\frac{D_t}{D_{t-1}}\right)\frac{CPI_{t-1}}{CPI_t} - 1 \tag{12}$$

$$RP_t = \frac{D_t}{P_{t-1}} + \left(\frac{P_t}{P_{t-1}}\right)\frac{CPI_{t-1}}{CPI_t} - 1 \tag{13}$$

where D_t are dividends, E_t are earnings, P_{t-1} is the lagged price of the S&P 500 and CPI_t is the consumer price index. We also consider the three measures constructed above, but subtracting the ten year nominal Treasury yield from each of them. The idea behind these measures is to impose –rather than assume— that stationary variables must eventually return to their long-run mean. As in models 16 and 17, at time t we use the predictor that has the best out-of-sample performance until $t-1$, which leads to different measures being used at different points in time. This model was proposed by Fama and French (2002) who argue that stock returns have been too high compared to dividend or earnings growth, and therefore must have been in part due to luck (positive shocks). A way to account for this sample-specific realization is to "correct" the dividend-price ratio as in equations (11), (12) and (13).

Model 19: The predictor is Baker and Wurgler's (2007) sentiment measure. The measure is constructed by finding the most predictive linear combination of five variables: the closed-end fund discount, NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium. Baker and Wurgler (2007) have a more detailed explanation.

4.5 Surveys

The survey approach consists in asking economic agents what they think the ERP is. Surveys incorporate the views of many people, some of which are very sophisticated and/or make real investment decisions based on the level of the ERP. Surveys should also be good forecasters of the ERP because in principle stock prices are determined by supply and demand of investors such as the ones taking the surveys. On the other hand, Greenwood and Shleifer (2012) document that investor expectations of future stock market returns are positively correlated with past stock returns and with the current level of the stock market, but strongly *negatively* correlated with model-based expected returns and future realized stock market returns. Other studies such as Easton and Sommers (2007) also argue that survey measures of the ERP can be systematically biased.

The term structure of the ERP can only be obtained from surveys to the extent that questions are asked about the ERP at different horizons. To the authors' knowledge, the only consistent survey with a long enough time-series for analysis that asks about point estimates of the ERP at different horizons is the Duke CFO survey by Graham and Harvey (2012), which we use for the next model.

Model 20: Chief financial officers (CFOs) are asked about the one and ten-year-ahead ERP. A typical question in the survey is the following:

> *On November 19, 2007 the annual yield on 10-yr treasury bonds was 4.1 percent. Please complete the following: Over the next year, I expect the average annual S&P 500 return will be:*

The survey has grown over time and now has around 600 respondents. We take the mean of all responses as our measure of the ERP[9]. We construct the term-structure of the ERP by linearly interpolating the one and ten-year-ahead ERP estimates given by respondents. For this model, we do not construct ERP measures for horizons shorter than a year.

5. The Equity Risk Premium has Reached a Historic High

We summarize the behavior of the twenty models we consider by their first principal component. Let X be the matrix containing the demeaned ERP estimates at a monthly horizon from the different models we consider, with columns corresponding to models and rows corresponding to observation periods. The matrix is a 643-by-20 matrix, since we have 643 monthly observations and 20 models. The first principal component is the eigenvector of the variance-covariance matrix of X associated with the largest eigenvalue. Because X was demeaned, this principal component has mean zero. We take as our preferred ERP estimate the sum of the first principal component and the unconditional mean of ERP estimates across all models (i.e. the average of all elements of X). We repeat this process using ERP estimates at different horizons to obtain a single ERP time-series for each horizon. We call these estimates our *preferred measures*. The share of the variance explained by these measures ranges between 81% and 94%, suggesting that they are good summary statistics for the behavior of the models.

One challenge that arises in computing the principal component is that the matrix X has missing observations, either because some models can only be obtained at frequencies lower than monthly or because the necessary data is not available for all time periods (Appendix A contains a detailed description of when this happens). To overcome this challenge, we use an iterative linear projection method[10]. On the first iteration, we make a guess for the principal component and regress the non-missing elements of each row of X on the guess and a constant. We then find the first principal component of the variance-covariance matrix of the fitted values of these regressions, and use it as the guess for the next iteration. The process ends when the norm of the difference between consecutive estimates is small enough.

Figure 1 displays our preferred measures for the one-month and one-year-ahead ERP in blue and red, respectively. Recessions are indicated by shaded bars. The correlation between the two measures is

[9] Taking the median does not substantially alter results.
[10] We thank Richard Crump for suggesting this method and providing code for its implementation.

86%, but we do see that the one-month ahead ERP is sometimes above and sometimes below the one-year-ahead ERP, indicating that the slope of the term structure of the ERP is time-varying. As expected, the ERP measures tend to peak during financial turmoil, recessions and periods of low real GDP growth or high inflation. The ERP tends to bottom out after periods of sustained bullish stock markets and high real GDP growth.

The one year ahead ERP is at 14.5 percent in July of 2013, the highest it has even been. The one month ahead ERP is at 11.5 percent in July 2013, nearing the record levels obtained in February 2009, July 2012 and the early 1980s, but still below the peak of 15 percent in September of 1974.

The current high levels of the ERP are unusual in that we are not currently in a recession and we have just experienced an extended period of high stock returns, with 60 percent returns since July 2010 and almost 20 percent since the beginning of 2013. During previous periods, the ERP has always decreased during periods of sustained high realized returns. This is also the only period in which the ERP is elevated and the one-year ahead ERP is significantly higher than the one month ahead ERP.

Figure 2 displays in red the standard deviation of one-year-ahead ERP estimates across models for each time period. The standard deviation has been steadily decreasing since 2000 except for a few months during the financial crisis and has reached an all-time low in the last three months. A low standard deviation can be interpreted as models displaying a high degree of agreement – in this case, agreement that the ERP is high. Figure 2 also shows the reason for the recent increase in agreement and in the ERP by plotting the 25^{th} and 75^{th} percentiles of the distribution of models in blue and green, respectively. The interquartile range –the difference between the 25^{th} and 75^{th} percentiles— has compressed, mostly because the models in the bottom of the distribution have had higher ERP estimates since 2010. It is also interesting to note that the 75^{th} percentile has remained fairly constant over the last 10 years, and is actually somewhat below its long-run mean.

6. The Term Structure of Equity Risk Premia

In Section 4, we described how each of the different models can trace out a term structure of the ERP – what expected excess returns are over different time horizons[11]. Figure 3 plots our preferred ERP measures as a function of investment horizon (rather than time) for some selected dates. The black line shows the average of the term structure across all periods. It is slightly upward sloping, with a short-term ERP at just over 6% and a three-year ERP at almost 7%. We selected the other dates because they are typical dates for when the ERP was unusually high or unusually low at the one-month horizon. We see that the ERP is strongly mean-reverting, with the term structure sloping downward for high one-month ERP periods, and sloping upward for low one-month ERP periods. In contrast, the ERP in July 2013 is upward sloping, something that has never happened in periods of elevated ERP.

7. Why is the Equity Risk Premium High?

The last two sections showed evidence that the ERP is high at all horizons, and that this is an unusual occurrence given the current economic and financial environment. There are two reasons why the ERP can be high: low discount rates and high current or expected future cash flows.

[11] For other ways to estimate the term structure of the ERP using equilibrium models or derivatives, see Ait-Sahalia, Karaman and Mancini (2012), Ang and Ulrich (2012), Berg (2013), Boguth, Carlson, Fisher and Simutin (2012), Croce, Lettau and Ludvigson (2012), Lemke and Werner (2009), Lettau and Wachter (2011), Muir (2013), among others.

Figure 4 shows that earnings are likely not the reason why the ERP is high. The blue line shows the realized monthly growth rates of real earnings for the S&P500 expressed in annualized percentage points. Since 2010, earnings growth has been declining, hovering around zero for the last few months of the sample. It currently stands at 2.5%, which is near its long-run average. Perhaps more importantly for the equity premium, the expectations of future earnings growth since 2010 have also been moderate to low. The red line in Figure 4 shows the year-on-year change in the mean expectation of one-year-ahead earnings per share for the S&P500. Similarly to realized earnings growth, earnings per share have been declining over the last three years, making expected earnings growth an unlikely reason for why the ERP is near its all-time high.

Nominal and real bond yields, on the other hand, have been exceptionally low since the end of the financial crisis. Figure 5 displays the term structure of the ERP under two counterfactual scenarios, in addition to the mean and current term structures already displayed in Figure 4. In the first counterfactual scenario, we leave expected stock returns unmodified but change the risk-free rates from the current values of nominal bond yields to the average nominal bond yields over 1960-2013. In other words, we replace R^f_{t+k} in equation (2) by the mean of R^f_{t+k} over t. The result of this counterfactual is shown in Figure 5 in orange. Using average levels of bond yields brings the whole term structure of the ERP much closer to its mean level (the black line), especially at short horizons. This shows that a "normalization" of bond yields, everything else being equal, would bring the ERP down substantially. In our second counterfactual exercise, we do not keep expected stock returns unchanged, but instead estimate the following regression:

$$\Delta ERP_t(k) = a(k) + b(k) \times \Delta y^1_t + c(k) \times \Delta y^5_t + d(k) \times \Delta y^{10}_t + error_t(k) \qquad (14)$$

where y^1_t, y^5_t and y^{10}_t are nominal yields for one, five and ten-year constant maturity bonds, and Δ is the first-difference operator, i.e. $\Delta x_t = x_t - x_{t-1}$ for any variable x. Equation (14) can be thought of as regressing $\Delta ERP_t(k)$ on basic level, slope and curvature factors of the nominal yield curve since these factors are linear combinations of the three bond yields y^1_t, y^5_t and y^{10}_t. We chose to run regression (14) in differences to avoid spurious regression bias, since bond yields and the ERP are persistent variables[12]. We then add to the current term structure of the ERP the fitted values of regression (14) that result from plugging in the values of $\Delta y^1_t, \Delta y^5_t$ and Δy^{10}_t that would bring bond yields from their current levels to their historical levels:

$$\Delta y^j_t = \bar{y}^j - y^j_{July\ 2013}$$

for $j = 1, 5$ and 10 years, where $\bar{y}^j$ is the time-average of y^j_t. The resulting counterfactual term structure of the ERP is shown in green in Figure 5. Unlike the case in which expected returns were held constant, this counterfactual assumes that expected returns respond to changes in yields in the same way that they have responded in the past. The resulting counterfactual term structure of the ERP is now flat and substantially below its average value. This means that if yields increased to their average levels and expected returns reacted to this increase as they have in the past, the ERP would decrease below its average levels at all horizons. This exercise shows that the current environment of exceptionally low bond yields is capable, quantitatively speaking, of causing an ERP as high as we are currently observing.

[12] An augmented Dickey-Fuller test fails to reject a unit root in the EPR at the 5% level using any number of lags between 5 and 18 (the maximum lag chosen by the Schwert criterion). The tau test statistic using 15 lags (the optimal number of lags obtained by the Ng-Perron procedure) is -1.96, which is smaller than the critical value of -2.83. A similar analysis for bond yields also fails to reject the null hypothesis of a unit root.

8. Are excess returns predictable?

In this section, we analyze how ERP models perform when trying to predict the realized equity premium. There is substantial debate in the academic literature[13] on whether any model can explain or predict the ERP better than the historical mean of realized excess returns. For this reason, we choose performance metrics that are relative to the historical mean. The historical mean itself is only a weak explanatory and predictive variable. Its correlation with the realized ERP is about 6 percent, and the R^2 in one to sixty month predictive regressions is less than 1 percent.

The measure we use for how well a model predicts the ERP is the out-of-sample R^2, popularized by Campbell and Thompson (2008):

$$R_{OOS}^2 = 100 \times \left(1 - \frac{\sum_{t=1}^{T}(R_t^e - Model\ ERP_t)^2}{\sum_{t=1}^{T}\left(R_t^e - Historical\ mean_t\right)^2}\right) \tag{15}$$

Here R_t^e are the realized excess returns at time t, $Model\ ERP_t$ is the time t real-time estimate of the ERP given by some model and $Historical\ mean_t$ is the real-time mean of the ERP since the beginning of the sample. Because $Model\ ERP_t$ is computed using information available at $t-1$ but R_t^e is only realized at time t, we interpret R_{OOS}^2 as a measure of how well the model predicts the ERP compared to the historical mean.

The out-of sample R^2 in equation (15) ranges from minus infinity to +100 percent. If the R^2 is 0, the historical mean and the model in question perform equally well, whereas a positive R^2 implies that the model outperforms the historical mean. Note that this measure is not a traditional R^2, so its units cannot be interpreted as the percentage of the variance of the dependent variable explained by the model. However, the R^2 numbers have an intuitive economic interpretation. A mean-variance investor with coefficient of relative risk aversion equal to γ, when using the $Model\ ERP_t$ as the measure of expected excess returns, will earn returns over the whole sample in excess of those predicted by the historical mean equal to R_{OOS}^2/γ.[14] For example, if $\gamma = 2$, and $R_{OOS}^2 = 10$ percent, then the investor can, though better predictability, earn extra returns of 5 percent over the T periods considered. Of course, if R_{OOS}^2 is negative, using the model would lead to returns that are lower than those obtained by using the historical mean as the sole estimate for the ERP.

Figures 6 and 7 show the results. In Figure 6, we show the R_{OOS}^2 for the models that perform the best within each category. For DDM, the Shiller model (model 4) outperformed all other DDM at all horizons. For cross-sectional regressions, the model by Adrian, Crump and Moench (model 14) did best at the one month to two year horizons, but was outperformed by Fama-French and momentum (model 12) at the three and four year horizons, and by the cross-sectional regression with inflation (model 13) at the five year horizon. For time-series regressions, the results were mixed: The Goyal and Welch predictors (model 16) were the best at the one month and three year horizons; The predictors in Fama and French (model 18) were the best at the four and five year horizons; The dividend-price ratio (model 15) was the best at the other horizons. For surveys, all results correspond to CFO surveys (model 20), which is the only survey we analyze.

The main conclusion is that the R-squares are small, which means that none of the models drastically outperform the historical mean out of sample. For example, cross-sectional regressions (the green line with crosses) starts at almost 12%, which means that the mean-variance investor with coefficient of

[13] See, for example, Goyal and Welch (2008), Lettau and Van Nieuwerburgh (2008), Campbell and Thompson (2008), Cochrane (2011), Fama and French (2002) and references therein.
[14] For a derivation of this fact, see Campbell and Thompson (2008).

relative risk aversion equal to one would have made an extra 12% over the last 53 years compared to just using the historical mean as a measure of the ERP. This amounts to about 23 basis points per year. In addition, although surveys are clearly inferior predictors, all other predictors are comparable at all horizons.

Figure 7 displays the same analysis but using principal components instead of the individual models for each category. The principal components were computed in real time following the procedure explained in Section 4. The predictability of the principal component need not be better or worse than the best model in each category. One the one hand, the principal component may reduce noise and aggregate useful information from the many models. On the other hand, it puts some weight on models that have worse predictability than the best models. Figure 7 shows that, as a group, dividend discount models perform substantially better than other models at short horizons but are worst at long horizons, while cross-sectional regressions perform best at long horizons but are worst at short horizons. The principal component of all models, in the solid purple line, performs well across all horizons and is always close to the model with the best predictability. The good performance of the principal component reinforces its usefulness as a summary statistic.

9. Conclusion

Estimates for the ERP as high as we have found should give policymakers pause. We have argued that it is unusual for the ERP to be at its present level in the current stage of the business cycle, especially when expectations are that it will continue to rise over the next three years. Because the ERP is a key input in many important decisions of economic agents, an unusually high ERP can herald unusual behavior. Our analysis provides evidence that is consistent with a bond-driven ERP: expected excess stock returns are high not because stocks are expected to have high returns, but because bond yields are exceptionally low. In such an environment, we should expect monetary policy –both conventional and unconventional—to have a large impact on asset prices and hence the real economy.

Our study of the ERP has many limitations. The main one is that stocks returns are very difficult to predict, if they are predictable at all. We have shown how to improve upon current estimates by using principal components yet still found weak evidence in favor of predictability, at least at horizons shorter than five years. Any conclusions that rely on ERP estimates must be weighted by how strongly it predicts future returns. Another limitation is that even though we have conducted all of our out-of-sample tests in real time (using information available at time t for $t + 1$ estimates), some of the models we use had not been yet proposed for many periods of our sample, so there is some selection and forward looking biases. Finally, we have not focused on the possibility that a bubble –rational or irrational— could be a further driver of the recent high realized and expected returns, a topic we consider outside of the scope of the broadly used models we consider.

References

Adrian, T., Crump, R., Moench E., 2012. Efficient regression-based estimation of dynamic asset pricing models. Federal Reserve Bank of New York Staff Reports, Number 439, May 2011.

Ait-Sahalia, Y., Karaman, M., Mancini, L., 2012. The term structure of variance swaps, risk premia and the expectation hypothesis. Risk Premia and the Expectation Hypothesis (August 24, 2012).

Ang, A., Kristensen, D., 2012. Testing conditional factor models. NBER Working Paper 17561.

Ang, A., Ulrich, M., 2012. Nominal Bonds, Real Bonds, and Equity.

Asness, C., 2003. Fight the Fed model. Journal of Portfolio Management (Fall).11-24.

Berg, T., 2013. The Term Structure of Risk Premia: Evidence from CDS Spreads.

Bloom, N., 2009. The impact of uncertainty shocks. Econometrica, . 77 (3), 623-685.

Boguth, O., Carlson, M., Fisher, A. J., Simutin, M., 2012. Leverage and the limits of arbitrage pricing: Implications for dividend strips and the term structure of equity risk premia. manuscript, Arizona State University.

Campbell, J.Y., Thompson, S., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21 (4), 1509-1531.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57–82.

Cochrane, J. H., 2001. Asset pricing. Princeton University Press, New Jersey.

Cochrane, J.H., 2011. Presidential address: discount rates. Journal of Finance 66 (4), 1047-1108.

Croce, M., Lettau, M., Ludvigson, S., 2012. Investor information, long-run risk, and the term structure of equity. Unpublished paper, New York University.[40, 49].

Duarte, F., 2013. Inflation and the cross-section of stock returns. Federal Reserve Bank of New York Staff Reports, Number 621.

Duarte, F., Kogan, L., Livdan, D., 2013. Aggregate Investment and Stock Returns. Unpublished paper, M.I.T.

Damodaran, A., 2012. Equity Risk Premiums (ERP): Determinants, Estimation and Implications – The 2012 Edition. Working paper New York University.

Easton, P.D., Sommers, G.A., 2007. Effect of analysts' optimism on estimates of the expected return implied by earnings forecasts. Journal of Accounting Research 45, 983-1016.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. Journal of Finance 47 (2), 427-465.

Fama, E.F., French, K.R., 2002. The equity premium, Journal of Finance, 57 (2), 637-659

Fuller, R., Hsia, C., 1984. A Simplified Common Stock Valuation Model, Financial Analysts Journal, September-October, 1984.

Gordon, M. J., 1962. The investment, financing, and valuation of the corporation. Greenwood Press.

Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium predictions. Review of Financial Studies 21 (4), 1455-1508.

Graham, J., Harvey, C., 2012. The Equity Risk Premium in 2012. Working paper, Duke University.

Greenwood, R., Shleifer, A., 2012. Expectations of returns and expected returns. Working paper, Harvard University.

Gürkaynak, R. S., Sack, B., & Wright, J. H., 2007. The US Treasury yield curve: 1961 to the present. Journal of Monetary Economics, 54(8), 2291-2304.

Hall, R. E., 2013. The Routes into and out of the Zero Lower Bound. Working paper, Stanford University.

Kuehn, L. A., Petrosky-Nadeau, N., & Zhang, L., 2012. An equilibrium asset pricing model with labor market search (No. w17742). National Bureau of Economic Research

Lettau, M., Van Nieuwerburgh, S., 2008. Reconciling the return predictability evidence. The Review of Financial Studies 21 (4), 1607-1652.

Lettau, M., Wachter, J. A., 2011. The term structures of equity and interest rates. Journal of Financial Economics, 101(1), 90-113.

Ludvigson, S.C., Lettau, M., 2001. Consumption, aggregate wealth, and expected stock returns. Journal of Finance 56 (3), 815-849.

Merton, R.C., 1973. An intertemporal capital asset pricing model. Econometrica 41 (5), 867-887.

Muir, T., 2013. Financial Crises, Risk Premia, and the Term Structure of Risky Assets.

Panigirtzoglou, N., Loeys, J., 2005. A fair value model for US bonds, credit and equities. JP Morgan, June 2005.

Polk, C., Thompson, S., Vuolteenaho, T., 2006. Cross-sectional forecasts of the equity premium. Journal of Financial Economics 81, 101-141.

Appendix A: Data Variables

Variable	Source	Original Frequency	First period	Last Period
Sentiment measure from "Investor Sentiment in the Stock Market," Journal of Economic Perspectives, 2007	Jeffrey Wurgler	monthly	07/01/65	12/01/10
Equity issuance	Jeffrey Wurgler	monthly	01/01/58	12/01/10
Debt issuance	Jeffrey Wurgler	monthly	01/01/58	12/01/10
CAY	Martin Lettau	quarterly	03/01/52	09/01/12
One-year ahead ERP from CFO survey	Duke CFO Survey	quarterly	06/06/00	06/05/13
Ten-year ahead ERP from CFO survey	Duke CFO Survey	quarterly	06/06/00	06/05/13
Book value per share	Compustat	annual	12/31/77	12/31/12
S&P 500 closing price	Compustat	monthly	02/28/62	06/30/13
ERP from a dividend discount model	Damodaran	annual	12/01/61	12/01/12
ERP from a dividend discount model using free cash flow	Damodaran	annual	12/01/61	12/01/12
Size and book-to-market sorted portfolios	Fama-French	monthly	07/01/26	06/01/13
Realized excess returns for the market	Fama-French	monthly	07/01/26	06/01/13
Size factor	Fama-French	monthly	07/01/26	06/01/13
Book-to-market factor	Fama-French	monthly	07/01/26	06/01/13
Risk free rate	Fama-French	monthly	07/01/26	06/01/13
Baa minus Aaa bond yield spread	FRED	monthly	01/01/19	07/01/13
NBER recession indicator	FRED	monthly	01/02/00	06/01/13
Momentum portfolios	Fama-French	monthly	01/01/27	12/01/12
Momentum factor	Fama-French	monthly	01/01/27	06/01/13
ERP as constructed in Adrian, Crump and Moench (2013)	NY Fed	monthly	01/01/63	07/01/13
Nominal price for the S&P 500	Shiller	monthly	01/02/00	07/01/13
Nominal dividends for the S&P 500	Shiller	monthly	01/02/00	06/01/13
Nominal earnings for the S&P 500	Shiller	monthly	01/02/00	03/01/13
Consumer Price Index	Shiller	monthly	01/02/00	07/01/13
10 year nominal treasury yield	Shiller	monthly	01/02/00	07/01/13
Real price for the S&P 500	Shiller	monthly	01/02/00	07/01/13
Real dividends for the S&P 500	Shiller	monthly	01/02/00	06/01/13
Real earnings for the S&P 500	Shiller	monthly	01/02/00	03/01/13
Cyclycally Adjusted Price-Earnings ratio	Shiller	monthly	01/02/00	07/01/13
Realized Earnings per Share for the S&P 500	Thomson Reuters I/B/E/S	annual	12/31/81	12/31/12
Mean analyst forecast for Earnings per share for the S&P500	Thomson Reuters I/B/E/S	monthly	01/14/82	04/18/13
Yield on TIPS	Fed Board	monthly	01/01/03	07/01/13
Nominal yields	Fed Board (Gurkaynak, Sack and Wright)	daily	06/14/61	08/12/13

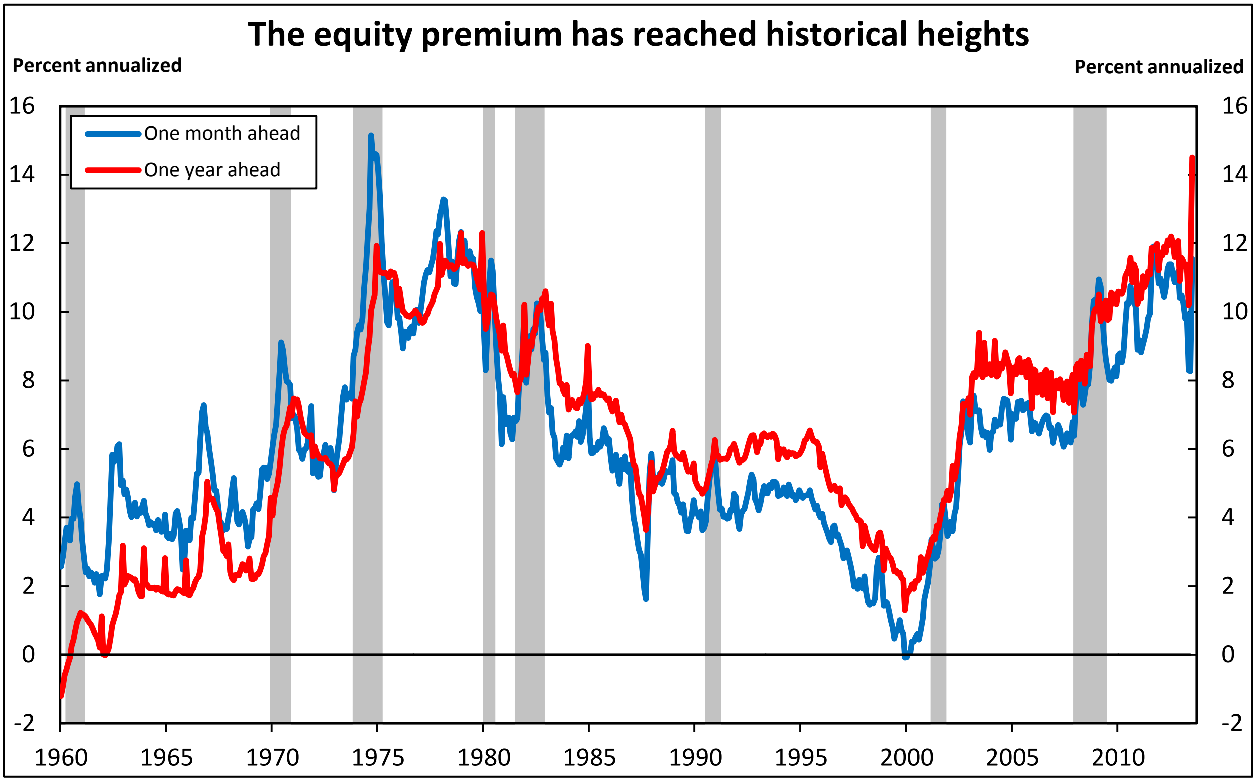


Figure 1 The equity risk premium (expected excess returns) over a one year ahead and one month ahead horizons are the first principal components of 20 models of the equity premium. The models include time-series and cross-sectional regressions, dividend discount models and surveys. Shaded bars are NBER recessions.

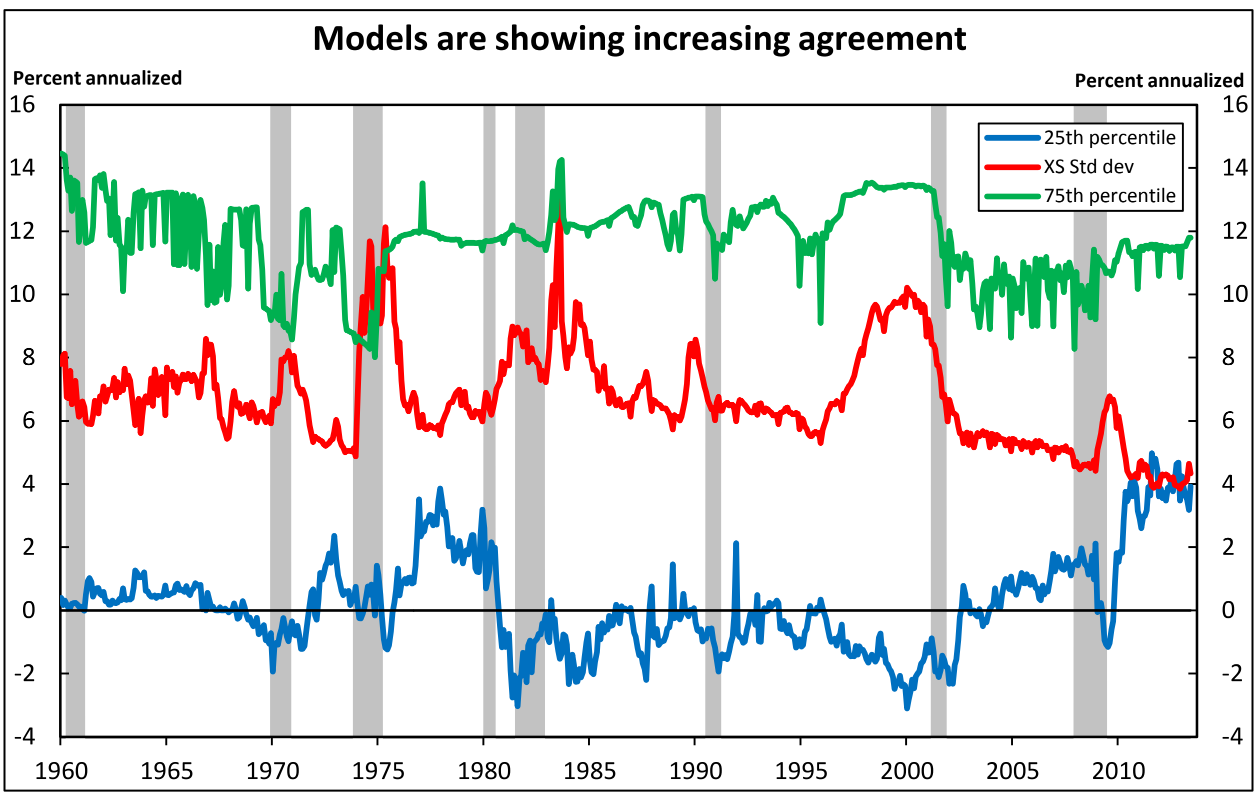


Figure 2 The cross-sectional standard deviation (labeled "XS Std dev", in red) computes, at each time period, the standard deviation of the 20 equity risk premium estimates given by the different models. The 25th and 75th percentiles (in blue and green, respectively) give the corresponding quartile of the 20 estimates for each time period. Shaded bars are NBER recessions.

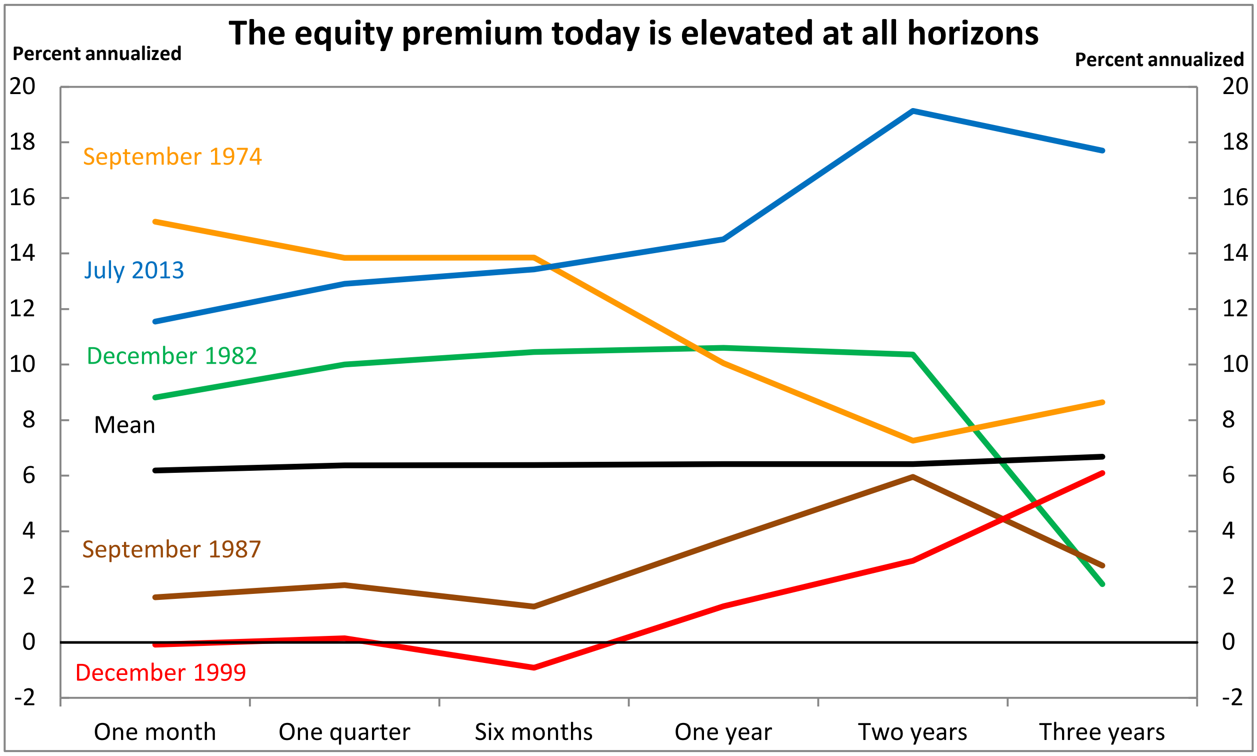


Figure 3 The equity risk premium at different horizons are the first principal component of 20 estimates of expected excess returns at different horizons. The estimates are obtained from cross-sectional and time-series regressions, dividend discount models and surveys. The black line (labeled "Mean") shows the mean of expected excess returns at different horizons over the sample 1960-2013. The most recent estimates of the term structure of the equity risk premium (labeled "July 2013" in blue), does not show mean reversion, unlike other periods when the equity risk premium was substantially above or below its mean at the one-month horizon.

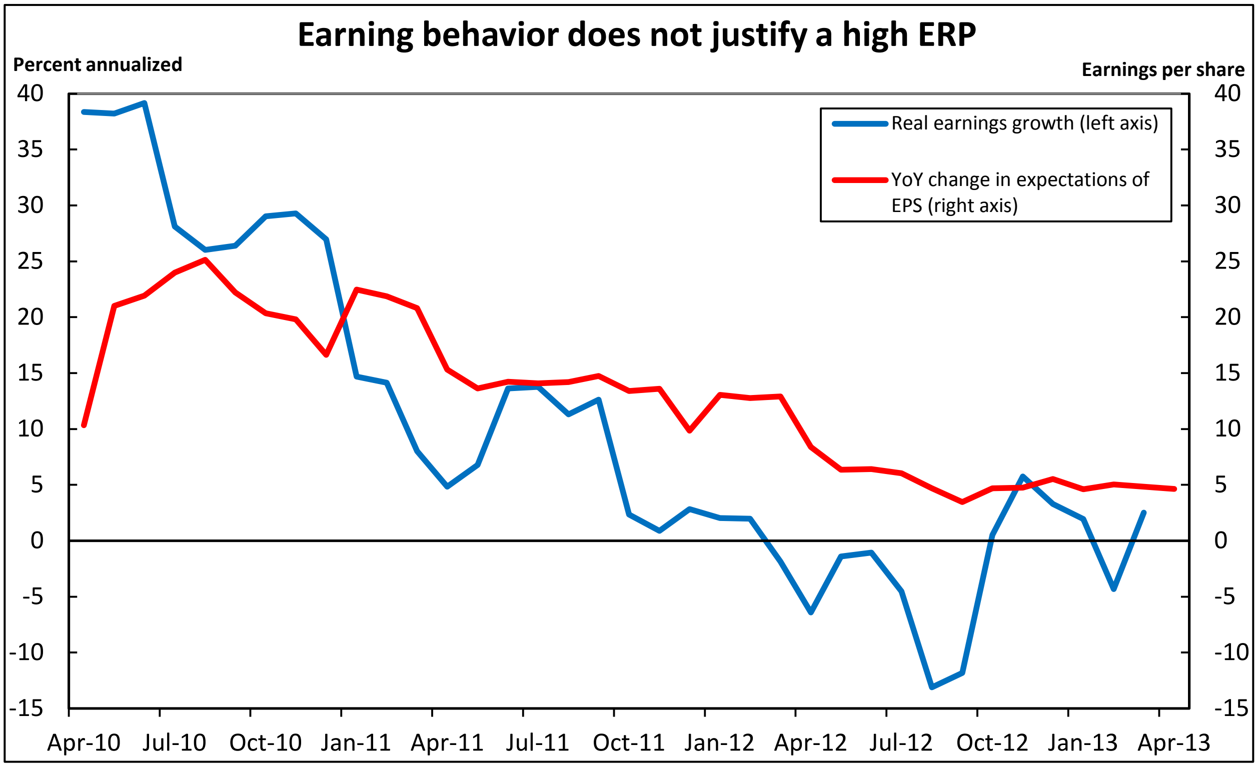


Figure 4 The blue line shows the monthly growth rate of real S&P 500 earnings, annualized and in percentage points. The red line shows the year-on-year change in the mean expectation of one-year ahead earnings per share for the S&P 500 from analysts from Thomson Reuters I/B/E/S.

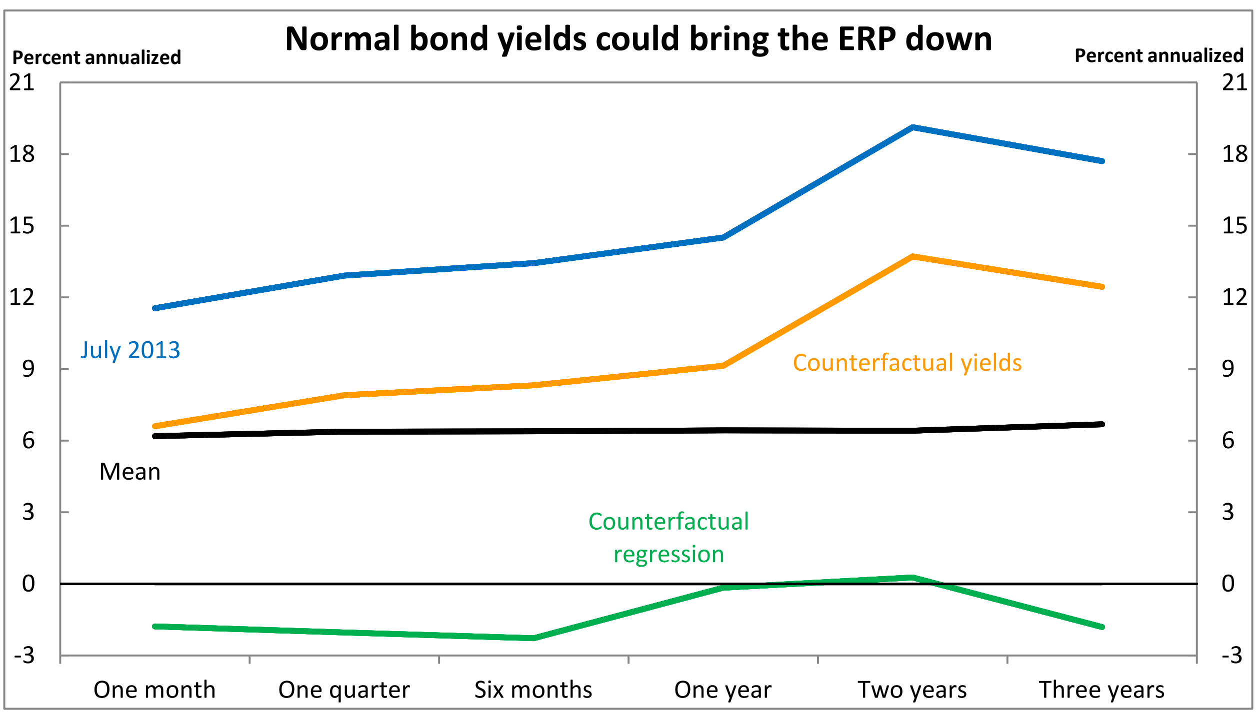


Figure 5 The black line (labeled "Mean") shows the mean term structure of the equity risk premium over the sample 1960-2013. The blue line (labeled "July 2013") shows the most recent estimates. The orange line (labeled "Counterfactual yields") shows what the term structure of the equity risk premium would be in July 2013 if instead of subtracting today's yield curve from expected returns we subtracted the average yield curve for the period 1960-2013. The green line shows an estimate of what the term structure of the equity premium would be if yields rose to their average historical levels and expected stock returns co-moved with yields with the same correlation as during 1960-2013.

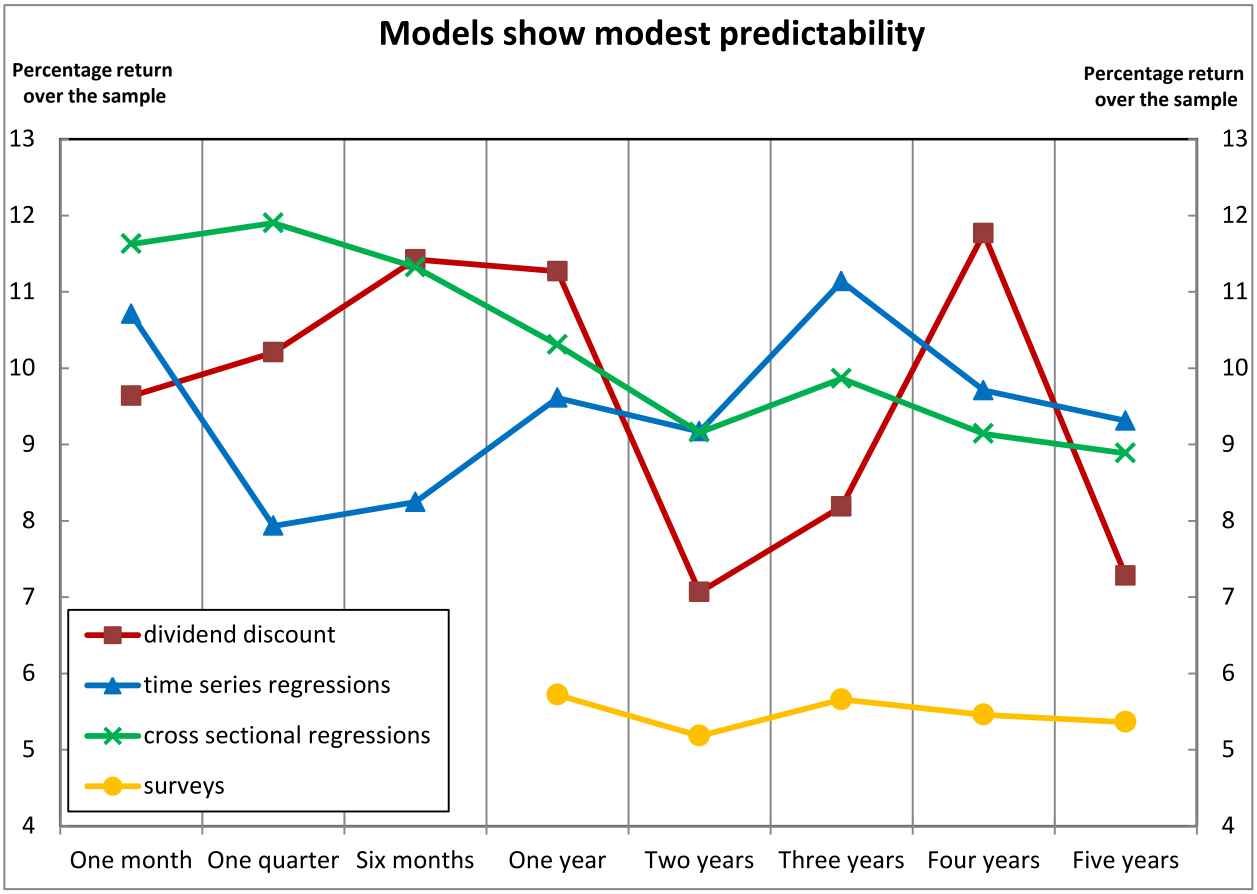


Figure 6 Each data point corresponds to the returns that a mean-variance investor with unit coefficient of relative risk aversion would have earned over the period 1960-2013 if she had used one of the equity risk premium models over and above the returns she would have made if she had assumed that expected excess returns are equal to their historical mean at each point in time. The x-axis shows the investment horizon of the investor (how often the portfolio is rebalanced and hence how far ahead excess returns must be forecast).

For each class of model (dividend discount, time-series regressions, cross-sectional regressions, surveys) we report the model that had the best predictability. For dividend discount models, the Shiller model (model 4) outperformed all other discount models at all horizons. For cross-sectional regressions, the model by Adrian, Crump and Moench (model 14) did best at the one month to two year horizon, but was outperformed by Fama-French and momentum (model 12) at the three and four year horizons, and by the cross-section with inflation (model 13) at the five year horizon. For time-series regressions, the results were mixed: The Goyal and Welch predictors (model 16) were the best at a one month and three year horizons; The predictors in Fama and French (model 18) were the best at the four and five year horizons; The dividend-price ratio (model 15) was the best at the other horizons. For surveys, all results correspond to CFO surveys (model 20), which is the only survey we analyze.

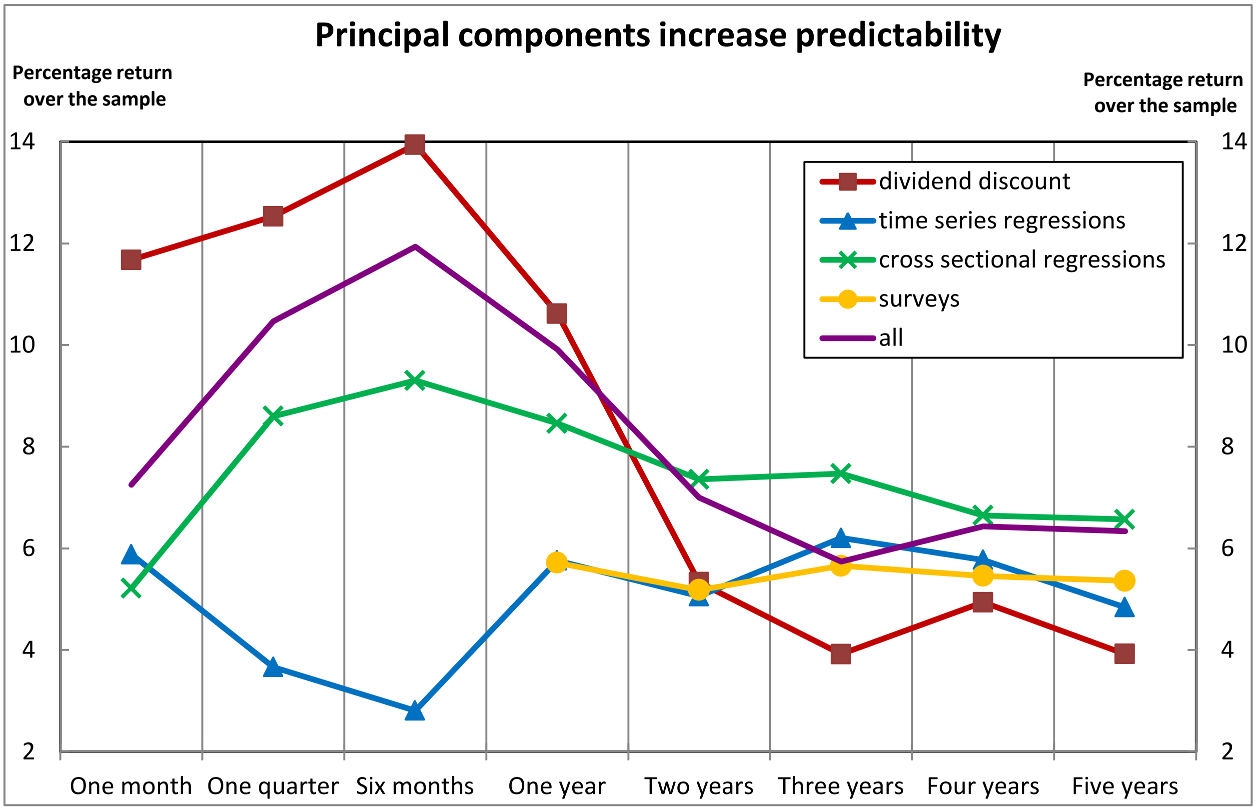


Figure 7 Each data point corresponds to the returns that a mean-variance investor with unit coefficient of relative risk aversion would have earned over the period 1960-2013 if she had used the first principal component of all models within a certain class (dividend discount, time-series regressions, cross-sectional regressions, surveys) over and above the returns she would have made if she had assumed that expected excess returns are equal to their historical mean at each point in time. The x-axis shows the investment horizon of the investor (how often the portfolio is rebalanced and hence how far ahead excess returns must be forecast). The line labeled "all" corresponds to the principal component of all models (our preferred measure).

Federal Reserve Bank of New York
Staff Reports

The Equity Risk Premium: A Review of Models

Fernando Duarte
Carlo Rosa

Staff Report No. 714
February 2015

The Equity Risk Premium: A Review of Models

Fernando Duarte and Carlo Rosa

Federal Reserve Bank of New York Staff Reports, no. 714
February 2015
JEL classification: C58, G00, G12, G17

Abstract

We estimate the equity risk premium (ERP) by combining information from twenty models. The ERP in 2012 and 2013 reached heightened levels—of around 12 percent—not seen since the 1970s. We conclude that the high ERP was caused by unusually low Treasury yields.

Key words: equity premium, stock returns

Duarte, Rosa: Federal Reserve Bank of New York (e-mail: fernando.duarte@ny.frb.org, carlo.rosa@ny.frb.org). The authors thank Tobias Adrian and James Egelhof for helpful comments on earlier drafts. This article is an update of, and a more comprehensive and rigorous treatment than, our blog post in Liberty Street Economics (http://libertystreeteconomics.newyorkfed.org/2013/05/are-stocks-cheap-a-review-of-the-evidence.html).

1. Introduction

The equity risk premium —the expected return on stocks in excess of the risk-free rate— is a fundamental quantity in all of asset pricing, both for theoretical and practical reasons. It is a key measure of aggregate risk-aversion and an important determinant of the cost of capital for corporations, savings decisions of individuals and budgeting plans for governments. Recently, the equity risk premium (ERP) has also returned to the forefront as a leading indicator of the evolution of the economy, a potential explanation for jobless recoveries and a gauge of financial stability[3].

In this article, we estimate the ERP by combining information from twenty prominent models used by practitioners and featured in the academic literature. Our main finding is that the ERP has reached heightened levels. The first principal component of all models –a linear combination that explains as much of the variance of the underlying data as possible– places the one-year-ahead ERP in June 2012 at 12.2 percent, above the 10.5 percent that was reached during the financial crisis in 2009 and at levels similar to those in the mid and late 1970s. Since June 2012 and until the end of our sample in June 2013, the ERP has remained little changed, despite substantial positive realized returns. It is worth keeping in mind, however, that there is considerable uncertainty around these estimates. In fact, the issue of whether stock returns are predictable is still an active area of research.[4] Nevertheless, we find that the dispersion in estimates across models, while quite large, has been shrinking, potentially signaling increased agreement

[3] As an indicator of future activity, a high ERP at short horizons tends to be followed by higher GDP growth, higher inflation and lower unemployment. See, for example, Piazzesi and Schneider (2007), Stock and Watson (2003), and Damodaran (2012). Bloom (2009) and Duarte, Kogan and Livdan (2013) study connections between the ERP and real aggregate investment. As a potential explanation of the jobless recovery, Hall (2014) and Kuehn, Petrosky-Nadeau and Zhang (2012) propose that increased risk-aversion has prevented firms from hiring as much as would be expected in the post-crisis macroeconomic environment. Among many others, Adrian, Covitz and Liang (2013) analyze the role of equity and other asset prices in monitoring financial stability.

[4] A few important references among a vast literature are Ang and Bekaert (2007), Goyal and Welch (2008), Campbell and Thompson (2008), Kelly and Pruitt (2013), Chen, Da and Zhao (2013), Neely, Rapach, Tu and Zhou (2014).

even when the models are substantially different from each other and use more than one hundred different economic variables.

In addition to estimating the level of the ERP, we investigate the reasons behind its recent behavior. Because the ERP is the difference between expected stock returns and the risk-free rate, a high estimate can be due to expected stock returns being high or risk-free rates being low. We conclude the ERP is high because Treasury yields are unusually low. Current and expected future dividend and earnings growth play a smaller role. In fact, expected stock returns are close to their long-run mean. One implication of a bond-yield-driven ERP is that traditional indicators of the ERP like the price-dividend or price-earnings ratios, which do not use data from the term structure of risk-free rates, may not be as good a guide to future excess returns as they have been in the past.

As a second contribution, we present a concise and coherent taxonomy of ERP models. We categorize the twenty models into five groups: predictors that use historical mean returns only, dividend-discount models, cross-sectional regressions, time-series regressions and surveys. We explain the methodological and practical differences among these classes of models, including the assumptions and data sources that each require.

2. The Equity Risk Premium: Definition

Conceptually, the ERP is the compensation investors require to make them indifferent at the margin between holding the risky market portfolio and a risk-free bond. Because this compensation depends on the future performance of stocks, the ERP incorporates expectations of future stock market returns, which are not directly observable. At the end of the day, any model of the ERP is a model of investor expectations. One challenge in estimating the ERP is that it is not clear what truly constitutes the market return and the risk-free rate in the real world. In practice, the most common measures of total market returns are based on broad stock market indices, such as the S&P 500 or the Dow Jones Industrial

Average, but those indices do not include the whole universe of traded stocks and miss several other components of wealth such as housing, private equity and non-tradable human capital. Even if we restricted ourselves to all traded stocks, we still have several choices to make, such as whether to use value or equal-weighted indices, and whether to exclude penny or infrequently traded stocks. A similar problem arises with the risk-free rate. While we almost always use Treasury yields as measures of risk-free rates, they are not completely riskless since nominal Treasuries are exposed to inflation[5] and liquidity risks even if we were to assume there is no prospect of outright default. In this paper, we want to focus on how expectations are estimated in different models, and not on measurement issues regarding market returns and the risk-free rate. Thus, we follow common practice and always use the S&P 500 as a measure of stock market prices and either nominal or real Treasury yields as risk-free rates so that our models are comparable with each other and with most of the literature.

While implementing the concept of the ERP in practice has its challenges, we can precisely define the ERP mathematically. First, we decompose stock returns[6] into an expected component and a random component:

$$R_{t+k} = E_t[R_{t+k}] + error_{t+k}. \tag{1}$$

In equation (1), R_{t+k} are *realized* returns between *t* and *t*+*k*, and $E_t[R_{t+k}]$ are the returns that were expected from *t* to $t + k$ using information available at time t. The variable $error_{t+k}$ is a random variable that is unknown at time t and realized at $t + k$. Under rational expectations, $error_{t+k}$ has a mean of zero and is orthogonal to $E_t[R_{t+k}]$. We keep the discussion as general as possible and do not assume rational

[5] Note that inflation risk in an otherwise risk-free nominal asset does not invalidate its usefulness to compute the ERP. If stock returns and the risk-free rate are expressed in nominal terms, their difference has little or no inflation risk. This follows from the following formula, which holds exactly in continuous time and to a first order approximation in discrete time: real stock returns – real risk-free rate = (nominal stock returns – expected inflation) – (nominal risk-free rate – expected inflation) = nominal stock returns– nominal risk-free rate. Hence, there is no distinction between a nominal and a real ERP.
[6] Throughout this article, all returns are *net* returns. For example, a five percent return corresponds to a net return of 0.05 as opposed to a *gross* return of 1.05.

expectations at this stage, although it will be a feature of many of the models we consider. The ERP at time t for horizon k is defined as

$$ERP_t(k) = E_t[R_{t+k}] - R^f_{t+k}, \tag{2}$$

where R^f_{t+k} is the risk-free rate for investing from t to $t+k$ (which, being risk-free, is known at time t).

This definition shows three important aspects of the ERP. First, future expected returns and the future ERP are stochastic, since expectations depend on the arrival of new information that has a random component not known in advance[7]. Second, the ERP has an investment horizon k embedded in it, since we can consider expected excess returns over, say, one month, one year or five years from today. If we fix t, and let k vary, we trace the *term structure* of the equity risk premium. Third, if expectations are rational, because the unexpected component $error_{t+k}$ is stochastic and orthogonal to expected returns, the ERP is always less volatile than realized excess returns. In this case, we expect ERP estimates to be smoother than realized excess returns.

3. Models of the Equity Risk Premium

We describe twenty models of the equity risk premium, comparing their advantages, disadvantages and ease of implementation. Of course, there are many more models of the ERP than the ones we consider. We selected the models in our study based on the recent academic literature, their widespread use by practitioners and data availability. Table I describes the data we use and their sources, all of which are either readily available or standard in the literature[8]. With a few exceptions, all data is monthly from January 1960 to June 2013. Appendix A provides more details.

[Insert Table I here]

[7] More precisely, $E_t[R_{t+k}]$ and $ERP_t(k)$ are known at time t but random from the perspective of all earlier periods.
[8] In fact, except for data from I/B/E/S and Compustat, all sources are public.

We classify the twenty models into five categories based on their underlying assumptions; models in the same category tend to give similar estimates for the ERP. The five categories are: models based on the historical mean of realized returns, dividend discount models, cross-sectional regressions, time-series regressions and surveys.

All but one of the estimates of the ERP are constructed in real time, so that an investor who lived through the sample would have been able to construct the measures at each point in time using available information only[9]. This helps minimize look-ahead bias and makes any out-of-sample evaluation of the models more meaningful. Clearly, most of the models themselves were designed only recently and were not available to investors in real time, potentially introducing another source of forward-looking and selection biases that are much more difficult to quantify and eliminate.

3.1 Historical mean of realized returns

The easiest approach to estimating the ERP is to use the historical mean of realized market returns in excess of the contemporaneous risk-free rate. This model is very simple and, as shown in Goyal and Welch (2008), quite difficult to improve upon when considering out-of-sample predictability performance measures. The main drawbacks are that it is purely backward looking and assumes that the future will behave like the past, i.e. it assumes the mean of excess returns is either constant or very slow moving over time, giving very little time-variation in the ERP. The main choice is how far back into the past we should go when computing the historical mean. Table II shows the two versions of historical mean models that we use.

[Insert Table II here]

[9] The one exception is Adrian, Crump and Moench's (2014) cross-sectional model, which is constructed using full-sample regression estimates.

3.2 Dividend discount models (DDM)

All DDM start with the basic intuition that the value of a stock is determined by no more and no less than the cash flows it produces for its shareholders, as in Gordon (1962). Today's stock price should then be the sum of all expected future cash flows, discounted at an appropriate rate to take into account their riskiness and the time value of money. The formula that reflects this intuition is

$$P_t = \frac{D_t}{\rho_t} + \frac{E_t[D_{t+1}]}{\rho_{t+1}} + \frac{E_t[D_{t+2}]}{\rho_{t+2}} + \frac{E_t[D_{t+3}]}{\rho_{t+3}} + \cdots, \tag{3}$$

where P_t is the current price of the stock, D_t are current cash flows, $E_t[D_{t+k}]$ are the cash flows k periods from now expected as of time t, and ρ_{t+k} is the discount rate for time $t+k$ from the perspective of time t. Cash flows to stockholders certainly include dividends, but can also arise from spin-offs, buy-outs, mergers, buy-backs, etc. In general, the literature focuses on dividend distributions because they are readily available data-wise and account for the vast majority of cash flows. The discount rate can be decomposed into

$$\rho_{t+k} = 1 + R^f_{t+k} + ERP_t(k). \tag{4}$$

In this framework, the risk-free rate captures the discounting associated with the time value of money and the ERP captures the discounting associated with the riskiness of dividends. When using a DDM, we refer to $ERP_t(k)$ as the *implied* ERP. The reason is that we plug in prices, risk-free rates and estimated expected future dividends into equation (3), and then derive what value of $ERP_t(k)$ makes the right-hand side equal to the left-hand side in the equation, i.e. what ERP value is *implied* by equation (3).

DDM are forward looking and are consistent with no arbitrage. In fact, equation (3) must hold in any economy with no arbitrage[10]. Another advantage of DDM is that they are easy to implement. A drawback of DDM is that the results are sensitive to how we compute expectations of future dividends. Table III displays the DDM we consider and a brief description of their different assumptions.

[Insert Table III here]

3.3 Cross-sectional regressions

This method exploits the variation in returns and exposures to the S&P 500 of different assets to infer the ERP[11]. Intuitively, cross-sectional regressions find the ERP by answering the following question: what is the level of the ERP that makes expected returns on a variety of stocks consistent with their exposure to the S&P 500? Because we need to explain the relationship between returns and exposures for multiple stocks with a single value for the ERP (and perhaps a small number of other variables), this model imposes tight restrictions on estimates of the ERP.

The first step is to find the exposures of assets to the S&P 500 by estimating an equation of the following form:

$$R_{t+k}^{i} - R_{t+k}^{f} = \alpha^{i} \times state\ variables_{t+k} + \beta^{i} \times risk\ factors_{t+k} + idiosyncratic\ risk_{t+k}^{i}. \qquad (5)$$

In equation (5), R_{t+k}^{i} is the realized return on a stock or portfolio i from time t to $t+k$. $State\ variables_{t+k}$ are any economic indicators that help identify the state of the economy and its likely future path. $Risk\ factors_{t+k}$ are any measures of systematic contemporaneous co-variation in returns across all stocks or portfolios. Of course, some economic indicators can be both state variables and risk

[10] Note that when performing the infinite summation in equation (3) we have not assumed the n^{th} term goes to zero as n tends to infinity, which allows for rational bubbles. In this sense, DDM do allow for a specific kind of bubble.

[11] See Polk, Thompson and Vuolteenaho (2006) and Adrian, Crump and Moench (2014) for a detailed description of this method.

factors at the same time. Finally, $idiosyncratic\ risk^i_{t+k}$ is the component of returns that is particular to each individual stock or portfolio that is not explained by $state\ variables_{t+k}$ or $risk\ factors_{t+k}$ (both of which, importantly, are common to all stocks and hence not indexed by i). Examples of state variables are inflation, unemployment, the yield spread between Aaa and Baa bonds, the yield spread between short and long term Treasuries, and the S&P 500's dividend-to-price ratio. The most important risk factor is the excess return on the S&P 500, which we must include if we want to infer the ERP consistent with the cross-section of stock returns. Other risk-factors usually used are the Fama-French (1992) factors and the momentum factor of Carhart (1997). The values in the vector α^i give the strength of asset-specific return predictability and the values in the vector β^i give the asset-specific exposures to risk factors[12]. For the cross-section of assets indexed by i, we can use the whole universe of traded stocks, a subset of them, or portfolios of stocks grouped, for example, by industry, size, book-to-market, or recent performance. It is important to point out that equation (5) is not a predictive regression; the left and right-hand side variables are both associated with time $t + k$.

The second step is to find the ERP associated with the S&P 500 by estimating the cross-sectional equations

$$R^i_{t+k} - R^f_{t+k} = \lambda_t(k) \times \hat{\beta}^i \ , \tag{6}$$

where $\hat{\beta}^i$ are the values found when estimating equation (5). Equation (6) attempts to find, at each point in time, the vector of numbers $\lambda_t(k)$ that makes exposures β^i as consistent as possible with realized excess returns of all stocks or portfolios considered. The element in the vector $\hat{\lambda}_t(k)$ that is multiplied by

[12] The vectors α^i and β^i could also be time-varying, reflecting a more dynamic relation between returns and their explanatory variables. In this case, the estimation of equation (5) is more complicated and requires making further assumptions. The model by Adrian, Crump and Moench (2014) is the only cross-sectional model we examine that uses time-varying α^i and β^i .

the element in the $\hat{\beta}^i$ vector corresponding to the S&P 500 is $ERP_t(k)$, the equity risk premium we are seeking.

One advantage of cross-sectional regressions is that they use information from more asset prices than other models. Cross-sectional regressions also have sound theoretical foundations, since they provide one way to implement Merton's (1973) Intertemporal Capital Asset Pricing Model. Finally, this method nests many of the other models considered. The two main drawbacks of this method are that results are dependent on what portfolios, state variables and risk factors are used (Harvey, Liu and Zhu (2014)), and that it is not as easy to implement as most of the other options. Table IV displays the cross-sectional models in our study, together with the state variables and risk factors they use.

[Insert Table IV here]

3.4 Time-series regressions

Time-series regressions use the relationship between economic variables and stock returns to estimate the ERP. The idea is to run a predictive linear regression of realized excess returns on lagged "fundamentals":

$$R_{t+k} - R_{t+k}^{f} = a + b \times Fundamental_t + error_t. \qquad (7)$$

Once estimates $\hat{a}$ and $\hat{b}$ for a and b are obtained, the ERP is obtained by ignoring the error term:

$$ERP_t(k) = \hat{a} + \hat{b} \times Fundamental_t. \qquad (8)$$

In other words, we estimate only the forecastable or expected component of excess returns. This method attempts to implement equations (1) and (2) as directly as possible in equations (7) and (8), with the assumption that "fundamentals" are the right sources of information to look at when computing expected returns, and that a linear equation is the correct functional specification.

The use of time-series regressions requires minimal assumptions; there is no concept of equilibrium and no absence of arbitrage necessary for the method to be valid[13]. In addition, implementation is quite simple, since it only involves running ordinary least-square regressions. The challenge is to select what variables to include on the right-hand side of equation (7), since results can change substantially depending on what variables are used to take the role of "fundamentals". In addition, including more than one predictor gives poor out-of-sample predictions even if economic theory may suggest a role for many variables to be used simultaneously (Goyal and Welch (2008)). Finally, time-series regressions ignore information in the cross-section of stock returns. Table V shows the time-series regression models that we study.

[Insert Table V here]

3.5 Surveys

The survey approach consists of asking economic agents about the current level of the ERP. Surveys incorporate the views of many people, some of which are very sophisticated and/or make real investment decisions based on the level of the ERP. Surveys should also be good predictors of excess returns because in principle stock prices are determined by supply and demand of investors such as the ones taking the surveys. On the other hand, Greenwood and Shleifer (2014) document that investor expectations of future stock market returns are positively correlated with past stock returns and with the current level of the stock market, but strongly *negatively* correlated with model-based expected returns and future realized stock market returns. Other studies such as Easton and Sommers (2007) also argue that survey measures of the ERP can be systematically biased. In this paper, we use the survey of CFOs by Graham and Harvey (2012), which to our knowledge is the only large-scale ERP survey that has more than just a few years of data (see Table VI).

[Insert Table VI here]

[13] However, the Arbitrage Pricing Theory of Ross (1976) provides a strong theoretical underpinning for time-series regressions by using no-arbitrage conditions.

4. Estimation of the Equity Risk Premium

We now study the behavior of the twenty models we consider by conducting principal component analysis. Since forecast accuracy can be substantially improved through the combination of multiple forecasts[14], the optimal strategy to forecast excess stock returns may consist of combining together all these models. The first principal component of the twenty models that we use is the linear combination of ERP estimates that captures as much of the variation in the data as possible. The second, third, and successive principal components are the linear combinations of the twenty models that explain as much of the variation of the data as possible and are also uncorrelated to all the preceding principal components. If the first few principal components —say one or two— account for most of the variation of the data, then we can use them as a good summary for the variation in all the measures over time, reducing the dimensionality from twenty to one or two. In addition, in the presence of classical measurement error, the first few principal components can achieve a higher signal-to-noise ratio than other summary measures like the cross-sectional mean of all models (Geiger and Kubin (2013)).

To compute the first principal component, we proceed in three steps. We first de-mean all ERP estimates and find their variance-covariance matrix. In the second step, we find the linear combination that explains as much of the variance of the de-meaned models as possible. The weights in the linear combination are the elements of the eigenvector associated with the largest eigenvalue of the variance-covariance matrix found in the first step. In the third step, we add to the linear combination just obtained, which has mean zero, the average of ERP estimates across all models and all time periods. Under the assumption that each of the models is an unbiased and consistent estimator of the ERP, the average across all models and all time periods is an unbiased and consistent estimator of the unconditional mean of the ERP. The time

[14] See, *inter alia*, Clemen (1989), Diebold and Lopez (1996) and Timmermann (2006).

variation in the first principal component then provides an estimate of the conditional ERP[15]. The share of the variance of the underlying models explained by this principal component is 76 percent, suggesting that there is not too much to gain from examining principal components beyond the first[16].

We now focus on the one-year-ahead ERP estimates and study other horizons in the next section.

The first two columns in Table VII show the mean and standard deviation of each model's estimates. The unconditional mean of the ERP across all models is 5.7 percent, with an average standard deviation of 3.2 percent. DDM give the lowest mean ERP estimates and have moderate standard deviations. In contrast, cross-sectional models tend to have mean ERP estimates on the high end of the distribution and very smooth time-series. Mean ERP estimates for time-series regressions are mixed, with high and low values depending on the predictors used, but uniformly large variances. The survey of CFOs has a mean and standard deviation that are both about half as large as in the overall population of models. The picture that emerges from Table VII is that there is considerable heterogeneity across model types, and even sometimes within model types, thereby underscoring the difficulty inherent in finding precise estimates of the ERP.

---

[15] As is customary in the literature, we perform the analysis using ERP estimates in levels, even though they are quite persistent. Results in first-differences do not give economically reasonable estimates since they feature a pro-cyclical ERP and unreasonable magnitudes.

One challenge that arises in computing the principal component is when we have missing observations, either because some models can only be obtained at frequencies lower than monthly or because the necessary data is not available for all time periods (Appendix A contains a detailed description of when this happens). To overcome this challenge, we use an iterative linear projection method, which conceptually preserves the idea behind principal components. Let X be the matrix that has observations for different models in its columns and for different time periods in its rows. On the first iteration, we make a guess for the principal component and regress the non-missing elements of each row of X on the guess and a constant. We then find the first principal component of the variance-covariance matrix of the fitted values of these regressions, and use it as the guess for the next iteration. The process ends when the norm of the difference between consecutive estimates is small enough. We thank Richard Crump for suggesting this method and providing the code for its implementation.

[16] The second and third principal components account for 13 and 8 percent of the variance, respectively.

[Insert Table VII here]

Figure 1 shows the time-series for all one-year-ahead ERP model estimates, with each class of models in a different panel. The green lines are the ERP estimates from the twenty underlying models. The black line, reproduced in each of the panels, is the principal component of all twenty models. The shaded areas are NBER recessions. The figure gives a sense of how the time-series move together, and how much they co-vary with the first principal component. Table VIII shows the correlations among models. Figure 1 and Table VIII give the same message: despite some outliers, there is a fairly strong correlation within each of the five classes of models. Across classes, however, correlations are small and even negative. Interestingly, the correlation between some DDM and cross-sectional models is as low as -91 percent. This negative correlation, however, disappears if we look at lower frequencies. When aggregated to quarterly frequency, the smallest correlation between DDM and cross-sectional models is -22 percent, while at the annual frequency it is 12 percent.

[Insert Figure 1 here]

[Insert Table VIII here]

Figure 1 also shows that the first principal component co-varies negatively with historical mean models, but positively with DDM and cross-sectional regression models. Time-series regression models are also positively correlated with the first principal component, although this is not so clearly seen in Panel 4 of Figure 1 because of the high volatility of time-series ERP estimates. The last panel shows that the survey of CFOs does track the first principal component quite well at low frequencies (e.g. annual), although any conclusions about survey estimates should be interpreted with caution given the short length of the sample.

As explained earlier, the first principal component is a linear combination of the twenty underlying ERP models:

$$PC_t^{(1)} = \sum_{m=1}^{20} w^{(m)} ERP_t^{(m)}. \tag{9}$$

In the above equation, m indexes the different models, $PC_t^{(1)}$ is the first principal component, $ERP_t^{(m)}$ is the estimate from model m and $w^{(m)}$ is the weight that the principal component places on model m. The third column in Table VII, labeled "PC coefficients", shows the weights $w^{(m)}$ normalized to sum up to one to facilitate comparison, i.e. the table reports the weights $\widehat{w}^{(m)}$ where

$$\widehat{w}^{(m)} = \frac{w^{(m)}}{\sum_{m=1}^{20} w^{(m)}}. \tag{10}$$

The first principal component puts positive weight on models based on the historical mean, cross-sectional regressions and the survey of CFOs. It weights DDM and time-series regressions mostly negatively. The absolute values of the weights are very similar for many of the models, and there is no single model or class of models that dominates. This means that the first principal component uses information from many of the models.

The last column in Table VII, labeled "Exposure to PC", shows the extent to which models *load* on the first principal component. By construction, each of the twenty ERP models can be written as a linear combination of twenty principal components:

$$ERP_t^{(m)} = \sum_{i=1}^{20} load_i^{(m)} PC_t^{(i)}, \tag{11}$$

where m indexes the model and i indexes the principal components. The values in the last column of Table VII are the loadings on the first principal component ($i = 1$) for each model ($m = 1, 2, \dots, 20$), again normalized to one for ease of comparability:

$$\widehat{load}_1^{(m)} = \frac{load_1^{(m)}}{\sum_{m=1}^{20} load_i^{(m)}}. \tag{12}$$

Most models have a positive loading on the first principal component; whenever the loading is negative, it tends to be relatively small. This means the first principal component, as expected, is a good explanatory variable for most models. Looking at the third and fourth columns of Table VII together, we can obtain additional information. For example, a model with a very high loading (fourth column) accompanied by a very small PC coefficient (third column) is likely to mean that the model is almost redundant, in the sense that it is close to being a linear combination of all other models and does not provide much independent information to the principal component. On the other hand, if the PC coefficient and loading are both high, the corresponding model is likely providing information not contained in other measures.

Figure 2 shows the first principal component of all twenty models in black, with recessions indicated by shaded bars (the black line is the same principal component shown in black in each of the panels of Figure 1). As expected, the principal component tends to peak during financial turmoil, recessions and periods of low real GDP growth or high inflation. It tends to bottom out after periods of sustained bullish stock markets and high real GDP growth. Evaluated by the first principal component, the one-year-ahead ERP reaches a local peak in June of 2012 at 12.2 percent. The surrounding months have ERP estimates of similar magnitude, with the most recent estimate in June 2013 at 11.2 percent. This behavior is not so clearly seen by simply looking at the collection of individual models in Figure 1, highlighting the usefulness of principal components analysis. Similarly high levels were seen in the mid and late 1970s, during a period of stagflation, while the recent financial crisis had slightly lower ERP estimates closer to 10 percent.

[Insert Figure 2 here]

Figure 2 also displays the 10^{th}, 25^{th}, 75^{th} and 90^{th} percentiles of the cross-sectional distribution of models. These bands can be interpreted as confidence intervals, since they give the range of the distribution of ERP estimates at each point in time. However, they do not incorporate other relevant sources of uncertainty, such as the errors that occur during the estimation of each individual model, the degree of doubt in the correctness of each model, and the correlation structure between these and all other kinds of errors. Standard error bands that capture all sources of uncertainty are therefore likely to be wider.

The difference in high and low percentiles can also be interpreted as measures of agreement across models. The interquartile range –the difference between the 25^{th} and 75^{th} percentiles— has compressed, mostly because the models in the bottom of the distribution have had higher ERP estimates since 2010. It is also interesting to note that the 75^{th} percentile has remained fairly constant over the last 10 years at a level somewhat below its long-run mean. The cross-sectional standard deviation in ERP estimates (not shown in the graph) also decreased from 10.2% in January of 2000 to 4.3% in June of 2013, confirming that the disagreement among models has decreased.

Another *a priori* reasonable summary statistic for the ERP is the cross-sectional mean of estimates across models. In Figure 3, we can see that by this measure the ERP has also been increasing since the crisis. However, unlike the principal component, it has not reached elevated levels compared to past values. The cross-sectional mean can be useful, but it has a few undesirable features as an overall measure of the ERP compared to the first principal component. First, it is procyclical, which contradicts the economic intuition that expected returns are highest in recessions, when risk aversion is high and future prospects look brighter than current ones. Second, it overloads on DDM simply because there is a higher number of DDM models in our sample. Lastly, it has a smaller correlation with the realized returns it is supposed to predict.

[Insert Figure 3 here]

5. The Term Structure of Equity Risk Premia

In Section 2, we described the term structure of the ERP – what expected excess returns are over different investment horizons. In practical terms, we estimate the ERP at different horizons by using the inputs for all the models at the corresponding horizons[17]. For example, if we want to take the historical mean of returns as our estimate, we can take the mean of returns over one month, six months, or a one-year period. In cross-sectional and time-series regressions, we can predict monthly, quarterly or annual returns using monthly, quarterly or annual right-hand side variables. DDM, on the other hand, have little variation across horizons. In fact, all the DDM we consider have a constant term structure of expected stock returns, and the only term structure variation in ERP estimates comes from risk-free rates[18].

Figure 4 plots the first principal components of the ERP as a function of investment horizon for some selected dates. We picked the dates because they are typical dates for when the ERP was unusually high or unusually low at the one-month horizon. As was the case for one-year-ahead ERP estimates, we can capture the majority of the variance of the underlying models at all horizons by a single principal component. The shares of the variance explained by the first principal components at horizons of one month to three years range between 68 and 94 percent. The grey line in Figure 4 shows the average of the term structure across all periods. It is slightly upward sloping, with a short-term ERP at just over 6 percent and a three-year ERP at almost 7 percent.

[Insert Figure 4 here]

[17] For other ways to estimate the term structure of the ERP using equilibrium models or derivatives, see Ait-Sahalia, Karaman and Mancini (2014), Ang and Ulrich (2012), van Binsbergen, Hueskes, Koijen and Vrugt (2014), Boguth, Carlson, Fisher and Simutin (2012), Durham (2013), Croce, Lettau and Ludvigson (2014), Lemke and Werner (2009), Lettau and Wachter (2011), Muir (2013), among others.

[18] In equation (3), ρ_{t+k} is assumed to be the same for all k, while risk-free rates are allowed to vary over the investment horizon k in equation (4). Of course, with additional assumptions, it is possible to have DDM with a non-constant term structure of expected excess returns.

The first observation is that the term structure of the ERP has significant time variation and can be flat, upward or downward sloping. Figure 4 also shows some examples that hint at lower future expected excess returns when the one-month-ahead ERP is elevated and the term structure is downward sloping, and higher future expected excess returns when the one-month-ahead ERP is low and the term structure is upward sloping. In fact, this is generally true: There is a strong negative correlation between the level and the slope of the ERP term structure of -71 percent. Figure 5 plots monthly observations of the one-month-ahead ERP against the slope of the ERP term structure (the three-year-ahead minus the one-month-ahead ERP) together with the corresponding ordinary least squares regression line in black. Of course, this is only a statistical pattern and should not be interpreted as a causal relation.

[Insert Figure 5 here]

6. *Why is the Equity Risk Premium High?*

There are two reasons why the ERP can be high: low discount rates and high current or expected future cash flows.

Figure 6 shows that earnings are unlikely to be the reason why the ERP is high. The green line shows the year-on-year change in the mean expectation of one-year-ahead earnings per share for the S&P 500. These expectations are obtained from surveys conducted by the Institutional Brokers' Estimate System (I/B/E/S) and available from Thomson Reuters. Expected earnings per share have been declining from 2010 to 2013, making earnings growth an unlikely reason for why the ERP was high in the corresponding period. The black line shows the realized monthly growth rates of real earnings for the S&P 500 expressed in annualized percentage points. Since 2010, earnings growth has been declining, hovering around zero for the last few months of the sample. It currently stands at 2.5 percent, which is near its long-run average.

[Insert Figure 6 here]

Another way to examine whether a high ERP is due to discount rates or cash flows is shown in Figure 7. The black line is the same one-year-ahead ERP estimate shown in Figure 2. The green line simply adds the realized one-year Treasury yield to obtain expected stock returns. The figure shows expected stock returns have increased since 2000, similarly to the ERP. However, unlike the ERP, expected stock returns are close to their long-run mean, and nowhere near their highest levels, achieved in 1980. The discrepancies between the two lines are due to exceptionally low bond yields since the end of the financial crisis.

[Insert Figure 7 here]

Figure 8 displays the term structure of the ERP under a simple counterfactual scenario, in addition to the mean and current term structures already displayed in Figure 4. In this scenario, we leave expected stock returns unmodified but change the risk-free rates in June 2012 from their actual values to the average nominal bond yields over 1960-2013. In other words, we replace R_{t+k}^{f} in equation (2) by the mean of R_{t+k}^{f} over t. The result of this counterfactual is shown in Figure 8 in green. Using average levels of bond yields brings the whole term structure of the ERP much closer to its mean level (the grey line), especially at intermediate horizons. This shows that a “normalization” of bond yields, everything else being equal, would bring the ERP close to its historical norm. This exercise shows that the current environment of low bond yields is capable, quantitatively speaking, of significantly contributing to an ERP as high as was observed in 2012-2013.

[Insert Figure 8 here]

7. Conclusion

We have analyzed twenty different models of the ERP by considering the assumptions and data required to implement them, and how they relate to each other. When it comes to the ERP, we find that there is substantial heterogeneity in estimation methodology and final estimates. We then extract the first principal component of the twenty models, which signals that the ERP in 2012 and 2013 is at heightened levels compared to previous periods. Our analysis provides evidence that the current level of the ERP is consistent with a bond-driven ERP: expected excess stock returns are elevated not because stocks are expected to have high returns, but because bond yields are exceptionally low. The models we consider suggest that expected stock returns, on their own, are close to average levels.

References

Adrian, T., D. M. Covitz, and N. Liang. 2013. "Financial stability monitoring." Federal Reserve Bank of New York Staff Reports, no. 601 (February).

Adrian, T., R. Crump, and E. Moench. 2014. "Regression-based estimation of dynamic asset pricing models." Federal Reserve Bank of New York Staff Reports, no. 493 (June).

Ait-Sahalia, Y., M. Karaman, and L. Mancini. 2014. "The term structure of variance swaps, risk premia and the expectation hypothesis. Risk Premia and the Expectation Hypothesis." (May 10, 2014). Available at SSRN: http://ssrn.com/abstract=2136820

Ang, A., and G. Bekaert. 2007. "Stock return predictability: Is it there?" Review of Financial Studies 20, no. 3 (May): 651-707.

Ang, A., and M. Ulrich. 2012. "Nominal Bonds, Real Bonds, and Equity." Unpublished paper, Columbia University.

Baker, M., and J. Wurgler. 2007. "Investor sentiment in the stock market." The Journal of Economic Perspectives: 129-51.

van Binsbergen, J., W. Hueskes, R. Koijen, and E. Vrugt. 2013. "Equity yields." Journal of Financial Economics 110, no. 3 (December): 503-19.

Bloom, N. 2009. "The impact of uncertainty shocks." Econometrica 77, no. 3 (May): 623-85.

Boguth, O., M. Carlson, A. J. Fisher, and M. Simutin. 2012. "Leverage and the limits of arbitrage pricing: Implications for dividend strips and the term structure of equity risk premia." Unpublished paper, Arizona State University.

Campbell, J.Y., and S. Thompson. 2008. "Predicting excess stock returns out of sample: Can anything beat the historical average?" Review of Financial Studies 21, no. 4 (July): 1509-31.

Carhart, M. M. 1997. "On persistence in mutual fund performance." Journal of Finance 52, no 1 (March): 57–82.

Chen, L., Z. Da, and X. Zhao. 2013. "What drives stock price movements?" Review of Financial Studies 26, no. 4 (April): 841-76.

Clemen, R.T., 1989. "Combining forecasts: a review and annotated bibliography". International Journal of Forecasting 5 (4), 559-83.

Croce, M., M. Lettau, and S. Ludvigson. Forthcoming. 2014. "Investor information, long-run risk, and the term structure of equity." Review of Financial Studies.

Damodaran, A. 2012. “Equity Risk Premiums (ERP): Determinants, Estimation and Implications – The 2012 Edition.” Unpublished paper, New York University.

Diebold, F.X., and J. Lopez, 1996. “Forecast Evaluation and Combination”. In G.S. Maddala and C.R. Rao (eds.), Handbook of Statistics, 241-268, 1996. Amsterdam: North-Holland.

Duarte, F. 2013. “Inflation and the cross-section of stock returns.” Federal Reserve Bank of New York Staff Reports, no. 621 (May).

Durham, J. B. 2013. “Arbitrage-free models of stocks and bonds.” Federal Reserve Bank of New York Staff Reports, no. 656 (December).

Duarte, F., L. Kogan, and D. Livdan. 2013. “Aggregate Investment and Stock Returns.” Unpublished paper, M.I.T.

Easton, P.D., and G. A. Sommers. 2007. “Effect of analysts’ optimism on estimates of the expected return implied by earnings forecasts.” Journal of Accounting Research 45, no. 5 (December): 983-1016.

Fama, E.F., and K. R. French. 1988. “Dividend yields and expected stock returns.” Journal of Financial Economics 22, no. 1 (October), 3-25.

Fama, E.F., and K. R. French. 1992. “The cross-section of expected stock returns.” Journal of Finance 47, no. 2 (June): 427-65.

Fama, E.F., and K. R. French. 2002. “The equity premium.” Journal of Finance 57, no. 2 (April): 637-59.

Geiger, B., and G. Kubin. 2013. "Signal Enhancement as Minimization of Relevant Information Loss". Proceedings of ITG Conference on Systems, Communication and Coding

Gordon, M. J. 1962. “The investment, financing, and valuation of the corporation.” Greenwood Press.

Goyal, A., and I. Welch. 2008. “A comprehensive look at the empirical performance of equity premium predictions.” Review of Financial Studies 21, no. 4 (July): 1455-1508.

Graham, J., and C. Harvey. 2012. “The Equity Risk Premium in 2012.” Unpublished paper, Duke University.

Greenwood, R., and A. Shleifer. 2014. “Expectations of returns and expected returns.” Review of Financial Studies 27, no. 3 (March): 714-46

Gurkaynak, R. S., B. Sack, and J. H. Wright. 2007. “The US Treasury yield curve: 1961 to the present.” Journal of Monetary Economics 54, no. 8 (November): 2291-2304.

Gurkaynak, R. S., S. Refet, B. Sack, and J. H. Wright. 2010. “The TIPS yield curve and inflation compensation.” American Economic Journal: Macroeconomics 2, no. 1 (January): 70-92.

Hall, R. E. Forthcoming. 2014. "The Routes into and out of the Zero Lower Bound." Federal Reserve Bank of Kansas City Symposium (Jackson Hole).

Harvey, C. R., and Y. Liu, and H. Zhu. 2014. "…and the Cross-Section of Expected Returns." (October 7, 2014). Available at SSRN: http://ssrn.com/abstract=2249314

Kelly, B., and S. Pruitt. 2013. "Market Expectations in the Cross-Section of Present Values." The Journal of Finance 68, no. 5 (October): 1721-56.

Kuehn, L. A., N. Petrosky-Nadeau, and L. Zhang. 2012. "An equilibrium asset pricing model with labor market search." NBER Working Paper no. 17742.

Lemke, W., and T. Werner. 2009. "The term structure of equity premia in an affine arbitrage-free model of bond and stock market dynamics." European Central Bank, Working Paper No.1045.

Lettau, M., and J. A. Wachter. 2011. "The term structures of equity and interest rates." Journal of Financial Economics 101, no. 1 (July): 90-113.

Merton, R.C. 1973. "An intertemporal capital asset pricing model." Econometrica 41, no. 5 (September): 867-87.

Muir, T. 2013. "Financial crises, risk premia, and the term structure of risky assets." Unpublished paper, Yale School of Management.

Neely, C. J., D.E. Rapach, J. Tu, and G. Zhou. 2014. "Forecasting the equity risk premium: the role of technical indicators." Management Science 60, no. 7 (July): 1772-91.

Panigirtzoglou, N., and J. Loeys. 2005. "A fair value model for US bonds, credit and equities." JP Morgan (January).

Piazzesi, M., and M. Schneider. 2007. "Equilibrium yield curves." In *NBER Macroeconomics Annual 2006* 21: 389-472. MIT Press, 2007.

Polk, C., S. Thompson, and T. Vuolteenaho. 2006. "Cross-sectional forecasts of the equity premium." Journal of Financial Economics 81, no. 1 (July): 101-41.

Ross, S. 1976. "The arbitrage theory of capital asset pricing." Journal of Economic Theory 13, no. 3 (December): 341–60.

Shiller, R. J. 2005. "Irrational Exuberance". Princeton University Press 2000, Broadway Books 2001, 2nd ed.

Stock, J.H., and M. W. Watson. 2003. "Forecasting output and inflation: The role of asset prices." Journal of Economic Literature 41, no. 3 (September): 788-829.

Timmermann, A. 2006. “Forecast combinations”. Handbook of Economic Forecasting, Volume 1 (chapter 4), 135-96.

Appendix A: Data Variables

Fama and French (1992)	http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html Monthly frequency; 1/1/1960 to 6/30/2013. We use 25 portfolios sorted on size and book to market, 10 portfolios sorted on momentum, realized excess market returns, HML, SMB, and the momentum factor.
Shiller (2005)	http://www.econ.yale.edu/~shiller/data.htm Monthly frequency; 1/1/1960 to 6/30/2013. We use the nominal and real price, nominal and real dividends and nominal and real earnings for the S&P 500, CPI, and 10 year nominal treasury yield.
Baker and Wurgler (2007)	http://people.stern.nyu.edu/jwurgler/data/Investor_Sentiment_Data_v23_POST.xlsx Monthly frequency; 7/1/1965 to 12/1/2010. We use the "sentiment measure".
Graham and Harvey (2012)	http://www.cfosurvey.org/index.htm Quarterly frequency; 6/6/2000 to 6/5/2013. We use the answer to the question "Over the next 10 years, I expect the average annual S&P 500 return will be: Expected return:" and the analogous one that asks about the next year.
Damodaran (2012)	http://www.stern.nyu.edu/~adamodar/pc/datasets/histimpl.xls Annual frequency; 1/1/1960 to 12/1/2012. We use the ERP estimates from his dividend discount models (one uses free-cash flow, the other one doesn't).
Gurkaynak, Sack and Wright (2007)	http://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html Daily frequency; starting on 6/14/61 for one- to seven-year yields; 8/16/71 for nine- and ten-year yields; 11/15/71 for eleven- to fifteen-year yields; 7/2/81 for sixteen- to twenty-year yields; 11/25/85 for twenty-one- to thirty-year yields. We use all series until 6/30/2013.
Gurkaynak, Refet, Sack and Wright (2010)	http://www.federalreserve.gov/econresdata/researchdata.htm Monthly frequency; 1/1/2003 to 7/1/2013. We use yields on TIPS of all maturities available.
Compustat	Variable BKVLPS Annual frequency; 12/31/1977 to 12/31/2012.
Thomson Reuters I/B/E/S	Variables EPS 1 2 3 4 5 Monthly frequency; 1/14/1982 to 4/18/2013 for current and next year forecasts; 9/20/84 to 4/18/2013 for two-year-ahead forecasts; 9/19/85 to 3/15/2012 for three-year-ahead forecasts; 2/18/88 to 3/15/07 for four-year-ahead forecasts.
FRED (St. Louis Federal Reserve)	http://research.stlouisfed.org/fred2/graph/?g=D9J and http://research.stlouisfed.org/fred2/graph/?g=KKk Monthly frequency. 1/1/1960 to 7/1/2013 for Baa minus Aaa bond yield spread and recession indicator.

Tables and Figures

Table I: Data sources	
Fama and French (1992)	Fama-French factors, momentum factor, twenty-five portfolios sorted on size and book-to-market
Shiller (2005)	Inflation and ten-year nominal treasury yield. Nominal price, real price, earnings, dividends and cyclically adjusted price-earnings ratio for the S&P 500
Baker and Wurgler (2007)	Debt issuance, equity issuance, sentiment measure
Graham and Harvey (2012)	ERP estimates from the Duke CFO survey
Damodaran (2012)	ERP estimates
Gurkaynak, Sack and Wright (2007)	Zero coupon nominal bond yields for all maturities[19]
Gurkaynak, Refet, Sack and Wright (2010)	Zero coupon TIPS yields for all maturities
Compustat	Book value per share for the S&P 500
Thomson Reuters I/B/E/S	Mean analyst forecast of expected earnings per share
FRED (St. Louis Federal Reserve)	Corporate bond Baa-Aaa spread and the NBER recession indicator

Note: All variables start in January 1960 (or later, if unavailable for early periods) and end in June 2013 (or until no longer available). CFO surveys are quarterly; book value per share and ERP estimates by Damodaran (2012) are annual; all other variables are monthly. Appendix A provides more details.

[19] Except for the 10-year yield, which is from Shiller (2005). We use the 10-year yield from Shiller (2005) for ease of comparability with the existing literature. Results are virtually unchanged if we use all yields, including the 10-year yield, from Gurkaynak, Sack and Wright (2007).

Table II: Models based on the historical mean of realized returns	
Long-run mean	Average of realized S&P 500 returns minus the risk-free rate using all available historical data
Mean of the previous five years	Average of realized S&P 500 returns minus the risk-free rate using only data for the previous five years

Table III: Dividend Discount Models	
Gordon (1962) with nominal yields	S&P 500 dividend-to-price ratio minus the ten-year nominal Treasury yield
Shiller (2005)	Cyclically adjusted price-earnings ratio (CAPE) minus the ten-year nominal Treasury yield
Gordon (1962) with real yields	S&P 500 dividend-to-price ratio minus the ten year real Treasury yield (computed as the ten-year nominal Treasury rate minus the ten year breakeven inflation implied by TIPS)
Gordon (1962) with earnings forecasts	S&P 500 expected earnings-to-price ratio minus the ten-year nominal Treasury yield
Gordon (1962) with real yields and earnings forecasts	S&P 500 expected earnings-to-price ratio minus the ten-year real Treasury yield (computed as the ten-year nominal Treasury rate minus the ten-year breakeven inflation implied by TIPS)
Panigirtzoglou and Loeys (2005)	Two-stage DMM. The growth rate of earnings over the first five years is estimated by using the fitted values in a regression of average realized earnings growth over the last five years on its lag and lagged earnings-price ratio. The growth rate of earnings from years six and onwards is 2.2 percent
Damodaran (2012)	A six-stage DDM. Dividend growth the first five stages are estimated from analyst's earnings forecasts. Dividend growth in the sixth stage is the ten-year nominal Treasury yield
Damodaran (2012) free cash flow	Same as Damodaran (2012), but uses free-cash-flow-to-equity as a proxy for dividends plus stock buybacks

Table IV: Models with cross-sectional regressions	
Fama and French (1992)	Uses the excess returns on the market portfolio, a size portfolio and a book-to-market portfolio as risk factors
Carhart (1997)	Identical to Fama and French (1992) but adds the momentum measure of Carhart (1997) as an additional risk factor
Duarte (2013)	Identical to Carhart (1997) but adds an inflation risk factor
Adrian, Crump and Moench (2014)	Uses the excess returns on the market portfolio as the single risk factor. The state variables are the dividend yield, the default spread, and the risk free rate

Table V: Models with time-series regressions	
Fama and French (1988)	Only predictor is the dividend-price ratio of the S&P 500
Goyal and Welch (2008)	Uses, at each point in time, the best out-of-sample predictor out of twelve predictive variables proposed by Goyal and Welch (2008)
Campbell and Thompson (2008)	Same as Goyal and Welch (2008), but imposes two restrictions on the estimation. First, the coefficient *b* in equation (9) is replaced by zero if it has the "wrong" theoretical sign. Second, we replace the estimate of the ERP by zero if the estimation otherwise finds a negative ERP
Fama and French (2002)	Uses, at each point in time, the best out-of-sample predictor out of three variables: the price-dividend ratio adjusted by the growth rate of earnings, dividends or stock prices
Baker and Wurgler (2007)	The predictor is Baker and Wurgler's (2007) sentiment measure. The measure is constructed by finding the most predictive linear combination of five variables: the closed-end fund discount, NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium

Table VI: Surveys	
Graham and Harvey (2012)	Chief financial officers (CFOs) are asked since 1996 about the one and ten-year-ahead ERP. We take the mean of all responses

Table VII: ERP models					
		Mean	Std. dev.	PC coefficients $\widehat{w}^{(m)}$	Exposure to PC $\widehat{load}_1^{(m)}$
Based on historical mean	Long-run mean	9.3	1.3	0.78	-0.065
	Mean of previous five years	5.7	5.8	0.42	-0.160
DDM	Gordon (1926): E/P minus nominal 10yr yield	-0.1	2.1	-0.01	0.001
	Shiller (2005): 1/CAPE minus nominal 10yr yield	-0.4	1.8	-0.10	0.011
	Gordon (1962): E/P minus real 10yr yield	3.5	2.1	0.69	-0.077
	Gordon (1962): Expected E/P minus real 10yr yield	5.3	1.7	-0.78	0.208
	Gordon (1962): Expected E/P minus nominal 10yr yield	0.4	2.3	-0.79	0.077
	Panigirtzoglou and Loeys (2005): Two-stage DDM	-1.0	2.3	0.07	-0.011
	Damodaran (2012): Six-stage DDM	3.4	1.3	-0.26	0.032
	Damodaran (2012): Six-stage free cash flow DDM	4.0	1.1	-0.62	0.053
Cross-sectional regressions	Fama and French (1992)	12.6	0.7	0.80	-0.040
	Carhart (1997): Fama-French and momentum	13.1	0.8	0.81	-0.042
	Duarte (2013): Fama-French, momentum and inflation	13.1	0.8	0.82	-0.044
	Adrian, Crump and Moench (2014)	6.5	6.9	-0.05	0.114
Time-series regressions	Fama and French (1988): D/P	2.4	4.0	-0.27	0.069
	Best predictor in Goyal and Welch (2008)	14.5	5.2	-0.07	0.023
	Best predictor in Campbell and Thompson (2008)	3.1	9.8	-0.12	0.081
	Best predictor in Fama French (2002)	11.9	6.8	-0.72	0.321
	Baker and Wurgler (2007) sentiment measure	3.0	4.7	-0.32	0.184
Surveys	Graham and Harvey (2012) survey of CFOs	3.6	1.8	0.72	0.264
	All models	5.7	3.2	0.78	-0.065

For each of the twenty models of the equity risk premium, we show four statistics. The first two are the time-series means and standard deviations for monthly observations from January 1960 to June 2013 (except for surveys, which are quarterly). The units are annualized percentage points. The third statistic, "PC coefficients $\widehat{w}^{(m)}$", is the weight that the first principal component places on each model (normalized to sum to one). The fourth is the "Exposure to PC $\widehat{load}_1^{(m)}$", the weight on the first principal component when each model is written as a weighted sum of all principal components (also normalized to sum to one).

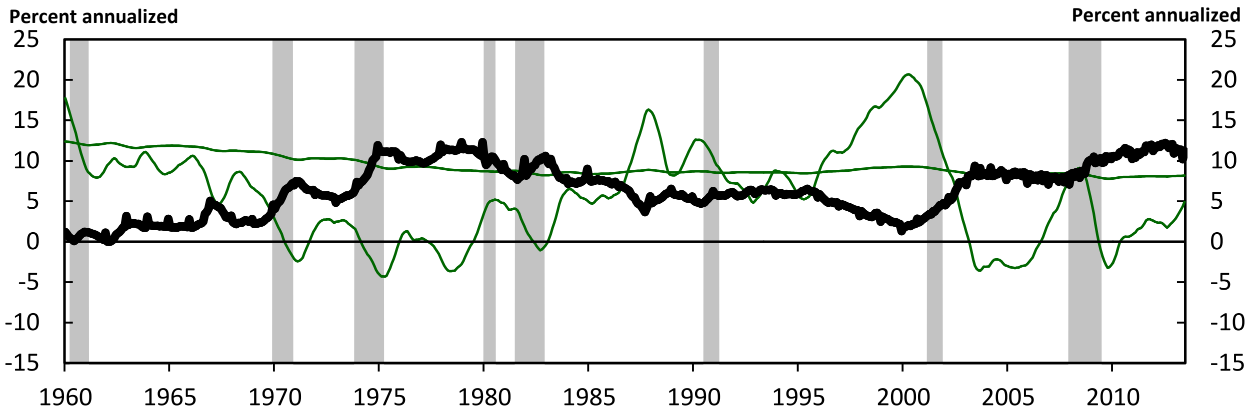
Table VIII: Correlation of ERP models

	LR mean	Mean past 5yr	E/P - 10yr	1/CAPE-10yr	E/P-real 10yr	Exp E/P-real 10yr	Exp E/P- 10yr	Two-stage DDM	Six-stage DDM	Free cash flow	FF	Carhart	Duarte	ACM	D/P	G and W	C and T	FF	Sentiment	CFO Survey
LR mean	100																			
Mean past 5yr	32	100																		
E/P - 10yr	8	15	100																	
1/CAPE-10yr	-9	0	78	100																
E/P-real 10yr	-11	25	98	23	100															
Exp E/P-real 10yr	-58	42	70	84	60	100														
Exp E/P- 10yr	-83	-61	84	95	46	98	100													
Two-stage DDM	17	27	88	54	89	66	79	100												
Six-stage DDM	3	-38	26	39	-30	32	52	-31	100											
Free cash flow	-43	-55	59	70	35	80	94	27	62	100										
FF	69	29	-8	-36	-21	-69	-91	9	-29	-77	100									
Carhart	71	30	-5	-31	-24	-71	-91	10	-25	-75	99	100								
Duarte	71	30	-3	-29	-22	-70	-91	11	-28	-74	99	100	100							
ACM	-1	-52	36	62	6	54	63	27	23	33	-28	-28	-25	100						
D/P	49	12	27	12	27	42	54	24	74	42	44	54	55	21	100					
G and W	25	12	25	21	-7	-36	-60	20	29	-9	7	13	14	-24	61	100				
C and T	27	31	14	-7	81	49	-60	28	-51	-40	60	57	58	-33	54	50	100			
FF	1	-30	-24	-29	37	-27	-37	-18	22	38	36	38	37	-9	40	23	43	100		
Sentiment	-10	33	-4	-20	68	-23	-29	27	-38	-20	18	17	18	-12	-38	-8	21	6	100	
CFO survey	-43	-33	12	30	1	1	13	16	5	-3	-36	-37	-39	60	14	-21	-32	-3	-36	100

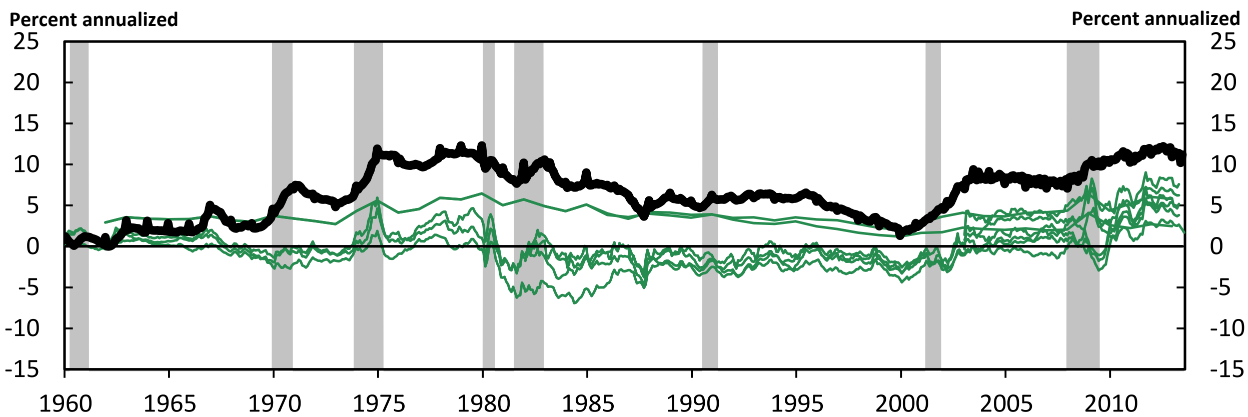
This table shows the correlation matrix of the twenty equity risk premium models we consider. Numbers are rounded to the nearest integer. Thick lines group models by their type (see Tables II to VI). Except for the CFO survey, the observations used to compute correlations are monthly for January 1960 to June 2013. For the CFO survey, correlations are computed by taking the last observation in the quarter for monthly series and then computing quarterly correlations.

Figure 1: ERP estimates for all models

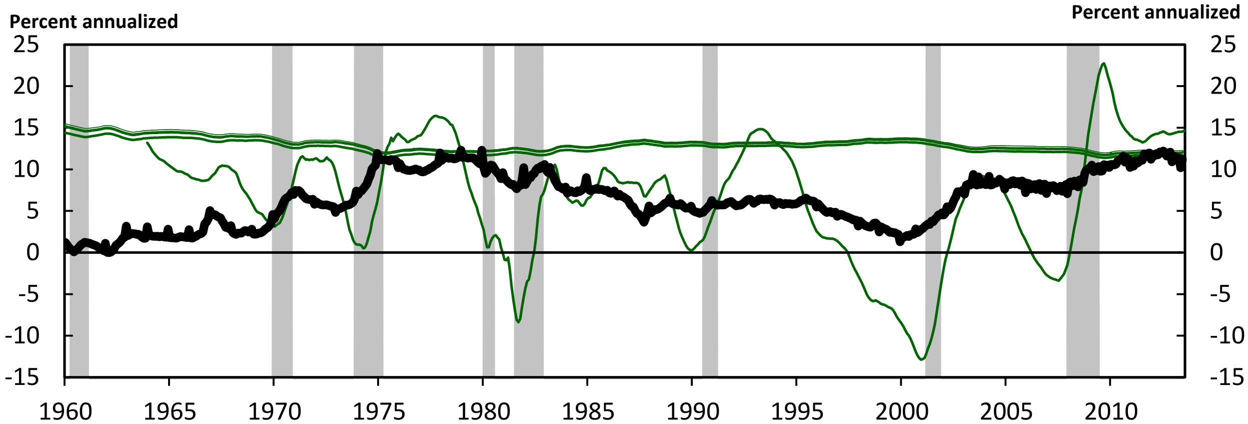
Panel 1: ERP models based on the historical mean of excess returns



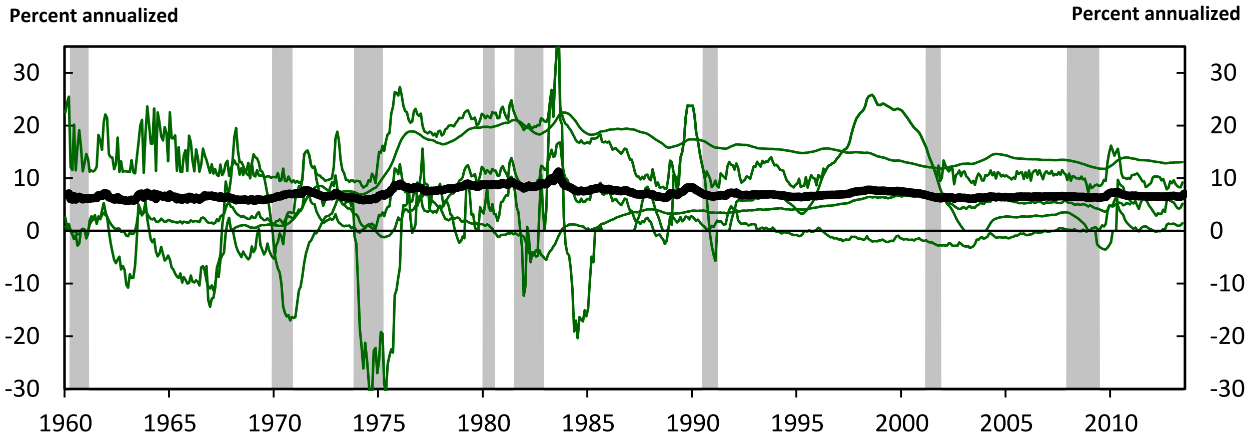
Panel 2: ERP dividend discount models (DDM)



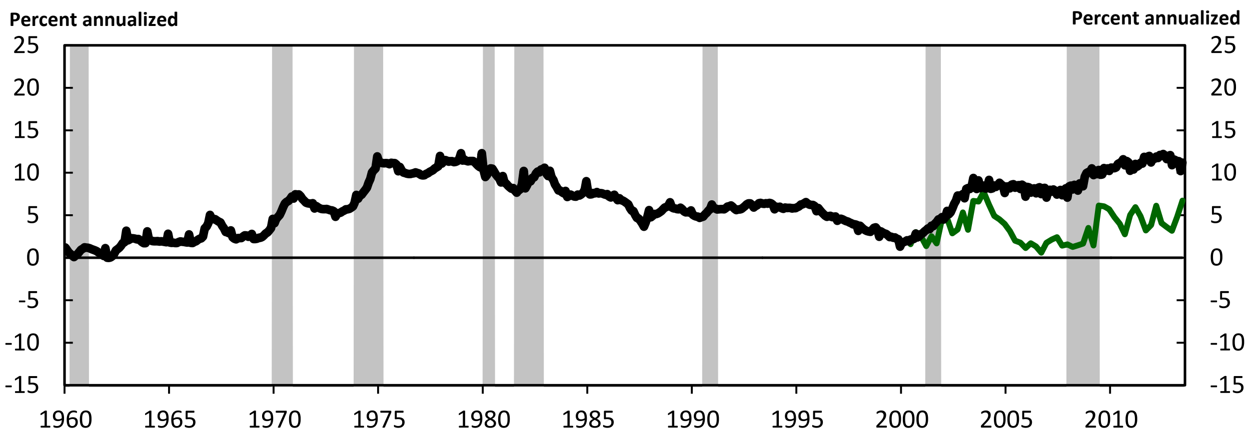
Panel 3: ERP cross sectional models



Panel 4: ERP time series models



Panel 5: ERP surveys

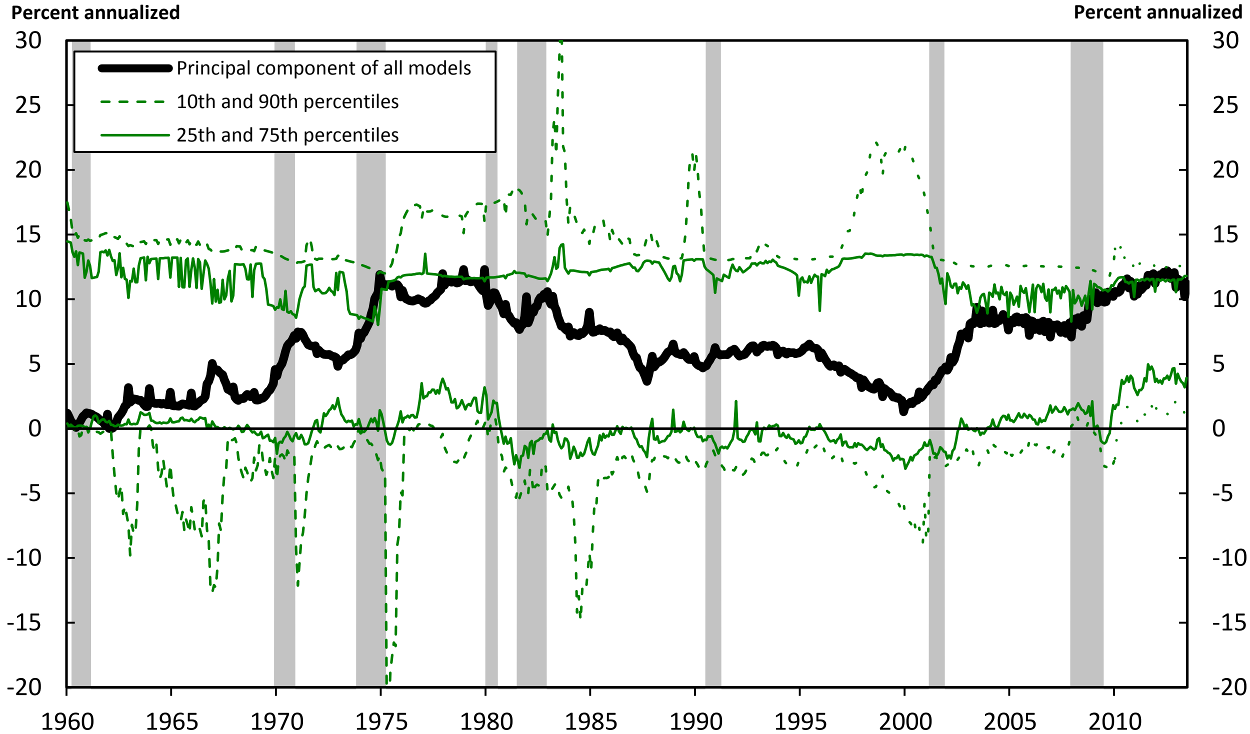


Each green line gives the one-year-ahead equity risk premium from each of the models listed in Tables II to VI. All numbers are in annualized percentage points.

Panel 1 shows the estimates for models based on the historical mean of excess returns, which are listed in Table II. Panel 2 shows estimates computed by the dividend discount models in Table III. Panel 3 uses the cross-sectional regression models from Table IV. Panel 4 shows the equity risk premium computed by the time-series regression models in Table V. Panel 5 gives the estimate obtained from the survey cited in Table VI.

In all panels, the black line is the first principal component of all twenty models (it can look different across panels due to different scales in the y-axis).

Figure 2: One-year-ahead ERP

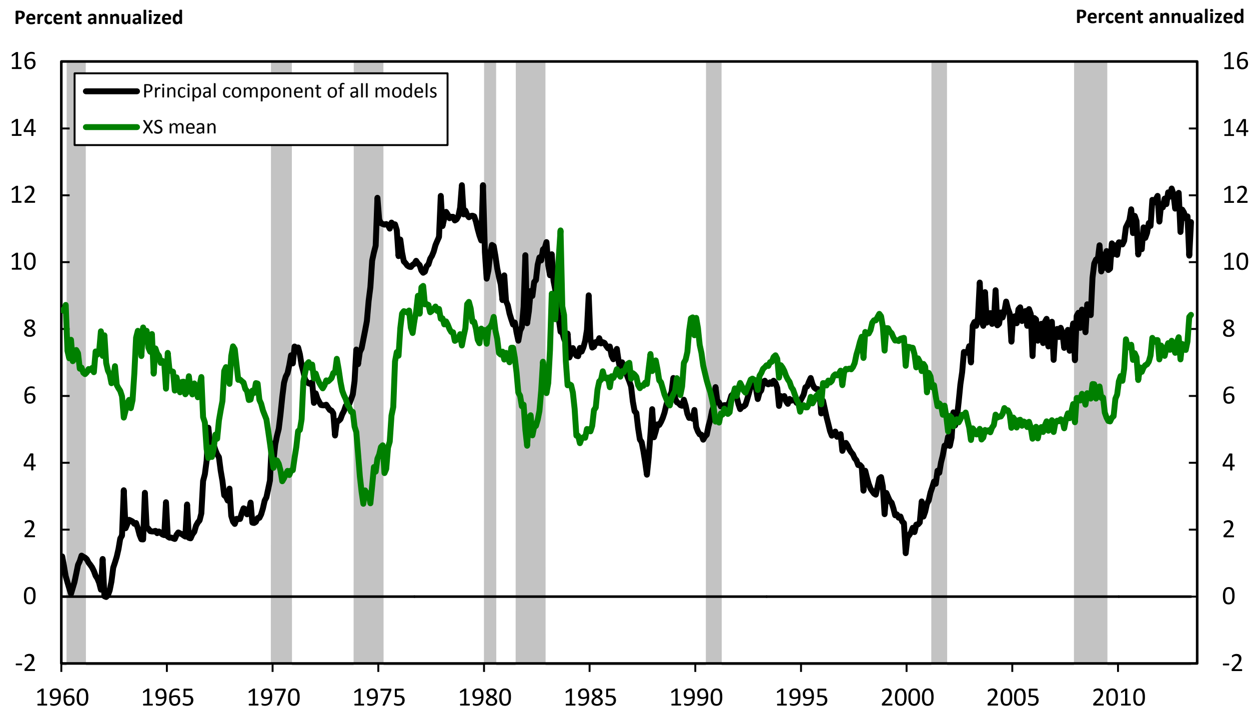


The black line is the first principal component of twenty models of the one-year-ahead equity risk premium (this is the same principal component shown in black in all panels of Figure 1). The models are listed in Tables II to VI.

The 25th and 75th percentiles (solid green lines) give the corresponding quartile of the 20 estimates for each time period, and similarly for the 10th and 90th percentiles (dashed green line).

Shaded bars indicate NBER recessions.

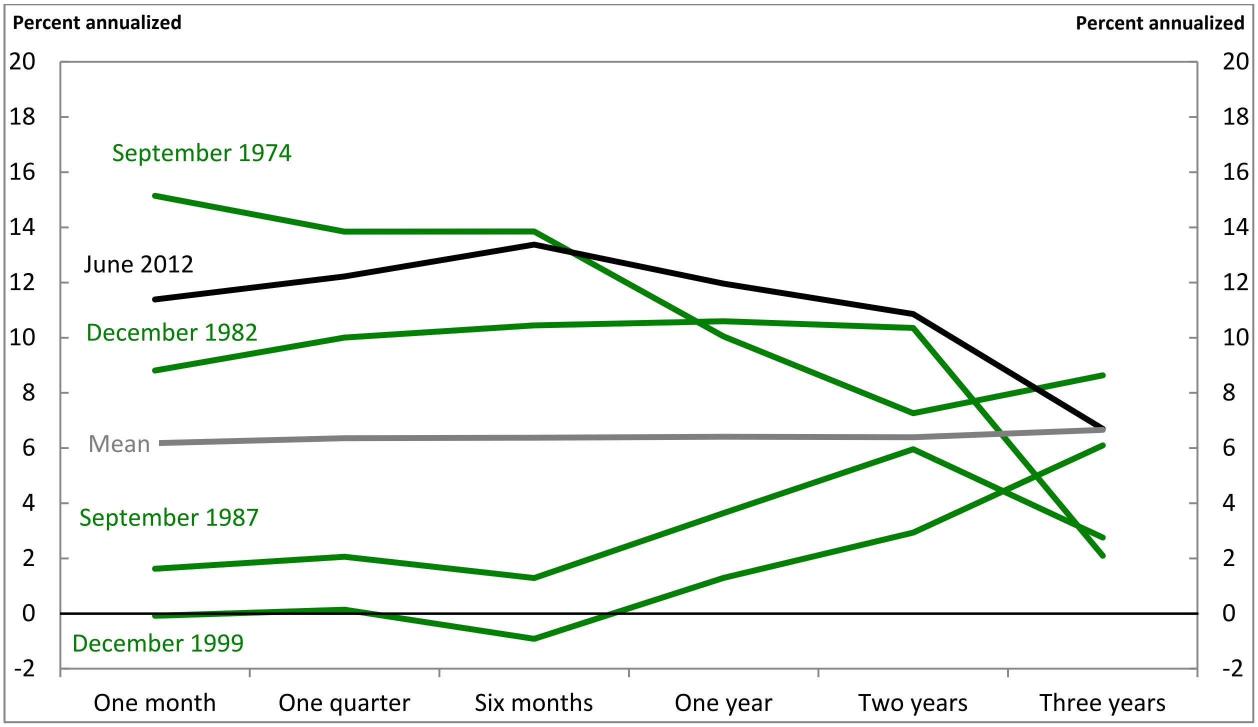
Figure 3: One-year-ahead ERP and cross-sectional mean of models



The black line is the first principal component of twenty models of the one-year-ahead equity risk premium (also shown in Figures 1 and 2). The green line is the cross-sectional average of models for each time period.

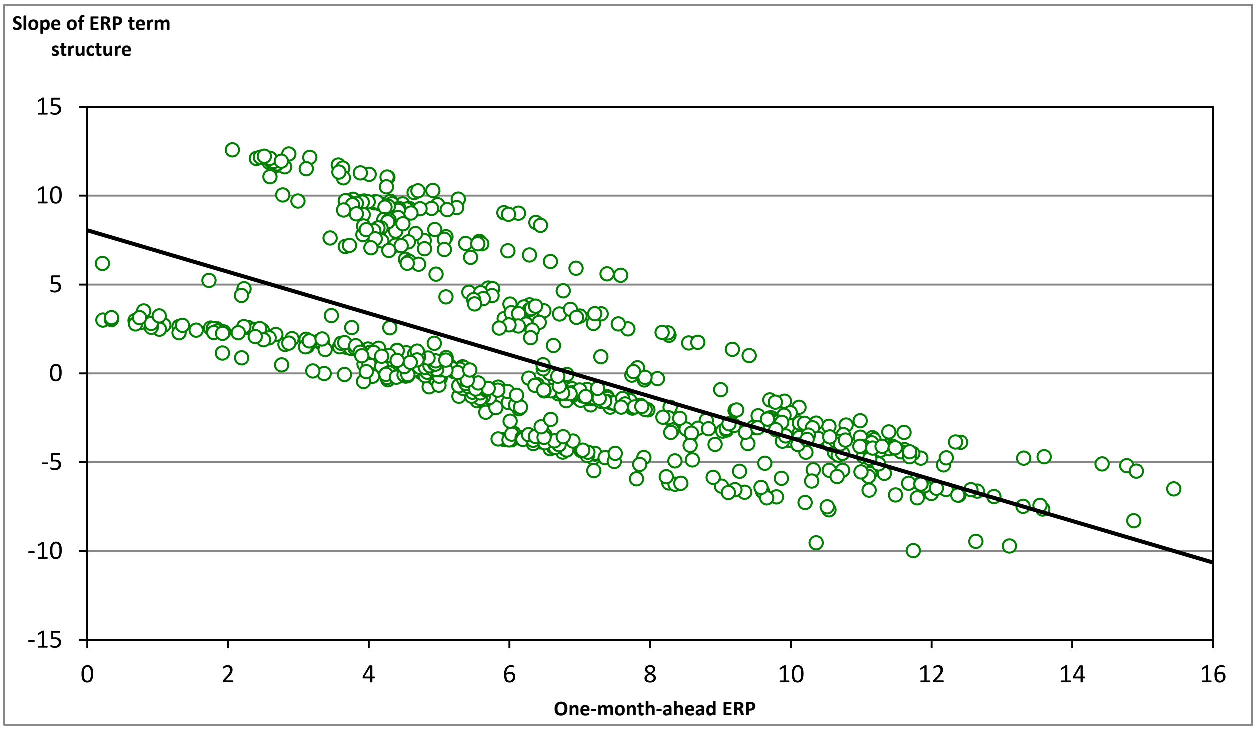
Shaded bars are NBER recessions.

Figure 4: Term structure of the ERP



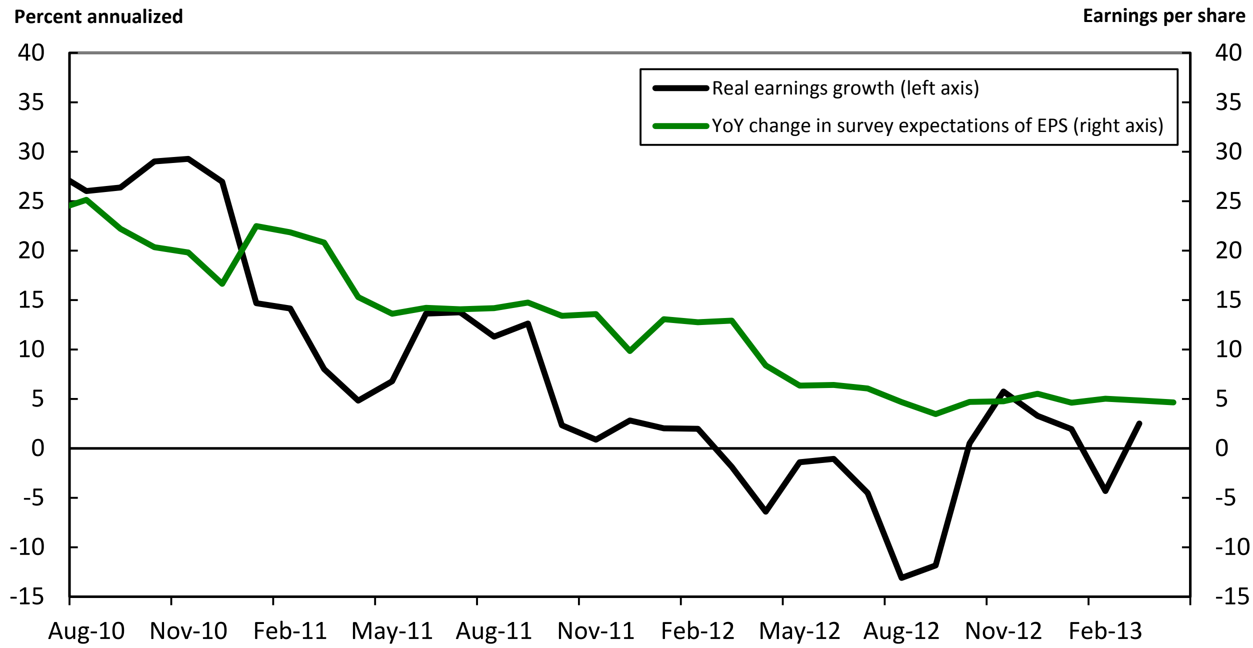
Each line, except for the grey one, shows equity risk premia as a function of investment horizon for some specific months in our sample. We consider horizons of one month, one quarter, six months, one year, two years and three years. The grey line (labeled "Mean") shows the average risk premium at different horizons over the whole sample January 1960 to June 2013. September 1987 and December 1999 were low points in one-month-ahead equity premia. In contrast, September 1974, December 1982 and June 2012 were peaks in the one-month-ahead equity premium.

Figure 5: Regression of the slope of the ERP term structure on one-month-ahead ERP



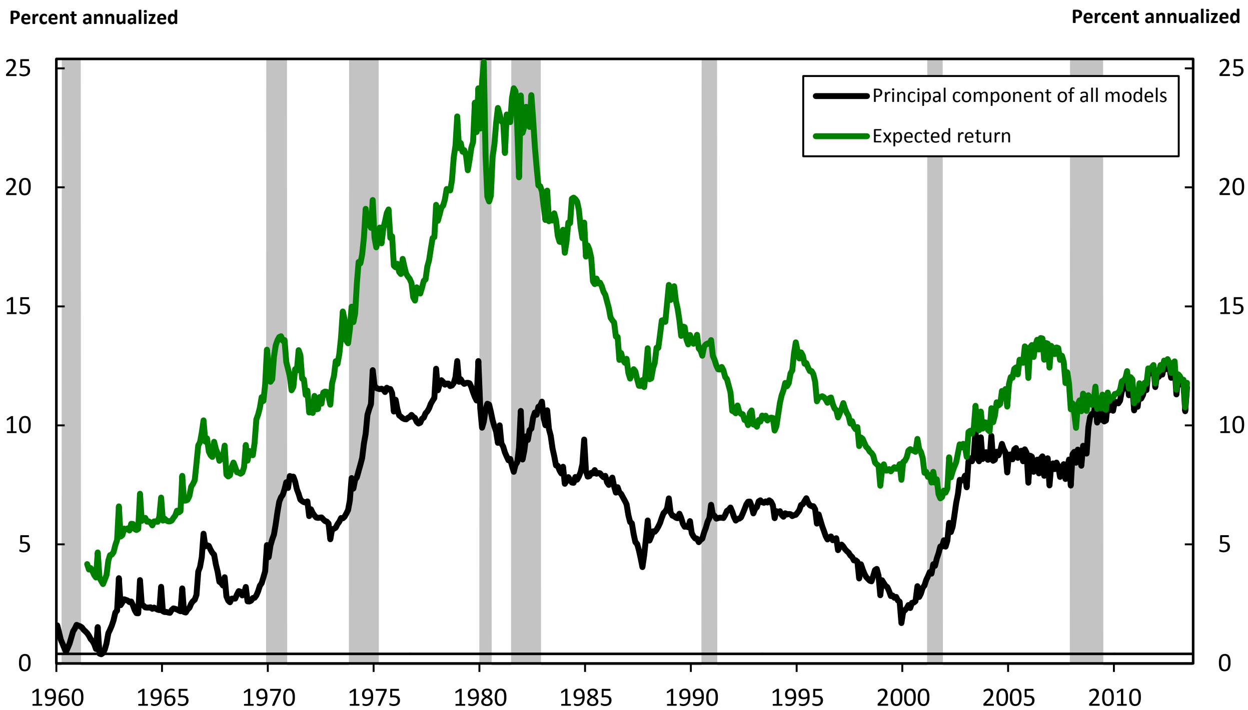
The figure shows monthly observations and the corresponding OLS regression for of the one-month-ahead ERP plotted against the slope of the ERP term structure for the period January 1960 to June 2013. The slope of the ERP term structure is the difference between the three-year-ahead ERP and the one-month-ahead ERP. All units are in annualized percentage points. The one-month-ahead and three-year-ahead ERP estimates used are the first principal components of twenty one-month-ahead or three-year-ahead ERP estimates from models described in Tables II-VI. The OLS regression slope is -1.17 (significant at the 99 percent level) and the R^2 is 50.1 percent.

Figure 6: Earnings behavior



The black line shows the monthly growth rate of real S&P 500 earnings, annualized and in percentage points. The green line shows the year-on-year change in the mean expectation of one-year-ahead earnings per share for the S&P 500 from a survey of analysts provided by Thomson Reuters I/B/E/S.

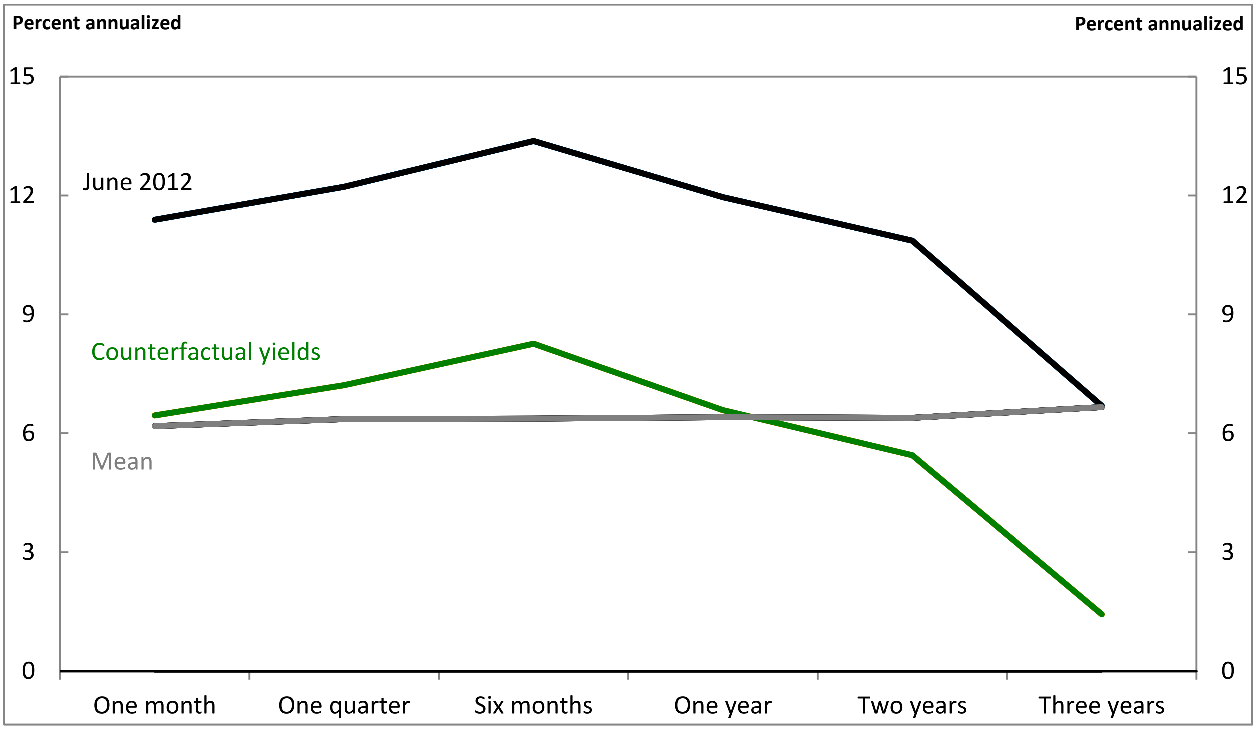
Figure 7: One-year-ahead ERP and expected returns



The black line is the first principal component of twenty models of the one-year-ahead equity risk premium (also shown in Figures 1, 2 and 3). The green line is the one-year-ahead expected return on the S&P 500, obtained by adding the realized one-year maturity Treasury yield from the principal component (the black line).

Shaded bars are NBER recessions.

Figure 8: Term structure of ERP using counterfactual bond yields



The grey line, labeled "Mean", shows the mean term structure of the equity risk premium over the sample January 1960 to June 2013. The black line, labeled "June 2012", shows the term structure for the most recent peak in the one-month-ahead ERP. These two lines are the same as in Figure 4. The green line, labeled "Counterfactual yields", shows what the term structure of equity risk premia would be in June 2012 if instead of subtracting June 2012's yield curve from expected returns we subtracted the average yield curve for January 1960 to June 2013.

Federal Reserve Bank of New York
Staff Reports

The Equity Risk Premium: A Review of Models

Fernando Duarte
Carlo Rosa

Staff Report No. 714
February 2015

The Equity Risk Premium: A Review of Models
Fernando Duarte and Carlo Rosa
Federal Reserve Bank of New York Staff Reports, no. 714
February 2015
JEL classification: C58, G00, G12, G17

Abstract

We estimate the equity risk premium (ERP) by combining information from twenty models. The ERP in 2012 and 2013 reached heightened levels—of around 12 percent—not seen since the 1970s. We conclude that the high ERP was caused by unusually low Treasury yields.

Key words: equity premium, stock returns

Duarte, Rosa: Federal Reserve Bank of New York (e-mail: fernando.duarte@ny.frb.org, carlo.rosa@ny.frb.org). The authors thank Tobias Adrian and James Egelhof for helpful comments on earlier drafts. This article is an update of, and a more comprehensive and rigorous treatment than, our blog post in Liberty Street Economics (http://libertystreeteconomics.newyorkfed.org/2013/05/are-stocks-cheap-a-review-of-the-evidence.html).

1. Introduction

The equity risk premium —the expected return on stocks in excess of the risk-free rate— is a fundamental quantity in all of asset pricing, both for theoretical and practical reasons. It is a key measure of aggregate risk-aversion and an important determinant of the cost of capital for corporations, savings decisions of individuals and budgeting plans for governments. Recently, the equity risk premium (ERP) has also returned to the forefront as a leading indicator of the evolution of the economy, a potential explanation for jobless recoveries and a gauge of financial stability[3].

In this article, we estimate the ERP by combining information from twenty prominent models used by practitioners and featured in the academic literature. Our main finding is that the ERP has reached heightened levels. The first principal component of all models –a linear combination that explains as much of the variance of the underlying data as possible– places the one-year-ahead ERP in June 2012 at 12.2 percent, above the 10.5 percent that was reached during the financial crisis in 2009 and at levels similar to those in the mid and late 1970s. Since June 2012 and until the end of our sample in June 2013, the ERP has remained little changed, despite substantial positive realized returns. It is worth keeping in mind, however, that there is considerable uncertainty around these estimates. In fact, the issue of whether stock returns are predictable is still an active area of research.[4] Nevertheless, we find that the dispersion in estimates across models, while quite large, has been shrinking, potentially signaling increased agreement

[3] As an indicator of future activity, a high ERP at short horizons tends to be followed by higher GDP growth, higher inflation and lower unemployment. See, for example, Piazzesi and Schneider (2007), Stock and Watson (2003), and Damodaran (2012). Bloom (2009) and Duarte, Kogan and Livdan (2013) study connections between the ERP and real aggregate investment. As a potential explanation of the jobless recovery, Hall (2014) and Kuehn, Petrosky-Nadeau and Zhang (2012) propose that increased risk-aversion has prevented firms from hiring as much as would be expected in the post-crisis macroeconomic environment. Among many others, Adrian, Covitz and Liang (2013) analyze the role of equity and other asset prices in monitoring financial stability.

[4] A few important references among a vast literature are Ang and Bekaert (2007), Goyal and Welch (2008), Campbell and Thompson (2008), Kelly and Pruitt (2013), Chen, Da and Zhao (2013), Neely, Rapach, Tu and Zhou (2014).

even when the models are substantially different from each other and use more than one hundred different economic variables.

In addition to estimating the level of the ERP, we investigate the reasons behind its recent behavior. Because the ERP is the difference between expected stock returns and the risk-free rate, a high estimate can be due to expected stock returns being high or risk-free rates being low. We conclude the ERP is high because Treasury yields are unusually low. Current and expected future dividend and earnings growth play a smaller role. In fact, expected stock returns are close to their long-run mean. One implication of a bond-yield-driven ERP is that traditional indicators of the ERP like the price-dividend or price-earnings ratios, which do not use data from the term structure of risk-free rates, may not be as good a guide to future excess returns as they have been in the past.

As a second contribution, we present a concise and coherent taxonomy of ERP models. We categorize the twenty models into five groups: predictors that use historical mean returns only, dividend-discount models, cross-sectional regressions, time-series regressions and surveys. We explain the methodological and practical differences among these classes of models, including the assumptions and data sources that each require.

2. The Equity Risk Premium: Definition

Conceptually, the ERP is the compensation investors require to make them indifferent at the margin between holding the risky market portfolio and a risk-free bond. Because this compensation depends on the future performance of stocks, the ERP incorporates expectations of future stock market returns, which are not directly observable. At the end of the day, any model of the ERP is a model of investor expectations. One challenge in estimating the ERP is that it is not clear what truly constitutes the market return and the risk-free rate in the real world. In practice, the most common measures of total market returns are based on broad stock market indices, such as the S&P 500 or the Dow Jones Industrial

Average, but those indices do not include the whole universe of traded stocks and miss several other components of wealth such as housing, private equity and non-tradable human capital. Even if we restricted ourselves to all traded stocks, we still have several choices to make, such as whether to use value or equal-weighted indices, and whether to exclude penny or infrequently traded stocks. A similar problem arises with the risk-free rate. While we almost always use Treasury yields as measures of risk-free rates, they are not completely riskless since nominal Treasuries are exposed to inflation[5] and liquidity risks even if we were to assume there is no prospect of outright default. In this paper, we want to focus on how expectations are estimated in different models, and not on measurement issues regarding market returns and the risk-free rate. Thus, we follow common practice and always use the S&P 500 as a measure of stock market prices and either nominal or real Treasury yields as risk-free rates so that our models are comparable with each other and with most of the literature.

While implementing the concept of the ERP in practice has its challenges, we can precisely define the ERP mathematically. First, we decompose stock returns[6] into an expected component and a random component:

$$R_{t+k} = E_t[R_{t+k}] + error_{t+k}. \tag{1}$$

In equation (1), R_{t+k} are *realized* returns between *t* and *t*+*k*, and $E_t[R_{t+k}]$ are the returns that were expected from *t* to $t+k$ using information available at time t. The variable $error_{t+k}$ is a random variable that is unknown at time t and realized at $t+k$. Under rational expectations, $error_{t+k}$ has a mean of zero and is orthogonal to $E_t[R_{t+k}]$. We keep the discussion as general as possible and do not assume rational

[5] Note that inflation risk in an otherwise risk-free nominal asset does not invalidate its usefulness to compute the ERP. If stock returns and the risk-free rate are expressed in nominal terms, their difference has little or no inflation risk. This follows from the following formula, which holds exactly in continuous time and to a first order approximation in discrete time: real stock returns – real risk-free rate = (nominal stock returns – expected inflation) – (nominal risk-free rate – expected inflation) = nominal stock returns– nominal risk-free rate. Hence, there is no distinction between a nominal and a real ERP.

[6] Throughout this article, all returns are *net* returns. For example, a five percent return corresponds to a net return of 0.05 as opposed to a *gross* return of 1.05.

expectations at this stage, although it will be a feature of many of the models we consider. The ERP at time t for horizon k is defined as

$$ERP_t(k) = E_t[R_{t+k}] - R_{t+k}^f, \tag{2}$$

where R_{t+k}^f is the risk-free rate for investing from t to $t+k$ (which, being risk-free, is known at time t).

This definition shows three important aspects of the ERP. First, future expected returns and the future ERP are stochastic, since expectations depend on the arrival of new information that has a random component not known in advance[7]. Second, the ERP has an investment horizon k embedded in it, since we can consider expected excess returns over, say, one month, one year or five years from today. If we fix t, and let k vary, we trace the *term structure* of the equity risk premium. Third, if expectations are rational, because the unexpected component $error_{t+k}$ is stochastic and orthogonal to expected returns, the ERP is always less volatile than realized excess returns. In this case, we expect ERP estimates to be smoother than realized excess returns.

3. Models of the Equity Risk Premium

We describe twenty models of the equity risk premium, comparing their advantages, disadvantages and ease of implementation. Of course, there are many more models of the ERP than the ones we consider. We selected the models in our study based on the recent academic literature, their widespread use by practitioners and data availability. Table I describes the data we use and their sources, all of which are either readily available or standard in the literature[8]. With a few exceptions, all data is monthly from January 1960 to June 2013. Appendix A provides more details.

[Insert Table I here]

[7] More precisely, $E_t[R_{t+k}]$ and $ERP_t(k)$ are known at time t but random from the perspective of all earlier periods.

[8] In fact, except for data from I/B/E/S and Compustat, all sources are public.

We classify the twenty models into five categories based on their underlying assumptions; models in the same category tend to give similar estimates for the ERP. The five categories are: models based on the historical mean of realized returns, dividend discount models, cross-sectional regressions, time-series regressions and surveys.

All but one of the estimates of the ERP are constructed in real time, so that an investor who lived through the sample would have been able to construct the measures at each point in time using available information only[9]. This helps minimize look-ahead bias and makes any out-of-sample evaluation of the models more meaningful. Clearly, most of the models themselves were designed only recently and were not available to investors in real time, potentially introducing another source of forward-looking and selection biases that are much more difficult to quantify and eliminate.

3.1 Historical mean of realized returns

The easiest approach to estimating the ERP is to use the historical mean of realized market returns in excess of the contemporaneous risk-free rate. This model is very simple and, as shown in Goyal and Welch (2008), quite difficult to improve upon when considering out-of-sample predictability performance measures. The main drawbacks are that it is purely backward looking and assumes that the future will behave like the past, i.e. it assumes the mean of excess returns is either constant or very slow moving over time, giving very little time-variation in the ERP. The main choice is how far back into the past we should go when computing the historical mean. Table II shows the two versions of historical mean models that we use.

[Insert Table II here]

[9] The one exception is Adrian, Crump and Moench’s (2014) cross-sectional model, which is constructed using full-sample regression estimates.

3.2 Dividend discount models (DDM)

All DDM start with the basic intuition that the value of a stock is determined by no more and no less than the cash flows it produces for its shareholders, as in Gordon (1962). Today's stock price should then be the sum of all expected future cash flows, discounted at an appropriate rate to take into account their riskiness and the time value of money. The formula that reflects this intuition is

$$P_t = \frac{D_t}{\rho_t} + \frac{E_t[D_{t+1}]}{\rho_{t+1}} + \frac{E_t[D_{t+2}]}{\rho_{t+2}} + \frac{E_t[D_{t+3}]}{\rho_{t+3}} + \cdots, \tag{3}$$

where P_t is the current price of the stock, D_t are current cash flows, $E_t[D_{t+k}]$ are the cash flows k periods from now expected as of time t, and ρ_{t+k} is the discount rate for time $t+k$ from the perspective of time t. Cash flows to stockholders certainly include dividends, but can also arise from spin-offs, buy-outs, mergers, buy-backs, etc. In general, the literature focuses on dividend distributions because they are readily available data-wise and account for the vast majority of cash flows. The discount rate can be decomposed into

$$\rho_{t+k} = 1 + R^f_{t+k} + ERP_t(k). \tag{4}$$

In this framework, the risk-free rate captures the discounting associated with the time value of money and the ERP captures the discounting associated with the riskiness of dividends. When using a DDM, we refer to $ERP_t(k)$ as the *implied* ERP. The reason is that we plug in prices, risk-free rates and estimated expected future dividends into equation (3), and then derive what value of $ERP_t(k)$ makes the right-hand side equal to the left-hand side in the equation, i.e. what ERP value is *implied* by equation (3).

DDM are forward looking and are consistent with no arbitrage. In fact, equation (3) must hold in any economy with no arbitrage[10]. Another advantage of DDM is that they are easy to implement. A drawback of DDM is that the results are sensitive to how we compute expectations of future dividends. Table III displays the DDM we consider and a brief description of their different assumptions.

[Insert Table III here]

3.3 Cross-sectional regressions

This method exploits the variation in returns and exposures to the S&P 500 of different assets to infer the ERP[11]. Intuitively, cross-sectional regressions find the ERP by answering the following question: what is the level of the ERP that makes expected returns on a variety of stocks consistent with their exposure to the S&P 500? Because we need to explain the relationship between returns and exposures for multiple stocks with a single value for the ERP (and perhaps a small number of other variables), this model imposes tight restrictions on estimates of the ERP.

The first step is to find the exposures of assets to the S&P 500 by estimating an equation of the following form:

$$R^i_{t+k} - R^f_{t+k} = \alpha^i \times state\ variables_{t+k} + \beta^i \times risk\ factors_{t+k} + idiosyncratic\ risk^i_{t+k}. \quad (5)$$

In equation (5), R^i_{t+k} is the realized return on a stock or portfolio i from time t to $t+k$. $State\ variables_{t+k}$ are any economic indicators that help identify the state of the economy and its likely future path. $Risk\ factors_{t+k}$ are any measures of systematic contemporaneous co-variation in returns across all stocks or portfolios. Of course, some economic indicators can be both state variables and risk

[10] Note that when performing the infinite summation in equation (3) we have not assumed the n^{th} term goes to zero as n tends to infinity, which allows for rational bubbles. In this sense, DDM do allow for a specific kind of bubble.

[11] See Polk, Thompson and Vuolteenaho (2006) and Adrian, Crump and Moench (2014) for a detailed description of this method.

factors at the same time. Finally, $idiosyncratic\ risk_{t+k}^{i}$ is the component of returns that is particular to each individual stock or portfolio that is not explained by $state\ variables_{t+k}$ or $risk\ factors_{t+k}$ (both of which, importantly, are common to all stocks and hence not indexed by i). Examples of state variables are inflation, unemployment, the yield spread between Aaa and Baa bonds, the yield spread between short and long term Treasuries, and the S&P 500's dividend-to-price ratio. The most important risk factor is the excess return on the S&P 500, which we must include if we want to infer the ERP consistent with the cross-section of stock returns. Other risk-factors usually used are the Fama-French (1992) factors and the momentum factor of Carhart (1997). The values in the vector α^{i} give the strength of asset-specific return predictability and the values in the vector β^{i} give the asset-specific exposures to risk factors[12]. For the cross-section of assets indexed by i, we can use the whole universe of traded stocks, a subset of them, or portfolios of stocks grouped, for example, by industry, size, book-to-market, or recent performance. It is important to point out that equation (5) is not a predictive regression; the left and right-hand side variables are both associated with time $t + k$.

The second step is to find the ERP associated with the S&P 500 by estimating the cross-sectional equations

$$R_{t+k}^{i} - R_{t+k}^{f} = \lambda_t(k) \times \hat{\beta}^{i} \ , \tag{6}$$

where $\hat{\beta}^{i}$ are the values found when estimating equation (5). Equation (6) attempts to find, at each point in time, the vector of numbers $\lambda_t(k)$ that makes exposures β^{i} as consistent as possible with realized excess returns of all stocks or portfolios considered. The element in the vector $\hat{\lambda}_t(k)$ that is multiplied by

[12] The vectors α^{i} and β^{i} could also be time-varying, reflecting a more dynamic relation between returns and their explanatory variables. In this case, the estimation of equation (5) is more complicated and requires making further assumptions. The model by Adrian, Crump and Moench (2014) is the only cross-sectional model we examine that uses time-varying α^{i} and β^{i} .

the element in the $\hat{\beta}^i$ vector corresponding to the S&P 500 is $ERP_t(k)$, the equity risk premium we are seeking.

One advantage of cross-sectional regressions is that they use information from more asset prices than other models. Cross-sectional regressions also have sound theoretical foundations, since they provide one way to implement Merton's (1973) Intertemporal Capital Asset Pricing Model. Finally, this method nests many of the other models considered. The two main drawbacks of this method are that results are dependent on what portfolios, state variables and risk factors are used (Harvey, Liu and Zhu (2014)), and that it is not as easy to implement as most of the other options. Table IV displays the cross-sectional models in our study, together with the state variables and risk factors they use.

[Insert Table IV here]

3.4 Time-series regressions

Time-series regressions use the relationship between economic variables and stock returns to estimate the ERP. The idea is to run a predictive linear regression of realized excess returns on lagged "fundamentals":

$$R_{t+k} - R^f_{t+k} = a + b \times Fundamental_t + error_t. \qquad (7)$$

Once estimates $\hat{a}$ and $\hat{b}$ for a and b are obtained, the ERP is obtained by ignoring the error term:

$$ERP_t(k) = \hat{a} + \hat{b} \times Fundamental_t. \qquad (8)$$

In other words, we estimate only the forecastable or expected component of excess returns. This method attempts to implement equations (1) and (2) as directly as possible in equations (7) and (8), with the assumption that "fundamentals" are the right sources of information to look at when computing expected returns, and that a linear equation is the correct functional specification.

The use of time-series regressions requires minimal assumptions; there is no concept of equilibrium and no absence of arbitrage necessary for the method to be valid[13]. In addition, implementation is quite simple, since it only involves running ordinary least-square regressions. The challenge is to select what variables to include on the right-hand side of equation (7), since results can change substantially depending on what variables are used to take the role of "fundamentals". In addition, including more than one predictor gives poor out-of-sample predictions even if economic theory may suggest a role for many variables to be used simultaneously (Goyal and Welch (2008)). Finally, time-series regressions ignore information in the cross-section of stock returns. Table V shows the time-series regression models that we study.

[Insert Table V here]

3.5 Surveys

The survey approach consists of asking economic agents about the current level of the ERP. Surveys incorporate the views of many people, some of which are very sophisticated and/or make real investment decisions based on the level of the ERP. Surveys should also be good predictors of excess returns because in principle stock prices are determined by supply and demand of investors such as the ones taking the surveys. On the other hand, Greenwood and Shleifer (2014) document that investor expectations of future stock market returns are positively correlated with past stock returns and with the current level of the stock market, but strongly *negatively* correlated with model-based expected returns and future realized stock market returns. Other studies such as Easton and Sommers (2007) also argue that survey measures of the ERP can be systematically biased. In this paper, we use the survey of CFOs by Graham and Harvey (2012), which to our knowledge is the only large-scale ERP survey that has more than just a few years of data (see Table VI).

[Insert Table VI here]

[13] However, the Arbitrage Pricing Theory of Ross (1976) provides a strong theoretical underpinning for time-series regressions by using no-arbitrage conditions.

4. Estimation of the Equity Risk Premium

We now study the behavior of the twenty models we consider by conducting principal component analysis. Since forecast accuracy can be substantially improved through the combination of multiple forecasts[14], the optimal strategy to forecast excess stock returns may consist of combining together all these models. The first principal component of the twenty models that we use is the linear combination of ERP estimates that captures as much of the variation in the data as possible. The second, third, and successive principal components are the linear combinations of the twenty models that explain as much of the variation of the data as possible and are also uncorrelated to all the preceding principal components. If the first few principal components —say one or two— account for most of the variation of the data, then we can use them as a good summary for the variation in all the measures over time, reducing the dimensionality from twenty to one or two. In addition, in the presence of classical measurement error, the first few principal components can achieve a higher signal-to-noise ratio than other summary measures like the cross-sectional mean of all models (Geiger and Kubin (2013)).

To compute the first principal component, we proceed in three steps. We first de-mean all ERP estimates and find their variance-covariance matrix. In the second step, we find the linear combination that explains as much of the variance of the de-meaned models as possible. The weights in the linear combination are the elements of the eigenvector associated with the largest eigenvalue of the variance-covariance matrix found in the first step. In the third step, we add to the linear combination just obtained, which has mean zero, the average of ERP estimates across all models and all time periods. Under the assumption that each of the models is an unbiased and consistent estimator of the ERP, the average across all models and all time periods is an unbiased and consistent estimator of the unconditional mean of the ERP. The time

[14] See, *inter alia*, Clemen (1989), Diebold and Lopez (1996) and Timmermann (2006).

variation in the first principal component then provides an estimate of the conditional ERP[15]. The share of the variance of the underlying models explained by this principal component is 76 percent, suggesting that there is not too much to gain from examining principal components beyond the first[16].

We now focus on the one-year-ahead ERP estimates and study other horizons in the next section.

The first two columns in Table VII show the mean and standard deviation of each model's estimates. The unconditional mean of the ERP across all models is 5.7 percent, with an average standard deviation of 3.2 percent. DDM give the lowest mean ERP estimates and have moderate standard deviations. In contrast, cross-sectional models tend to have mean ERP estimates on the high end of the distribution and very smooth time-series. Mean ERP estimates for time-series regressions are mixed, with high and low values depending on the predictors used, but uniformly large variances. The survey of CFOs has a mean and standard deviation that are both about half as large as in the overall population of models. The picture that emerges from Table VII is that there is considerable heterogeneity across model types, and even sometimes within model types, thereby underscoring the difficulty inherent in finding precise estimates of the ERP.

---

[15] As is customary in the literature, we perform the analysis using ERP estimates in levels, even though they are quite persistent. Results in first-differences do not give economically reasonable estimates since they feature a pro-cyclical ERP and unreasonable magnitudes.

One challenge that arises in computing the principal component is when we have missing observations, either because some models can only be obtained at frequencies lower than monthly or because the necessary data is not available for all time periods (Appendix A contains a detailed description of when this happens). To overcome this challenge, we use an iterative linear projection method, which conceptually preserves the idea behind principal components. Let X be the matrix that has observations for different models in its columns and for different time periods in its rows. On the first iteration, we make a guess for the principal component and regress the non-missing elements of each row of X on the guess and a constant. We then find the first principal component of the variance-covariance matrix of the fitted values of these regressions, and use it as the guess for the next iteration. The process ends when the norm of the difference between consecutive estimates is small enough. We thank Richard Crump for suggesting this method and providing the code for its implementation.

[16] The second and third principal components account for 13 and 8 percent of the variance, respectively.

[Insert Table VII here]

Figure 1 shows the time-series for all one-year-ahead ERP model estimates, with each class of models in a different panel. The green lines are the ERP estimates from the twenty underlying models. The black line, reproduced in each of the panels, is the principal component of all twenty models. The shaded areas are NBER recessions. The figure gives a sense of how the time-series move together, and how much they co-vary with the first principal component. Table VIII shows the correlations among models. Figure 1 and Table VIII give the same message: despite some outliers, there is a fairly strong correlation within each of the five classes of models. Across classes, however, correlations are small and even negative. Interestingly, the correlation between some DDM and cross-sectional models is as low as -91 percent. This negative correlation, however, disappears if we look at lower frequencies. When aggregated to quarterly frequency, the smallest correlation between DDM and cross-sectional models is -22 percent, while at the annual frequency it is 12 percent.

[Insert Figure 1 here]

[Insert Table VIII here]

Figure 1 also shows that the first principal component co-varies negatively with historical mean models, but positively with DDM and cross-sectional regression models. Time-series regression models are also positively correlated with the first principal component, although this is not so clearly seen in Panel 4 of Figure 1 because of the high volatility of time-series ERP estimates. The last panel shows that the survey of CFOs does track the first principal component quite well at low frequencies (e.g. annual), although any conclusions about survey estimates should be interpreted with caution given the short length of the sample.

As explained earlier, the first principal component is a linear combination of the twenty underlying ERP models:

$$PC_t^{(1)} = \sum_{m=1}^{20} w^{(m)} ERP_t^{(m)}. \tag{9}$$

In the above equation, m indexes the different models, $PC_t^{(1)}$ is the first principal component, $ERP_t^{(m)}$ is the estimate from model m and $w^{(m)}$ is the weight that the principal component places on model m. The third column in Table VII, labeled "PC coefficients", shows the weights $w^{(m)}$ normalized to sum up to one to facilitate comparison, i.e. the table reports the weights $\widehat{w}^{(m)}$ where

$$\widehat{w}^{(m)} = \frac{w^{(m)}}{\sum_{m=1}^{20} w^{(m)}}. \tag{10}$$

The first principal component puts positive weight on models based on the historical mean, cross-sectional regressions and the survey of CFOs. It weights DDM and time-series regressions mostly negatively. The absolute values of the weights are very similar for many of the models, and there is no single model or class of models that dominates. This means that the first principal component uses information from many of the models.

The last column in Table VII, labeled "Exposure to PC", shows the extent to which models *load* on the first principal component. By construction, each of the twenty ERP models can be written as a linear combination of twenty principal components:

$$ERP_t^{(m)} = \sum_{i=1}^{20} load_i^{(m)} PC_t^{(i)}, \tag{11}$$

where m indexes the model and i indexes the principal components. The values in the last column of Table VII are the loadings on the first principal component ($i = 1$) for each model ($m = 1, 2, \ldots, 20$), again normalized to one for ease of comparability:

$$\widehat{load}_1^{(m)} = \frac{load_1^{(m)}}{\sum_{m=1}^{20} load_i^{(m)}}. \tag{12}$$

Most models have a positive loading on the first principal component; whenever the loading is negative, it tends to be relatively small. This means the first principal component, as expected, is a good explanatory variable for most models. Looking at the third and fourth columns of Table VII together, we can obtain additional information. For example, a model with a very high loading (fourth column) accompanied by a very small PC coefficient (third column) is likely to mean that the model is almost redundant, in the sense that it is close to being a linear combination of all other models and does not provide much independent information to the principal component. On the other hand, if the PC coefficient and loading are both high, the corresponding model is likely providing information not contained in other measures.

Figure 2 shows the first principal component of all twenty models in black, with recessions indicated by shaded bars (the black line is the same principal component shown in black in each of the panels of Figure 1). As expected, the principal component tends to peak during financial turmoil, recessions and periods of low real GDP growth or high inflation. It tends to bottom out after periods of sustained bullish stock markets and high real GDP growth. Evaluated by the first principal component, the one-year-ahead ERP reaches a local peak in June of 2012 at 12.2 percent. The surrounding months have ERP estimates of similar magnitude, with the most recent estimate in June 2013 at 11.2 percent. This behavior is not so clearly seen by simply looking at the collection of individual models in Figure 1, highlighting the usefulness of principal components analysis. Similarly high levels were seen in the mid and late 1970s, during a period of stagflation, while the recent financial crisis had slightly lower ERP estimates closer to 10 percent.

[Insert Figure 2 here]

Figure 2 also displays the 10th, 25th, 75th and 90th percentiles of the cross-sectional distribution of models. These bands can be interpreted as confidence intervals, since they give the range of the distribution of ERP estimates at each point in time. However, they do not incorporate other relevant sources of uncertainty, such as the errors that occur during the estimation of each individual model, the degree of doubt in the correctness of each model, and the correlation structure between these and all other kinds of errors. Standard error bands that capture all sources of uncertainty are therefore likely to be wider.

The difference in high and low percentiles can also be interpreted as measures of agreement across models. The interquartile range –the difference between the 25th and 75th percentiles— has compressed, mostly because the models in the bottom of the distribution have had higher ERP estimates since 2010. It is also interesting to note that the 75th percentile has remained fairly constant over the last 10 years at a level somewhat below its long-run mean. The cross-sectional standard deviation in ERP estimates (not shown in the graph) also decreased from 10.2% in January of 2000 to 4.3% in June of 2013, confirming that the disagreement among models has decreased.

Another *a priori* reasonable summary statistic for the ERP is the cross-sectional mean of estimates across models. In Figure 3, we can see that by this measure the ERP has also been increasing since the crisis. However, unlike the principal component, it has not reached elevated levels compared to past values. The cross-sectional mean can be useful, but it has a few undesirable features as an overall measure of the ERP compared to the first principal component. First, it is procyclical, which contradicts the economic intuition that expected returns are highest in recessions, when risk aversion is high and future prospects look brighter than current ones. Second, it overloads on DDM simply because there is a higher number of DDM models in our sample. Lastly, it has a smaller correlation with the realized returns it is supposed to predict.

[Insert Figure 3 here]

5. The Term Structure of Equity Risk Premia

In Section 2, we described the term structure of the ERP – what expected excess returns are over different investment horizons. In practical terms, we estimate the ERP at different horizons by using the inputs for all the models at the corresponding horizons[17]. For example, if we want to take the historical mean of returns as our estimate, we can take the mean of returns over one month, six months, or a one-year period. In cross-sectional and time-series regressions, we can predict monthly, quarterly or annual returns using monthly, quarterly or annual right-hand side variables. DDM, on the other hand, have little variation across horizons. In fact, all the DDM we consider have a constant term structure of expected stock returns, and the only term structure variation in ERP estimates comes from risk-free rates[18].

Figure 4 plots the first principal components of the ERP as a function of investment horizon for some selected dates. We picked the dates because they are typical dates for when the ERP was unusually high or unusually low at the one-month horizon. As was the case for one-year-ahead ERP estimates, we can capture the majority of the variance of the underlying models at all horizons by a single principal component. The shares of the variance explained by the first principal components at horizons of one month to three years range between 68 and 94 percent. The grey line in Figure 4 shows the average of the term structure across all periods. It is slightly upward sloping, with a short-term ERP at just over 6 percent and a three-year ERP at almost 7 percent.

[Insert Figure 4 here]

[17] For other ways to estimate the term structure of the ERP using equilibrium models or derivatives, see Ait-Sahalia, Karaman and Mancini (2014), Ang and Ulrich (2012), van Binsbergen, Hueskes, Koijen and Vrugt (2014), Boguth, Carlson, Fisher and Simutin (2012), Durham (2013), Croce, Lettau and Ludvigson (2014), Lemke and Werner (2009), Lettau and Wachter (2011), Muir (2013), among others.

[18] In equation (3), ρ_{t+k} is assumed to be the same for all k, while risk-free rates are allowed to vary over the investment horizon k in equation (4). Of course, with additional assumptions, it is possible to have DDM with a non-constant term structure of expected excess returns.

The first observation is that the term structure of the ERP has significant time variation and can be flat, upward or downward sloping. Figure 4 also shows some examples that hint at lower future expected excess returns when the one-month-ahead ERP is elevated and the term structure is downward sloping, and higher future expected excess returns when the one-month-ahead ERP is low and the term structure is upward sloping. In fact, this is generally true: There is a strong negative correlation between the level and the slope of the ERP term structure of -71 percent. Figure 5 plots monthly observations of the one-month-ahead ERP against the slope of the ERP term structure (the three-year-ahead minus the one-month-ahead ERP) together with the corresponding ordinary least squares regression line in black. Of course, this is only a statistical pattern and should not be interpreted as a causal relation.

[Insert Figure 5 here]

6. Why is the Equity Risk Premium High?

There are two reasons why the ERP can be high: low discount rates and high current or expected future cash flows.

Figure 6 shows that earnings are unlikely to be the reason why the ERP is high. The green line shows the year-on-year change in the mean expectation of one-year-ahead earnings per share for the S&P 500. These expectations are obtained from surveys conducted by the Institutional Brokers' Estimate System (I/B/E/S) and available from Thomson Reuters. Expected earnings per share have been declining from 2010 to 2013, making earnings growth an unlikely reason for why the ERP was high in the corresponding period. The black line shows the realized monthly growth rates of real earnings for the S&P 500 expressed in annualized percentage points. Since 2010, earnings growth has been declining, hovering around zero for the last few months of the sample. It currently stands at 2.5 percent, which is near its long-run average.

[Insert Figure 6 here]

Another way to examine whether a high ERP is due to discount rates or cash flows is shown in Figure 7. The black line is the same one-year-ahead ERP estimate shown in Figure 2. The green line simply adds the realized one-year Treasury yield to obtain expected stock returns. The figure shows expected stock returns have increased since 2000, similarly to the ERP. However, unlike the ERP, expected stock returns are close to their long-run mean, and nowhere near their highest levels, achieved in 1980. The discrepancies between the two lines are due to exceptionally low bond yields since the end of the financial crisis.

[Insert Figure 7 here]

Figure 8 displays the term structure of the ERP under a simple counterfactual scenario, in addition to the mean and current term structures already displayed in Figure 4. In this scenario, we leave expected stock returns unmodified but change the risk-free rates in June 2012 from their actual values to the average nominal bond yields over 1960-2013. In other words, we replace R^f_{t+k} in equation (2) by the mean of R^f_{t+k} over t. The result of this counterfactual is shown in Figure 8 in green. Using average levels of bond yields brings the whole term structure of the ERP much closer to its mean level (the grey line), especially at intermediate horizons. This shows that a “normalization” of bond yields, everything else being equal, would bring the ERP close to its historical norm. This exercise shows that the current environment of low bond yields is capable, quantitatively speaking, of significantly contributing to an ERP as high as was observed in 2012-2013.

[Insert Figure 8 here]

7. Conclusion

We have analyzed twenty different models of the ERP by considering the assumptions and data required to implement them, and how they relate to each other. When it comes to the ERP, we find that there is substantial heterogeneity in estimation methodology and final estimates. We then extract the first principal component of the twenty models, which signals that the ERP in 2012 and 2013 is at heightened levels compared to previous periods. Our analysis provides evidence that the current level of the ERP is consistent with a bond-driven ERP: expected excess stock returns are elevated not because stocks are expected to have high returns, but because bond yields are exceptionally low. The models we consider suggest that expected stock returns, on their own, are close to average levels.

References

Adrian, T., D. M. Covitz, and N. Liang. 2013. “Financial stability monitoring.” Federal Reserve Bank of New York Staff Reports, no. 601 (February).

Adrian, T., R. Crump, and E. Moench. 2014. “Regression-based estimation of dynamic asset pricing models.” Federal Reserve Bank of New York Staff Reports, no. 493 (June).

Ait-Sahalia, Y., M. Karaman, and L. Mancini. 2014. “The term structure of variance swaps, risk premia and the expectation hypothesis. Risk Premia and the Expectation Hypothesis.” (May 10, 2014). Available at SSRN: http://ssrn.com/abstract=2136820

Ang, A., and G. Bekaert. 2007. “Stock return predictability: Is it there?” Review of Financial Studies 20, no. 3 (May): 651-707.

Ang, A., and M. Ulrich. 2012. “Nominal Bonds, Real Bonds, and Equity.” Unpublished paper, Columbia University.

Baker, M., and J. Wurgler. 2007. “Investor sentiment in the stock market.” The Journal of Economic Perspectives: 129-51.

van Binsbergen, J., W. Hueskes, R. Koijen, and E. Vrugt. 2013. “Equity yields.” Journal of Financial Economics 110, no. 3 (December): 503-19.

Bloom, N. 2009. “The impact of uncertainty shocks.” Econometrica 77, no. 3 (May): 623-85.

Boguth, O., M. Carlson, A. J. Fisher, and M. Simutin. 2012. “Leverage and the limits of arbitrage pricing: Implications for dividend strips and the term structure of equity risk premia.” Unpublished paper, Arizona State University.

Campbell, J.Y., and S. Thompson. 2008. “Predicting excess stock returns out of sample: Can anything beat the historical average?” Review of Financial Studies 21, no. 4 (July): 1509-31.

Carhart, M. M. 1997. “On persistence in mutual fund performance.” Journal of Finance 52, no 1 (March): 57–82.

Chen, L., Z. Da, and X. Zhao. 2013. “What drives stock price movements?” Review of Financial Studies 26, no. 4 (April): 841-76.

Clemen, R.T., 1989. “Combining forecasts: a review and annotated bibliography”. International Journal of Forecasting 5 (4), 559-83.

Croce, M., M. Lettau, and S. Ludvigson. Forthcoming. 2014. “Investor information, long-run risk, and the term structure of equity.” Review of Financial Studies.

Damodaran, A. 2012. “Equity Risk Premiums (ERP): Determinants, Estimation and Implications – The 2012 Edition.” Unpublished paper, New York University.

Diebold, F.X., and J. Lopez, 1996. “Forecast Evaluation and Combination”. In G.S. Maddala and C.R. Rao (eds.), Handbook of Statistics, 241-268, 1996. Amsterdam: North-Holland.

Duarte, F. 2013. “Inflation and the cross-section of stock returns.” Federal Reserve Bank of New York Staff Reports, no. 621 (May).

Durham, J. B. 2013. “Arbitrage-free models of stocks and bonds.” Federal Reserve Bank of New York Staff Reports, no. 656 (December).

Duarte, F., L. Kogan, and D. Livdan. 2013. “Aggregate Investment and Stock Returns.” Unpublished paper, M.I.T.

Easton, P.D., and G. A. Sommers. 2007. “Effect of analysts’ optimism on estimates of the expected return implied by earnings forecasts.” Journal of Accounting Research 45, no. 5 (December): 983-1016.

Fama, E.F., and K. R. French. 1988. “Dividend yields and expected stock returns.” Journal of Financial Economics 22, no. 1 (October), 3-25.

Fama, E.F., and K. R. French. 1992. “The cross-section of expected stock returns.” Journal of Finance 47, no. 2 (June): 427-65.

Fama, E.F., and K. R. French. 2002. “The equity premium.” Journal of Finance 57, no. 2 (April): 637-59.

Geiger, B., and G. Kubin. 2013. "Signal Enhancement as Minimization of Relevant Information Loss". Proceedings of ITG Conference on Systems, Communication and Coding

Gordon, M. J. 1962. “The investment, financing, and valuation of the corporation.” Greenwood Press.

Goyal, A., and I. Welch. 2008. “A comprehensive look at the empirical performance of equity premium predictions.” Review of Financial Studies 21, no. 4 (July): 1455-1508.

Graham, J., and C. Harvey. 2012. “The Equity Risk Premium in 2012.” Unpublished paper, Duke University.

Greenwood, R., and A. Shleifer. 2014. “Expectations of returns and expected returns.” Review of Financial Studies 27, no. 3 (March): 714-46

Gurkaynak, R. S., B. Sack, and J. H. Wright. 2007. “The US Treasury yield curve: 1961 to the present.” Journal of Monetary Economics 54, no. 8 (November): 2291-2304.

Gurkaynak, R. S., S. Refet, B. Sack, and J. H. Wright. 2010. “The TIPS yield curve and inflation compensation.” American Economic Journal: Macroeconomics 2, no. 1 (January): 70-92.

Hall, R. E. Forthcoming. 2014. “The Routes into and out of the Zero Lower Bound.” Federal Reserve Bank of Kansas City Symposium (Jackson Hole).

Harvey, C. R., and Y. Liu, and H. Zhu. 2014. “…and the Cross-Section of Expected Returns.” (October 7, 2014). Available at SSRN: http://ssrn.com/abstract=2249314

Kelly, B., and S. Pruitt. 2013. “Market Expectations in the Cross-Section of Present Values.” The Journal of Finance 68, no. 5 (October): 1721-56.

Kuehn, L. A., N. Petrosky-Nadeau, and L. Zhang. 2012. “An equilibrium asset pricing model with labor market search.” NBER Working Paper no. 17742.

Lemke, W., and T. Werner. 2009. “The term structure of equity premia in an affine arbitrage-free model of bond and stock market dynamics.” European Central Bank, Working Paper No.1045.

Lettau, M., and J. A. Wachter. 2011. “The term structures of equity and interest rates.” Journal of Financial Economics 101, no. 1 (July): 90-113.

Merton, R.C. 1973. “An intertemporal capital asset pricing model.” Econometrica 41, no. 5 (September): 867-87.

Muir, T. 2013. “Financial crises, risk premia, and the term structure of risky assets.” Unpublished paper, Yale School of Management.

Neely, C. J., D.E. Rapach, J. Tu, and G. Zhou. 2014. “Forecasting the equity risk premium: the role of technical indicators.” Management Science 60, no. 7 (July): 1772-91.

Panigirtzoglou, N., and J. Loeys. 2005. “A fair value model for US bonds, credit and equities.” JP Morgan (January).

Piazzesi, M., and M. Schneider. 2007. “Equilibrium yield curves.” In *NBER Macroeconomics Annual 2006* 21: 389-472. MIT Press, 2007.

Polk, C., S. Thompson, and T. Vuolteenaho. 2006. “Cross-sectional forecasts of the equity premium.” Journal of Financial Economics 81, no. 1 (July): 101-41.

Ross, S. 1976. “The arbitrage theory of capital asset pricing.” Journal of Economic Theory 13, no. 3 (December): 341–60.

Shiller, R. J. 2005. “Irrational Exuberance”. Princeton University Press 2000, Broadway Books 2001, 2nd ed.

Stock, J.H., and M. W. Watson. 2003. “Forecasting output and inflation: The role of asset prices.” Journal of Economic Literature 41, no. 3 (September): 788-829.

Timmermann, A. 2006. “Forecast combinations”. Handbook of Economic Forecasting, Volume 1 (chapter 4), 135-96.

Appendix A: Data Variables

Fama and French (1992)	http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html Monthly frequency; 1/1/1960 to 6/30/2013. We use 25 portfolios sorted on size and book to market, 10 portfolios sorted on momentum, realized excess market returns, HML, SMB, and the momentum factor.
Shiller (2005)	http://www.econ.yale.edu/~shiller/data.htm Monthly frequency; 1/1/1960 to 6/30/2013. We use the nominal and real price, nominal and real dividends and nominal and real earnings for the S&P 500, CPI, and 10 year nominal treasury yield.
Baker and Wurgler (2007)	http://people.stern.nyu.edu/jwurgler/data/Investor_Sentiment_Data_v23_POST.xlsx Monthly frequency; 7/1/1965 to 12/1/2010. We use the "sentiment measure".
Graham and Harvey (2012)	http://www.cfosurvey.org/index.htm Quarterly frequency; 6/6/2000 to 6/5/2013. We use the answer to the question "Over the next 10 years, I expect the average annual S&P 500 return will be: Expected return:" and the analogous one that asks about the next year.
Damodaran (2012)	http://www.stern.nyu.edu/~adamodar/pc/datasets/histimpl.xls Annual frequency; 1/1/1960 to 12/1/2012. We use the ERP estimates from his dividend discount models (one uses free-cash flow, the other one doesn't).
Gurkaynak, Sack and Wright (2007)	http://www.federalreserve.gov/pubs/feds/2006/200628/200628abs.html Daily frequency; starting on 6/14/61 for one- to seven-year yields; 8/16/71 for nine- and ten-year yields; 11/15/71 for eleven- to fifteen-year yields; 7/2/81 for sixteen- to twenty-year yields; 11/25/85 for twenty-one- to thirty-year yields. We use all series until 6/30/2013.
Gurkaynak, Refet, Sack and Wright (2010)	http://www.federalreserve.gov/econresdata/researchdata.htm Monthly frequency; 1/1/2003 to 7/1/2013. We use yields on TIPS of all maturities available.
Compustat	Variable BKVLPS Annual frequency; 12/31/1977 to 12/31/2012.
Thomson Reuters I/B/E/S	Variables EPS 1 2 3 4 5 Monthly frequency; 1/14/1982 to 4/18/2013 for current and next year forecasts; 9/20/84 to 4/18/2013 for two-year-ahead forecasts; 9/19/85 to 3/15/2012 for three-year-ahead forecasts; 2/18/88 to 3/15/07 for four-year-ahead forecasts.
FRED (St. Louis Federal Reserve)	http://research.stlouisfed.org/fred2/graph/?g=D9J and http://research.stlouisfed.org/fred2/graph/?g=KKk Monthly frequency. 1/1/1960 to 7/1/2013 for Baa minus Aaa bond yield spread and recession indicator.

Tables and Figures

Table I: Data sources	
Fama and French (1992)	Fama-French factors, momentum factor, twenty-five portfolios sorted on size and book-to-market
Shiller (2005)	Inflation and ten-year nominal treasury yield. Nominal price, real price, earnings, dividends and cyclically adjusted price-earnings ratio for the S&P 500
Baker and Wurgler (2007)	Debt issuance, equity issuance, sentiment measure
Graham and Harvey (2012)	ERP estimates from the Duke CFO survey
Damodaran (2012)	ERP estimates
Gurkaynak, Sack and Wright (2007)	Zero coupon nominal bond yields for all maturities[19]
Gurkaynak, Refet, Sack and Wright (2010)	Zero coupon TIPS yields for all maturities
Compustat	Book value per share for the S&P 500
Thomson Reuters I/B/E/S	Mean analyst forecast of expected earnings per share
FRED (St. Louis Federal Reserve)	Corporate bond Baa-Aaa spread and the NBER recession indicator

Note: All variables start in January 1960 (or later, if unavailable for early periods) and end in June 2013 (or until no longer available). CFO surveys are quarterly; book value per share and ERP estimates by Damodaran (2012) are annual; all other variables are monthly. Appendix A provides more details.

[19] Except for the 10-year yield, which is from Shiller (2005). We use the 10-year yield from Shiller (2005) for ease of comparability with the existing literature. Results are virtually unchanged if we use all yields, including the 10-year yield, from Gurkaynak, Sack and Wright (2007).

Table II: Models based on the historical mean of realized returns	
Long-run mean	Average of realized S&P 500 returns minus the risk-free rate using all available historical data
Mean of the previous five years	Average of realized S&P 500 returns minus the risk-free rate using only data for the previous five years

Table III: Dividend Discount Models	
Gordon (1962) with nominal yields	S&P 500 dividend-to-price ratio minus the ten-year nominal Treasury yield
Shiller (2005)	Cyclically adjusted price-earnings ratio (CAPE) minus the ten-year nominal Treasury yield
Gordon (1962) with real yields	S&P 500 dividend-to-price ratio minus the ten year real Treasury yield (computed as the ten-year nominal Treasury rate minus the ten year breakeven inflation implied by TIPS)
Gordon (1962) with earnings forecasts	S&P 500 expected earnings-to-price ratio minus the ten-year nominal Treasury yield
Gordon (1962) with real yields and earnings forecasts	S&P 500 expected earnings-to-price ratio minus the ten-year real Treasury yield (computed as the ten-year nominal Treasury rate minus the ten-year breakeven inflation implied by TIPS)
Panigirtzoglou and Loeys (2005)	Two-stage DMM. The growth rate of earnings over the first five years is estimated by using the fitted values in a regression of average realized earnings growth over the last five years on its lag and lagged earnings-price ratio. The growth rate of earnings from years six and onwards is 2.2 percent
Damodaran (2012)	A six-stage DDM. Dividend growth the first five stages are estimated from analyst's earnings forecasts. Dividend growth in the sixth stage is the ten-year nominal Treasury yield
Damodaran (2012) free cash flow	Same as Damodaran (2012), but uses free-cash-flow-to-equity as a proxy for dividends plus stock buybacks

Table IV: Models with cross-sectional regressions	
Fama and French (1992)	Uses the excess returns on the market portfolio, a size portfolio and a book-to-market portfolio as risk factors
Carhart (1997)	Identical to Fama and French (1992) but adds the momentum measure of Carhart (1997) as an additional risk factor
Duarte (2013)	Identical to Carhart (1997) but adds an inflation risk factor
Adrian, Crump and Moench (2014)	Uses the excess returns on the market portfolio as the single risk factor. The state variables are the dividend yield, the default spread, and the risk free rate

Table V: Models with time-series regressions	
Fama and French (1988)	Only predictor is the dividend-price ratio of the S&P 500
Goyal and Welch (2008)	Uses, at each point in time, the best out-of-sample predictor out of twelve predictive variables proposed by Goyal and Welch (2008)
Campbell and Thompson (2008)	Same as Goyal and Welch (2008), but imposes two restrictions on the estimation. First, the coefficient b in equation (9) is replaced by zero if it has the "wrong" theoretical sign. Second, we replace the estimate of the ERP by zero if the estimation otherwise finds a negative ERP
Fama and French (2002)	Uses, at each point in time, the best out-of-sample predictor out of three variables: the price-dividend ratio adjusted by the growth rate of earnings, dividends or stock prices
Baker and Wurgler (2007)	The predictor is Baker and Wurgler's (2007) sentiment measure. The measure is constructed by finding the most predictive linear combination of five variables: the closed-end fund discount, NYSE share turnover, the number and average first-day returns on IPOs, the equity share in new issues, and the dividend premium

Table VI: Surveys	
Graham and Harvey (2012)	Chief financial officers (CFOs) are asked since 1996 about the one and ten-year-ahead ERP. We take the mean of all responses

Table VII: ERP models

		Mean	Std. dev.	PC coefficients $\widehat{w}^{(m)}$	Exposure to PC $\widehat{load}_1^{(m)}$
Based on historical mean	Long-run mean	9.3	1.3	0.78	-0.065
	Mean of previous five years	5.7	5.8	0.42	-0.160
DDM	Gordon (1926): E/P minus nominal 10yr yield	-0.1	2.1	-0.01	0.001
	Shiller (2005): 1/CAPE minus nominal 10yr yield	-0.4	1.8	-0.10	0.011
	Gordon (1962): E/P minus real 10yr yield	3.5	2.1	0.69	-0.077
	Gordon (1962): Expected E/P minus real 10yr yield	5.3	1.7	-0.78	0.208
	Gordon (1962): Expected E/P minus nominal 10yr yield	0.4	2.3	-0.79	0.077
	Panigirtzoglou and Loeys (2005): Two-stage DDM	-1.0	2.3	0.07	-0.011
	Damodaran (2012): Six-stage DDM	3.4	1.3	-0.26	0.032
	Damodaran (2012): Six-stage free cash flow DDM	4.0	1.1	-0.62	0.053
Cross-sectional regressions	Fama and French (1992)	12.6	0.7	0.80	-0.040
	Carhart (1997): Fama-French and momentum	13.1	0.8	0.81	-0.042
	Duarte (2013): Fama-French, momentum and inflation	13.1	0.8	0.82	-0.044
	Adrian, Crump and Moench (2014)	6.5	6.9	-0.05	0.114
Time-series regressions	Fama and French (1988): D/P	2.4	4.0	-0.27	0.069
	Best predictor in Goyal and Welch (2008)	14.5	5.2	-0.07	0.023
	Best predictor in Campbell and Thompson (2008)	3.1	9.8	-0.12	0.081
	Best predictor in Fama French (2002)	11.9	6.8	-0.72	0.321
	Baker and Wurgler (2007) sentiment measure	3.0	4.7	-0.32	0.184
Surveys	Graham and Harvey (2012) survey of CFOs	3.6	1.8	0.72	0.264
	All models	5.7	3.2	0.78	-0.065

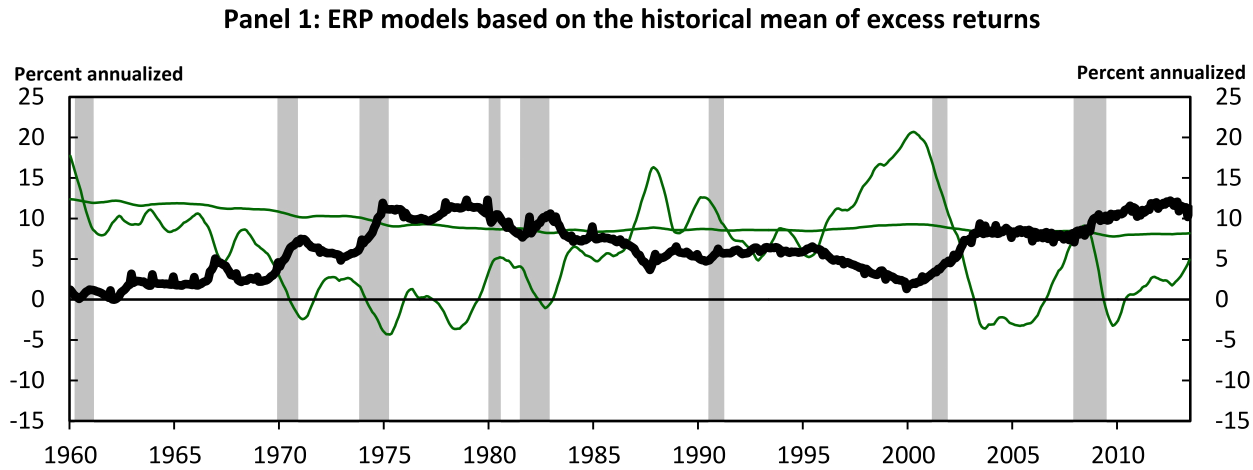
For each of the twenty models of the equity risk premium, we show four statistics. The first two are the time-series means and standard deviations for monthly observations from January 1960 to June 2013 (except for surveys, which are quarterly). The units are annualized percentage points. The third statistic, "PC coefficients $\widehat{w}^{(m)}$", is the weight that the first principal component places on each model (normalized to sum to one). The fourth is the "Exposure to PC $\widehat{load}_1^{(m)}$", the weight on the first principal component when each model is written as a weighted sum of all principal components (also normalized to sum to one).

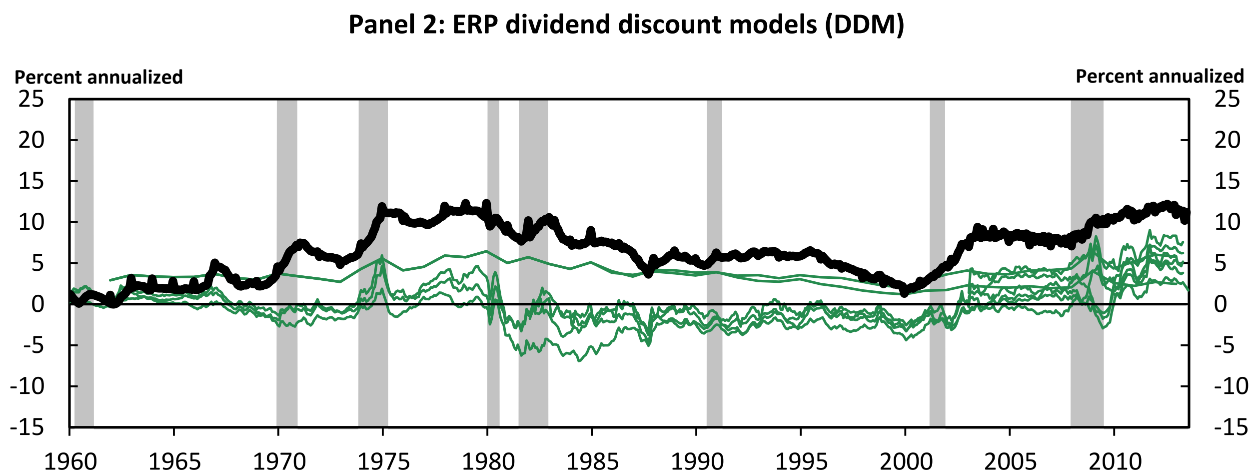
Table VIII: Correlation of ERP models

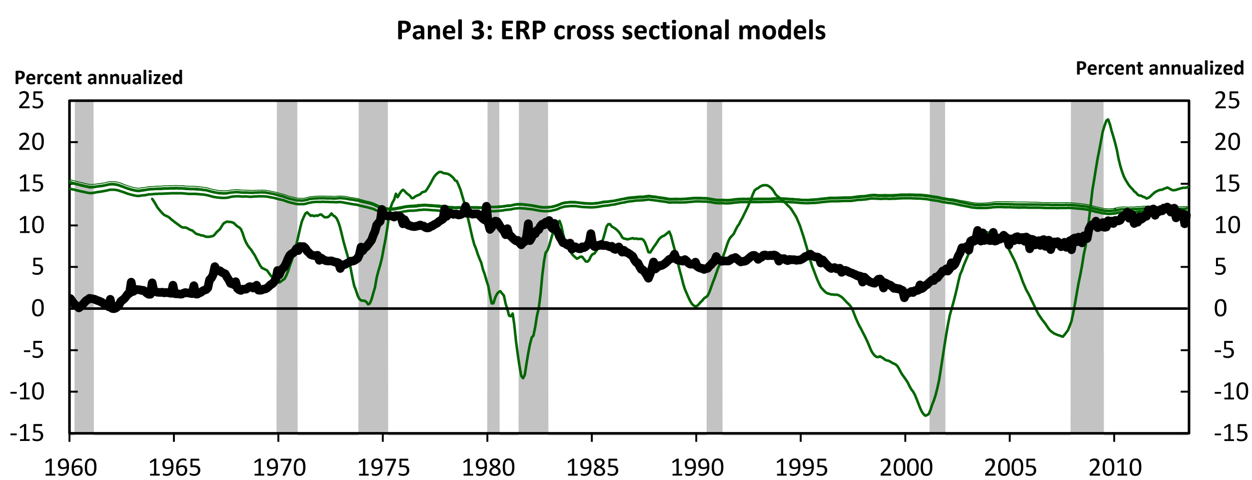
	LR mean	Mean past 5yr	E/P - 10yr	1/CAPE-10yr	E/P-real 10yr	Exp E/P-real 10yr	Exp E/P- 10yr	Two-stage DDM	Six-stage DDM	Free cash flow	FF	Carhart	Duarte	ACM	D/P	G and W	C and T	FF	Sentiment	CFO Survey
LR mean	100																			
Mean past 5yr	32	100																		
E/P - 10yr	8	15	100																	
1/CAPE-10yr	-9	0	78	100																
E/P-real 10yr	-11	25	98	23	100															
Exp E/P-real 10yr	-58	42	70	84	60	100														
Exp E/P- 10yr	-83	-61	84	95	46	98	100													
Two-stage DDM	17	27	88	54	89	66	79	100												
Six-stage DDM	3	-38	26	39	-30	32	52	-31	100											
Free cash flow	-43	-55	59	70	35	80	94	27	62	100										
FF	69	29	-8	-36	-21	-69	-91	9	-29	-77	100									
Carhart	71	30	-5	-31	-24	-71	-91	10	-25	-75	99	100								
Duarte	71	30	-3	-29	-22	-70	-91	11	-28	-74	99	100	100							
ACM	-1	-52	36	62	6	54	63	27	23	33	-28	-28	-25	100						
D/P	49	12	27	12	27	42	54	24	74	42	44	54	55	21	100					
G and W	25	12	25	21	-7	-36	-60	20	29	-9	7	13	14	-24	61	100				
C and T	27	31	14	-7	81	49	-60	28	-51	-40	60	57	58	-33	54	50	100			
FF	1	-30	-24	-29	37	-27	-37	-18	22	38	36	38	37	-9	40	23	43	100		
Sentiment	-10	33	-4	-20	68	-23	-29	27	-38	-20	18	17	18	-12	-38	-8	21	6	100	
CFO survey	-43	-33	12	30	1	1	13	16	5	-3	-36	-37	-39	60	14	-21	-32	-3	-36	100

This table shows the correlation matrix of the twenty equity risk premium models we consider. Numbers are rounded to the nearest integer. Thick lines group models by their type (see Tables II to VI). Except for the CFO survey, the observations used to compute correlations are monthly for January 1960 to June 2013. For the CFO survey, correlations are computed by taking the last observation in the quarter for monthly series and then computing quarterly correlations.

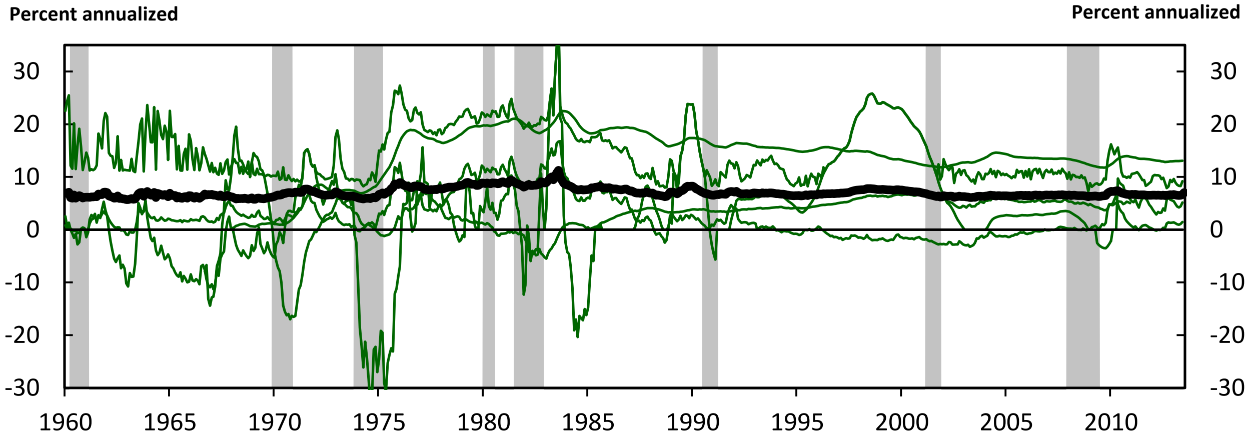
Figure 1: ERP estimates for all models



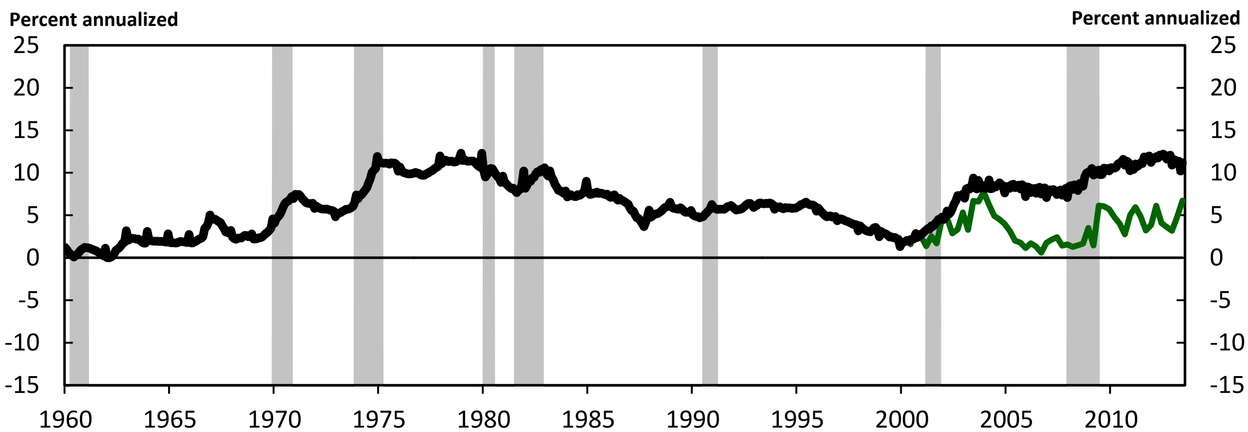




Panel 4: ERP time series models



Panel 5: ERP surveys

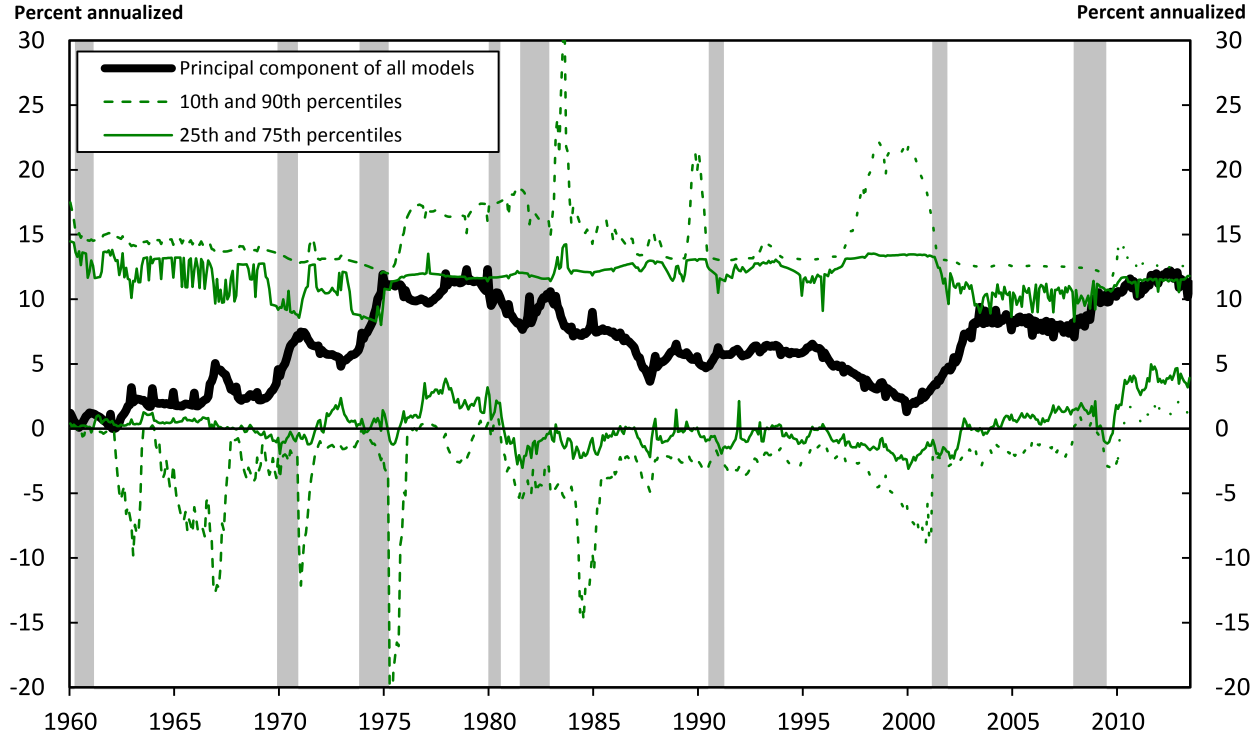


Each green line gives the one-year-ahead equity risk premium from each of the models listed in Tables II to VI. All numbers are in annualized percentage points.

Panel 1 shows the estimates for models based on the historical mean of excess returns, which are listed in Table II. Panel 2 shows estimates computed by the dividend discount models in Table III. Panel 3 uses the cross-sectional regression models from Table IV. Panel 4 shows the equity risk premium computed by the time-series regression models in Table V. Panel 5 gives the estimate obtained from the survey cited in Table VI.

In all panels, the black line is the first principal component of all twenty models (it can look different across panels due to different scales in the y-axis).

Figure 2: One-year-ahead ERP

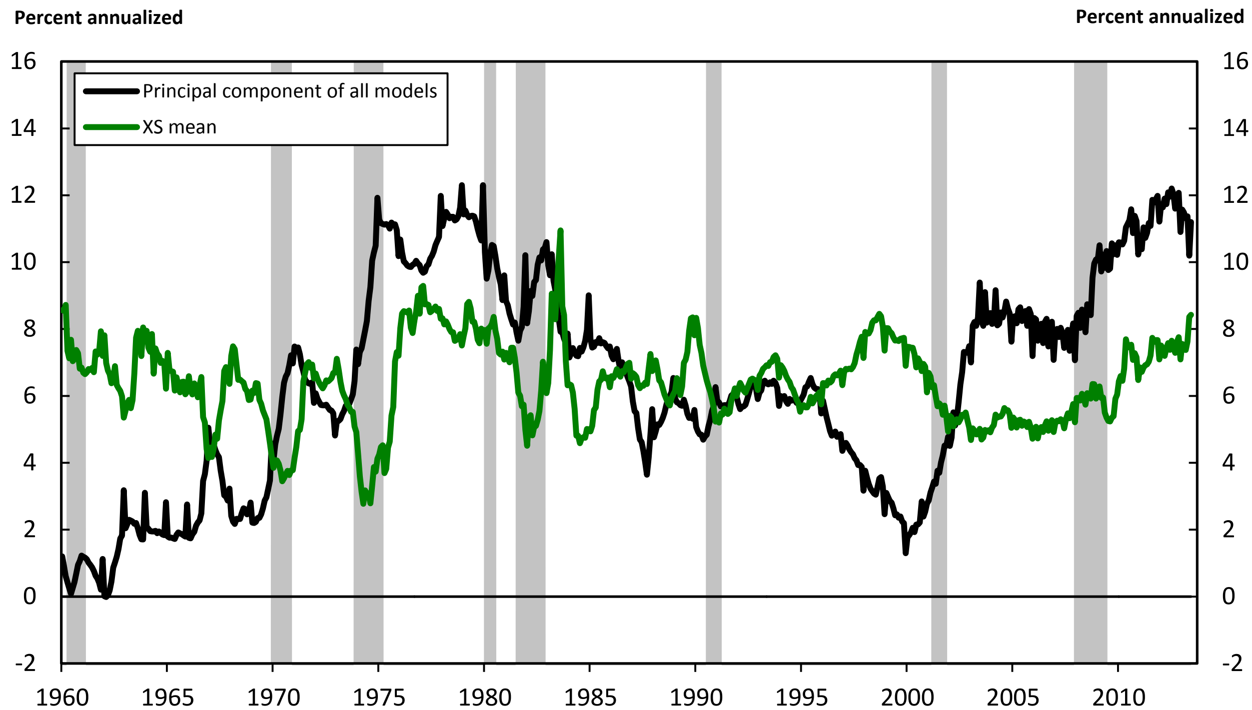


The black line is the first principal component of twenty models of the one-year-ahead equity risk premium (this is the same principal component shown in black in all panels of Figure 1). The models are listed in Tables II to VI.

The 25th and 75th percentiles (solid green lines) give the corresponding quartile of the 20 estimates for each time period, and similarly for the 10th and 90th percentiles (dashed green line).

Shaded bars indicate NBER recessions.

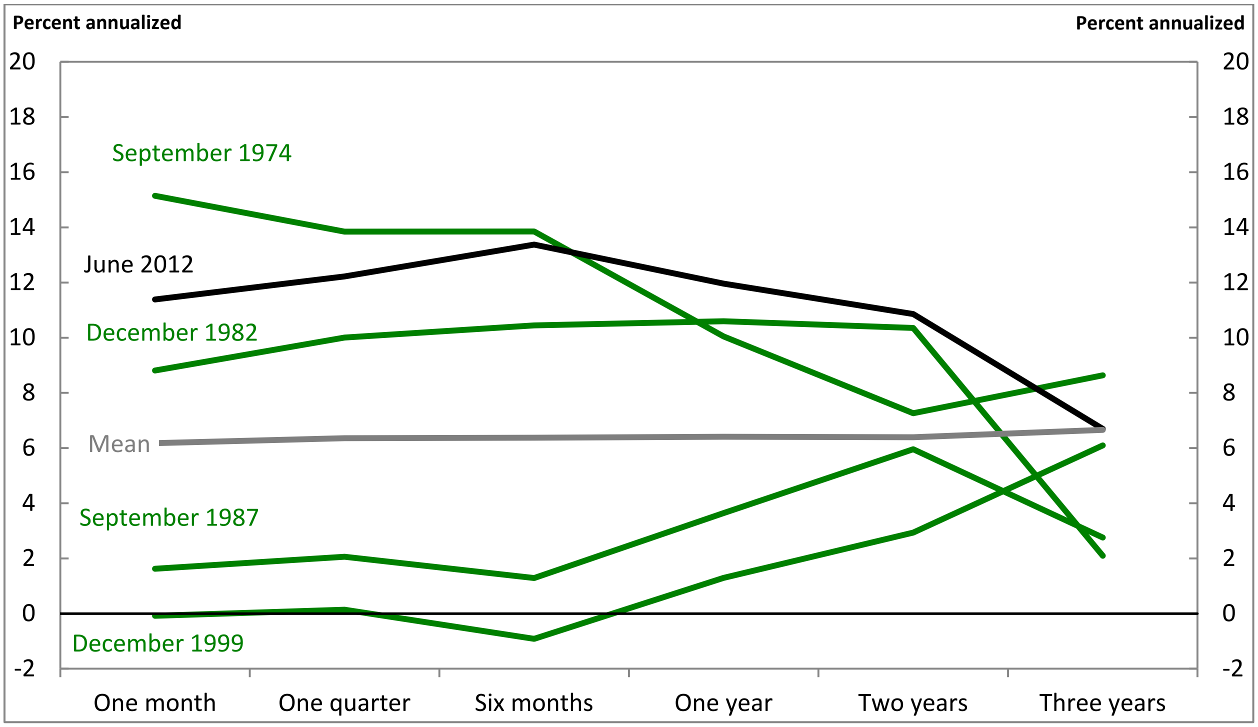
Figure 3: One-year-ahead ERP and cross-sectional mean of models



The black line is the first principal component of twenty models of the one-year-ahead equity risk premium (also shown in Figures 1 and 2). The green line is the cross-sectional average of models for each time period.

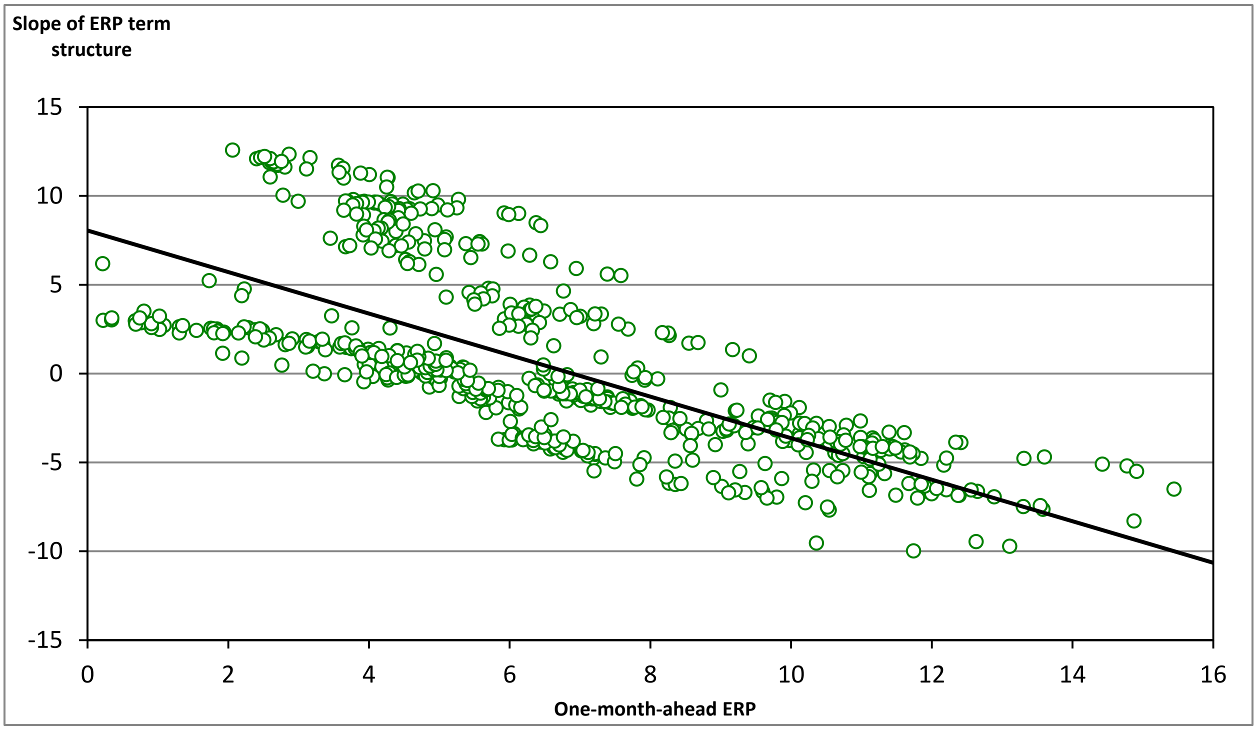
Shaded bars are NBER recessions.

Figure 4: Term structure of the ERP



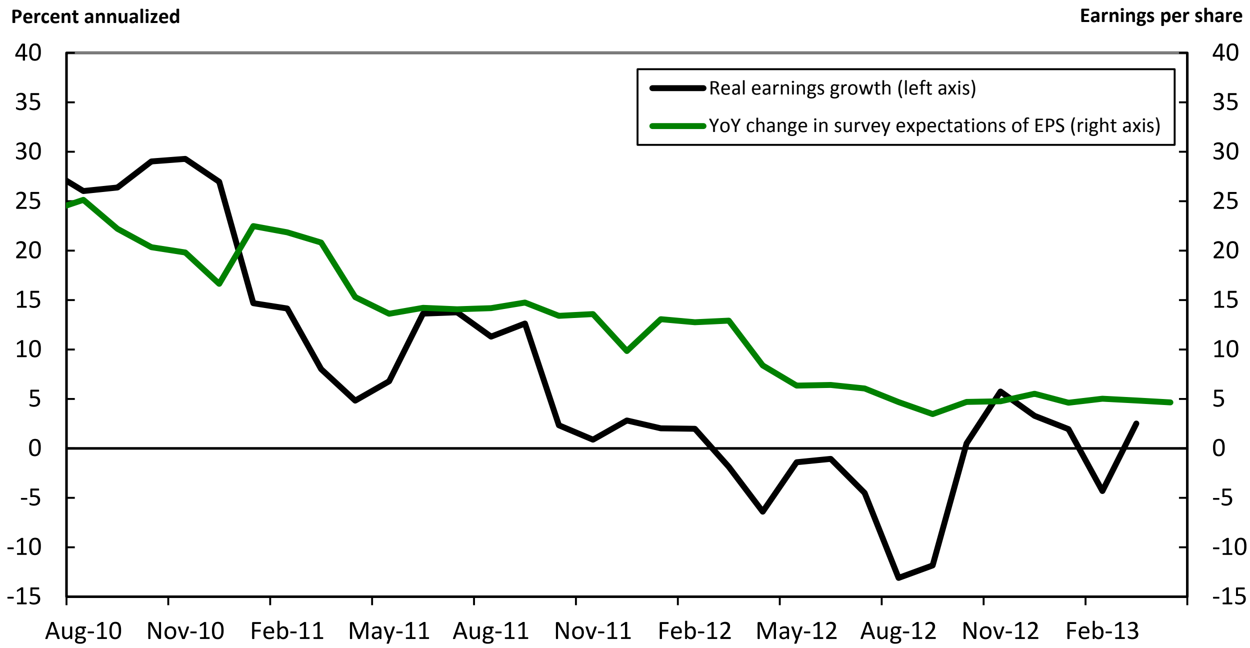
Each line, except for the grey one, shows equity risk premia as a function of investment horizon for some specific months in our sample. We consider horizons of one month, one quarter, six months, one year, two years and three years. The grey line (labeled "Mean") shows the average risk premium at different horizons over the whole sample January 1960 to June 2013. September 1987 and December 1999 were low points in one-month-ahead equity premia. In contrast, September 1974, December 1982 and June 2012 were peaks in the one-month-ahead equity premium.

Figure 5: Regression of the slope of the ERP term structure on one-month-ahead ERP



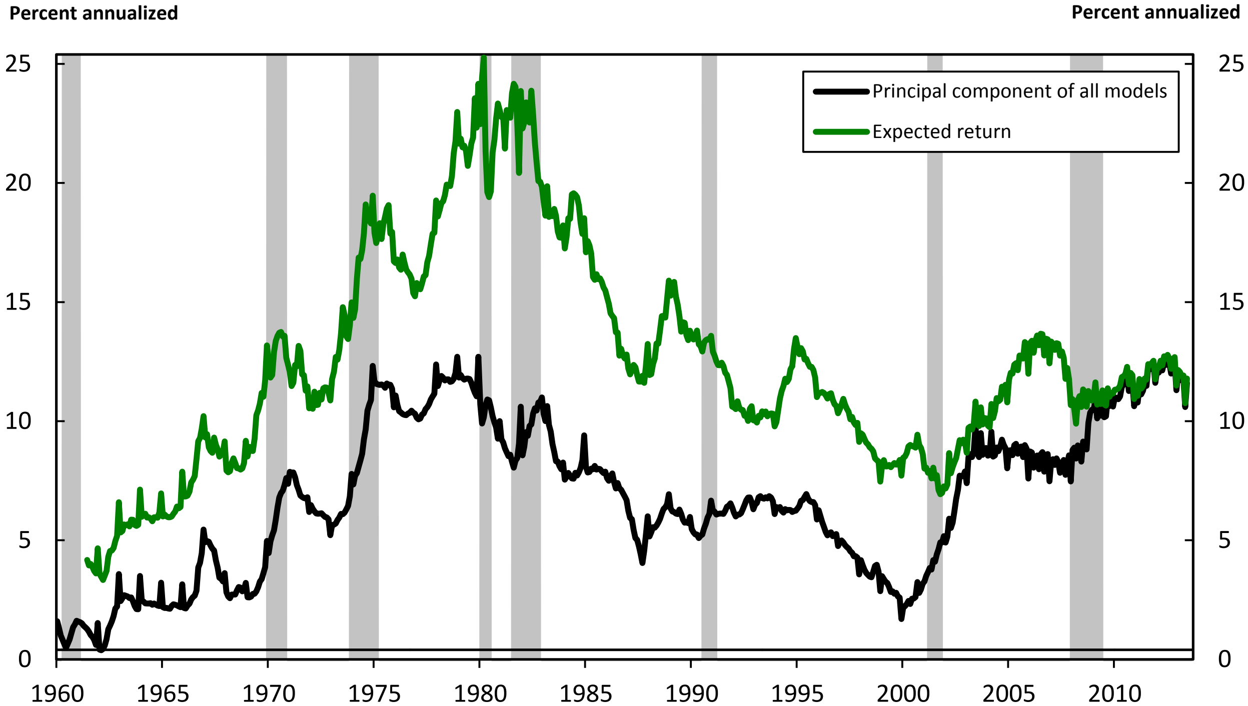
The figure shows monthly observations and the corresponding OLS regression for of the one-month-ahead ERP plotted against the slope of the ERP term structure for the period January 1960 to June 2013. The slope of the ERP term structure is the difference between the three-year-ahead ERP and the one-month-ahead ERP. All units are in annualized percentage points. The one-month-ahead and three-year-ahead ERP estimates used are the first principal components of twenty one-month-ahead or three-year-ahead ERP estimates from models described in Tables II-VI. The OLS regression slope is -1.17 (significant at the 99 percent level) and the R^2 is 50.1 percent.

Figure 6: Earnings behavior



The black line shows the monthly growth rate of real S&P 500 earnings, annualized and in percentage points. The green line shows the year-on-year change in the mean expectation of one-year-ahead earnings per share for the S&P 500 from a survey of analysts provided by Thomson Reuters I/B/E/S.

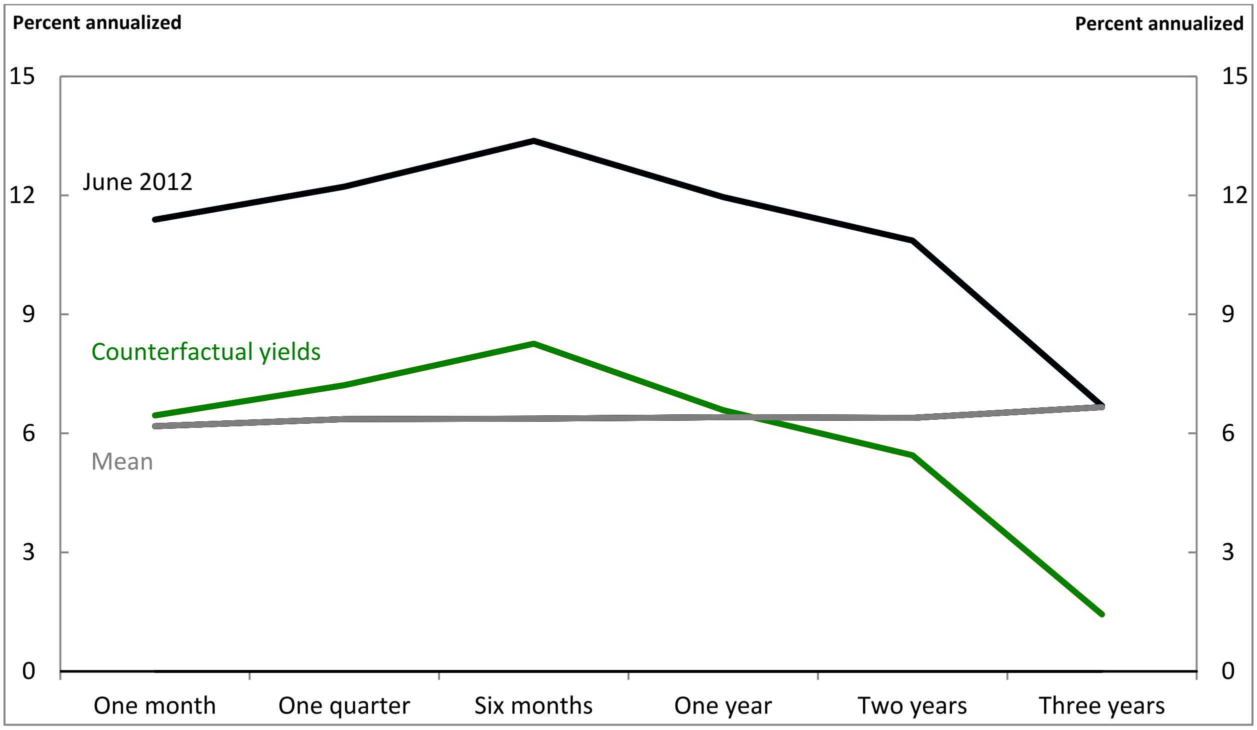
Figure 7: One-year-ahead ERP and expected returns



The black line is the first principal component of twenty models of the one-year-ahead equity risk premium (also shown in Figures 1, 2 and 3). The green line is the one-year-ahead expected return on the S&P 500, obtained by adding the realized one-year maturity Treasury yield from the principal component (the black line).

Shaded bars are NBER recessions.

Figure 8: Term structure of ERP using counterfactual bond yields



The grey line, labeled “Mean”, shows the mean term structure of the equity risk premium over the sample January 1960 to June 2013. The black line, labeled “June 2012”, shows the term structure for the most recent peak in the one-month-ahead ERP. These two lines are the same as in Figure 4. The green line, labeled “Counterfactual yields”, shows what the term structure of equity risk premia would be in June 2012 if instead of subtracting June 2012’s yield curve from expected returns we subtracted the average yield curve for January 1960 to June 2013.

Duff & Phelps

Table: Equity Risk Premium & Risk-Free Rates

March 6, 2014

Duff & Phelps Recommended Equity Risk Premium (ERP) and Corresponding Risk-Free Rates (R_f); January 2008–Present

For additional information, please visit www.duffandphelps.com/CostofCapital

	Duff & Phelps Recommended ERP	Risk-Free Rate
Current ERP Guidance February 28, 2013 – UNTIL FURTHER NOTICE	5.0%	4.0% Normalized 20-year Treasury yield *
Year-end 2013 Guidance December 31, 2013	5.0%	4.0% Normalized 20-year Treasury yield *
January 1, 2013 – February 27, 2013	5.5%	4.0% Normalized 20-year Treasury yield *
Year-end 2012 Guidance December 31, 2012	5.5%	4.0% Normalized 20-year Treasury yield *
Change in ERP Guidance January 15, 2012 – February 27, 2013	5.5%	4.0% Normalized 20-year Treasury yield *
Change in ERP Guidance September 30, 2011 – January 14, 2012	6.0%	4.0% Normalized 20-year Treasury yield *
July 1 2011 – September 29, 2011	5.5%	4.0% Normalized 20-year Treasury yield *
June 1, 2011 – June 30, 2011	5.5%	Spot 20-year Treasury Yield
May 1, 2011 – May 31, 2011	5.5%	4.0% Normalized 20-year Treasury yield *
December 1, 2010 – April 30, 2011	5.5%	Spot 20-year Treasury Yield
June 1, 2010 – November 30, 2010	5.5%	4.0% Normalized 20-year Treasury yield *
Change in ERP Guidance December 1, 2009 – May 31, 2010	5.5%	Spot 20-year Treasury Yield
June 1, 2009 – November 30, 2009	6.0%	Spot 20-year Treasury Yield
November 1, 2008 – May 31, 2009	6.0%	4.5% Normalized 20-year Treasury yield *
Change in ERP Guidance October 27, 2008 – October 31, 2008	6.0%	Spot 20-year Treasury Yield
January 1, 2008 – October 26, 2008	5.0%	Spot 20-year Treasury Yield

* Normalized in this context means that in months where the risk-free rate is deemed to be abnormally low, a proxy for a longer-term sustainable risk-free rate is used.

DUFF & PHELPS

Table: Equity Risk Premium & Risk-Free Rates

January 31, 2016

Duff & Phelps Recommended U.S. Equity Risk Premium (ERP) and Corresponding Risk-Free Rates (R_f); January 2008–Present

For additional information, please visit www.duffandphelps.com/CostofCapital

	Duff & Phelps Recommended ERP	*Risk-Free Rate*
Change in ERP Guidance (current guidance) January 31, 2016 – UNTIL FURTHER NOTICE	5.5%	4.0% Normalized 20-year Treasury yield *
Year-end 2015 Guidance December 31, 2015	5.0%	4.0% Normalized 20-year Treasury yield *
Change in ERP Guidance February 28, 2013 – January 30, 2016	5.0%	4.0% Normalized 20-year Treasury yield *
Change in ERP Guidance January 15, 2012 – February 27, 2013	5.5%	4.0% Normalized 20-year Treasury yield *
Change in ERP Guidance September 30, 2011 – January 14, 2012	6.0%	4.0% Normalized 20-year Treasury yield *
July 1 2011 – September 29, 2011	5.5%	4.0% Normalized 20-year Treasury yield *
June 1, 2011 – June 30, 2011	5.5%	Spot 20-year Treasury Yield
May 1, 2011 – May 31, 2011	5.5%	4.0% Normalized 20-year Treasury yield *
December 1, 2010 – April 30, 2011	5.5%	Spot 20-year Treasury Yield
June 1, 2010 – November 30, 2010	5.5%	4.0% Normalized 20-year Treasury yield *
Change in ERP Guidance December 1, 2009 – May 31, 2010	5.5%	Spot 20-year Treasury Yield
June 1, 2009 – November 30, 2009	6.0%	Spot 20-year Treasury Yield
November 1, 2008 – May 31, 2009	6.0%	4.5% Normalized 20-year Treasury yield *
Change in ERP Guidance October 27, 2008 – October 31, 2008	6.0%	Spot 20-year Treasury Yield
January 1, 2008 – October 26, 2008	5.0%	Spot 20-year Treasury Yield

* Normalized in this context means that in months where the risk-free rate is deemed to be abnormally low, a proxy for a longer-term sustainable risk-free rate is used.

Review of Recent Research on Improving Earnings Forecasts and Evaluating Accounting-based Estimates of the Expected Rate of Return on Equity Capital

Peter D. Easton
Center for Accounting Research and Education
305A Mendoza College of Business
Notre Dame, Indiana 46556-5646
United States
peaston@nd.edu

and

Steven J. Monahan
INSEAD
Boulevard de Constance
PMLS 1.24
F-7705 Fontainebleau Cedex, 77305
France
steven.monahan@insead.edu

March 2015

Written for a special issue of *Abacus* on Financial Statement Analysis and Valuation

Abstract

We extend Easton's (2007) review of the literature on accounting-based estimates of the expected rate of return on equity capital, which we refer to as the ERR. We begin by reiterating the reasons why accounting-based estimates are used. Next, we briefly review the recent literature that focuses on improving forecasts of expected earnings by either: (1) removing predictable errors from analysts' forecasts of earnings or (2) developing cross-sectional regression-based estimates of earnings using prior-period financial data. In the remainder of our review we discuss a recent debate on methods for evaluating estimates of the ERR. We highlight the key points in the debate so that the reader will find it easier to form an independent view of the relative merits of the proposed methods.

1. Why Use an Accounting-based Estimates of the Expected Rate of Return?

The answer to this question is straightforward: there is no reliable alternative estimate. Users of accounting-based estimates of the expected rate of return on equity capital, ERR, are making two implicit assumptions.[1] The first implicit assumption, which we refer to as IA1, is that neither firm- nor portfolio-level realized returns are a reliable measure of expected returns. The second implicit assumption, which we refer to as IA2, is that the factors that determine expected returns are unknown and/or that they cannot be estimated reliably. If the user is not making these assumptions, there is no need to use an accounting-based estimate. Rather, either realized returns or an estimate taken from an asset pricing model may be used.

1.1 IA1: Realized Returns are not Reliable Measures of Expected Returns

Users of accounting-based estimates of the ERR are implicitly assuming that: (1) firm-level realized returns are not a reliable measure of expected returns and/or (2) for their sample, it is infeasible to obtain reliable estimates of the ERR via temporal or cross-sectional averaging of firm-level returns. For example, a researcher may be interested in a small sample of firms with a short trading history, in which case cross-sectional and temporal averaging may be infeasible.[2] On the other hand, if the requisite data are available, accounting-based estimates of the ERR may be obtained for each firm in the sample.

IA1 is not an unreasonable assumption/conclusion. For instance, since Black, Jensen and Scholes (1972) and Fama and MacBeth (1973) it has been the norm in empirical asset pricing to use portfolio-level returns (e.g., value weighted averaging) instead of firm-level

[1] We refer to "users" of accounting-based proxies because much of our discussion is pertinent to people outside of academia.

[2] Alternatively, the researcher may have a long time-series of realized returns for each firm in the sample but may be concerned that the moments of the distribution are not stationary.

returns. But, portfolio-level returns are also suspect. For example, in his presidential address to the American Finance Association, Elton (1999) (p. 1199) states: ‘‘*The use of average realized returns as a proxy for expected returns relies on a belief that information surprises tend to cancel out over the period of a study and realized returns are therefore an unbiased measure of expected returns. However, I believe there is ample evidence that this belief is misplaced.*’’ We discuss this issue further in section 3 of this review.

1.2 IA2: Risk Factors are either Unknown or cannot be Reliably Estimated

This assumption is not controversial. On the contrary, the lack of consensus regarding the manner in which economic agents make risk-return trade-offs is well documented (e.g., chapter 20 of Cochrane (2001) and chapters six and seven of Campbell, Lo and MacKinley (1997) review the issues). While the four-factor model inspired by Fama and French (1993) and Carhart (1997) has become *de rigueur*, it is controversial; and, it is not based on a well-accepted theory of capital market equilibrium. Moreover, three of the four factors (i.e., size, book-to-market, and momentum) originally appeared in the literature under the guise of anomalies. These factors were later designated as risk factors purely on the basis of their ability to explain variation in returns.[3] For example, when discussing momentum in chapter 20 of his text Cochrane (2001) makes the following statement (p. 446). "*Momentum stocks move together, as do value and small stocks so a ‘momentum factor’ works to ‘explain’ momentum portfolio returns. This is so obviously ad-hoc (i.e. an APT factor that will only explain returns of portfolios organized on the same characteristic*

[3] The size, book-to-market and momentum effects were introduced by Banz (1981), Rosenberg, Reid and Lanstein (1985), and Jegadeesh and Titman (1993), respectively. Moreover, there is considerable evidence supporting the notion that the returns to these strategies are anomalous. For example, Lakonishok, Shleifer, and Vishny (1992), LaPorta et al. (1997) and Piotroski (2000)) provide evidence on the book-to-market effect.

as the factor) that nobody wants to add it as a risk factor." Nonetheless, momentum is now commonly included as a factor in empirical asset-pricing tests.

In addition, estimates of the ERR taken from factor models do not appear to be reliable. Evidence of this is provided by Fama and French (1997) who evaluate annual, industry-level estimates of the ERR and show that the temporal standard error is more than three percent for estimates based on the capital-asset pricing model and the three-factor model of Fama and French (1993). Hence, in the abstract to their paper they conclude that: "*Estimates of cost of equity for industries are imprecise. ... Estimates of the cost of equity for firms and projects are surely even less precise.*"

1.3 Summary

Implicit in the use of accounting-based estimates of the ERR is the assumption that alternative methods of estimating the ERR are infeasible. While this assumption is reasonable, its veracity is not the central issue. Rather, the central issue is that it is logically inconsistent to use an accounting-based estimate and then to proceed as if either IA1 or IA2 is invalid. Why? If one of these assumptions is invalid, a reliable ERR estimate may be obtained from either realized returns or a factor model. However, if this is possible, the reliability of accounting-based estimates is a moot point.

2. Improving Forecasts of Earnings

2.1 Models Based on Earnings Levels versus Models Based on Earnings Changes

Extant methods of estimating the implied expected rate of return using current market prices and earnings forecasts fall naturally into two groups: those based on forecasts of earnings levels and those based on forecasts of earnings changes. These methods are described in detail in numerous papers (e.g., Easton (2007) provides a comprehensive description and critique).[4] We do not repeat the details here; rather, we briefly describe the underlying models. We draw the distinction between methods based on earnings levels and those based on earnings changes because methods designed to improve earnings forecasts are more effective in the former than in the latter.[5]

2.1.1 Methods Based on Forecasts of Earnings Levels

The residual income valuation, RIV, model (generally based on a version of Claus and Thomas (2001) or Gebhardt et al. (2001)) is the most commonly-used earnings-levels-based model. Per the RIV model the ERR is the number that causes equity market value to equal the sum of: (1) equity book value and (2) the present value of expected future residual income. Residual income is estimated as expected earnings less the product of the ERR and beginning equity book value. Another earnings-levels-based model is described in Easton and Monahan (2005). In this model, the ERR is the rate that equates equity market value to the present value of multi-period forecasts of cum-dividend earnings levels.

2.1.2 Methods based on Forecasts of Earnings Changes

The models based on forecasts of earnings changes are based on the abnormal earnings growth, AEGV, model. Per the AEGV model, the ERR is the number that causes equity market value to equal the sum of: (1) capitalized expected earnings in year t+1 and (2) the present value of capitalized abnormal earnings growth, AEG, subsequent to year

[4] Although Easton (2007) has a publication date of 2007, it reviews the literature through 2009.

[5] The reason for this is two-fold: (1) the focus of extant research and (2) empirical properties. First, as discussed in this section 2.2, extant research typically focuses either on forecasting annual earnings *levels* for a several years—i.e., t+1 through t+h—or adjusting analysts' forecasts of earnings *levels* for years t+1 through t+h. Second, regarding the empirical issue, extant models typically generate forecasts of (adjustments to) earnings (analysts' forecasts of earnings) for year t+1 that are very similar to forecasts of (adjustments to) earnings (analysts' forecasts of earnings) for year t+h. Hence, the implied forecast of the *change* in earnings (adjustment to the *change* in analysts' forecasts of earnings) obtained from these models is essentially random noise.

t+1. AEG in year t equals the difference between: (1) the expected change in earnings in year t+1 and (2) the ERR multiplied by the difference between earnings in year t and dividends in year t. For example, per the *PEG* model, which is critiqued by Easton (2004), the ERR equals the square root of the ratio of the expected change in earnings in year t+1 divided by equity market value in year t. (Easton and Monahan (2005) discuss the other AEGV models used in the literature.)

2.2 Improving data on Forecasts of Earnings

Two quite different approaches have been taken to improving the data used as earnings forecasts: (1) removing predictable errors from analysts' forecasts and (2) developing forecasts from cross-sectional models.

2.2.1 Removing Predictable Errors

Two recent papers, Larocque (2013) and Mohanram and Gode (2013), estimate and then adjust for predictable errors in analysts' earnings forecasts. Both papers estimate predictable errors via a regression (using data that are available as of year t) of analysts' forecast errors on variables that they argue are predictors of these errors. Larocque's predictor variables are lagged forecast errors, lagged abnormal stock returns, lagged equity market value, and the abnormal return between the forecast date and the earnings announcement date. Mohanram and Gode (2013) use lagged accruals; lagged sales growth; the lagged analysts' forecast of long-term earnings growth; lagged change in property, plant and equipment; lagged change in other total assets; lagged stock returns; and the revision in analysts' forecasts of earnings over the prior year. Each paper then uses its respective predictors and the estimated regression coefficients to predict the error in analysts' forecasts of year t+1 and year t+2 earnings. Both methods are effective in removing errors in forecasts of earnings levels but, not surprisingly, they are less effective in removing errors in forecasts of earnings changes. Moreover, an obvious limitation of these methods is that they are only applicable to firms that are covered by analysts.

2.2.2 Using Mechanical Models to Forecast Earnings

Hou, van Dijk and Zhang (2012) (HVZ hereafter) extend the model in Fama and French (2002) to obtain forecasts of earnings for the next two years. HVZ's model is based on a regression of year t earnings on lagged financial statement data. Their claim that they provide improved earnings forecasts and, therefore, improved estimates of the ERR is valid inasmuch as they provide forecasts for a wider set of observations (i.e., beyond the subset of observations for which researchers have access to analysts' forecasts). However, they do not compare forecast for which there is both an analyst forecast and a forecast from their regression-based model and, it seems probable that the analysts' forecasts (and the analyst based estimates of the ERR) are superior for these firms. It is also important to note that for a large portion of the observations, forecasts of earnings levels and, particularly, forecasts of earnings changes from the method in HVZ will be negative. Hence, these forecasts are unusable in estimating the ERR. Furthermore, it is important to note that two papers (Gerakos and Gramacy (2013) and Mohanram and Li (2014)) show that the earnings forecast errors from the HVZ model are quite similar to errors obtained from a random walk model, which casts considerable doubt on whether HVZ's model should be used.

3. Evaluating Estimates of the ERR

In this section and the next section we clarify several key issues related to the use and evaluation of accounting-based estimates of the ERR. The impetus for our comments is four-fold. First, and foremost, the ERR is an important construct for practitioners, policy-makers, and academics. It is, however, unobservable and, thus, estimates are used as empirical proxies. In light of this fact, the properties and construct validity of various estimates should be carefully examined and understood.

Second, accounting-based estimates of the ERR are becoming commonplace in both the accounting and finance research literatures. There is, however, conflicting empirical evidence regarding their reliability and research that evaluates these estimates is often described as controversial. We believe that the root of the controversy is not well understood or at least poorly articulated. An aim of this paper is to clarify the issues and, thereby, resolve the controversy. We explain that the approach adopted by Botosan and Plumlee (2005) (BP hereafter), which is one of the two competing evaluation approaches, is logically inconsistent with a key, implicit assumption that motivates the use of accounting-based estimates. It follows that their approach cannot yield meaningful inferences.

Third, we revisit and elaborate on an earlier paper Easton and Monahan (2005) (EM hereafter), in which we developed and implemented an alternative approach to the one used by BP. This approach integrates the implicit assumptions that motivate the use of accounting-based estimates. Hence, we argue that, relative to BP's approach, EM's approach is a more appropriate way of evaluating the reliability of accounting-based estimates of the ERR. Of course, it behoves us to elaborate on our approach so that others may draw their own conclusions about its merits and shortcomings.

Finally, in a more recent paper, Botosan, Plumlee and Wen (2011) (BPW hereafter) assert that the empirical results in EM are (p. 1119) *"... attributable to an omitted variable* *bias arising from a lack of adequate controls for new information.*" We disagree with this statement and we explain why it is incorrect.

In this section, we first provide a brief overview of the "controversy." Next, we discuss BP's approach. Finally, we describe the approach used by EM. In section four, we explain why criticisms made by BPW of research design choices made by EM are unwarranted.

3.1. Overview

Presently there are two empirical approaches for evaluating the reliability of accounting-based estimates of the ERR: (1) the approach described in BP and (2) the approach developed and described in EM. These approaches rely on different methodologies and, to some extent, generate different results, possibly leading to the label, "controversial."

We believe the controversy regarding differences in the empirical results is minimal (at best) for two, related reasons. First, while BP infer that certain estimates are reliable for their sample of firms, EM also find that certain (different) estimates are reliable for nontrivial subsets of the sample they study. We believe this fact is often overlooked and that many are under the impression that EM conclude that accounting-based estimates are never reliable. They do not. For example, see the abstract on p. 501; discussions on p. 503 and pp. 526-531; and, results in Panel C of Table 9 of EM.

Second, we believe that the reliability, or lack thereof, of a particular proxy is likely sample specific. Hence, the results in BP and EM are less relevant than the relative merits of their methodologies. In particular, we believe that interested researchers should: (1)

focus on deciding which methodology is most appropriate and (2) use that methodology to evaluate the accounting-based estimates that they estimate for their sample.

Thus, we believe the heart of the controversy relates to methodological differences. Moreover, as we explain below, these methodological differences are rooted in different implicit assumptions made by BP and EM. BP implicitly assume that the factors that determine expected returns are known and that these factors can be reliably measured. As discussed in section 1.2, there are two problems with this assumption. First, it is not supported by the data or by extant theory. Second and, more importantly, it is logically inconsistent with the motivations underlying the use of accounting-based estimates of the ERR. That is, if the risk factors are known and can be reliably measured, why not simply use them instead of potentially unreliable accounting-based estimates?

EM, on the other hand, base their approach on the assumption that realized returns are biased and noisy measures of expected returns. This assumption is one of the primary motivations underlying the use of accounting-based estimates. Hence, EM's methodology is logically consistent with the underlying research question.

Finally, we note an important caveat. We argue that EM's approach is the best extant approach. That said we recognize that all empirical approaches have limitations and rely on assumptions. We conclude that EM's approach has less limitations and relies on less restrictive assumptions than the approach adopted by BP.

3.2. Discussion of Botosan and Plumlee (2005)

BP regress accounting-based estimates of the ERR on estimates of firm-specific variables (e.g., estimated CAPM beta, equity market value, book-to-market, etc.). They use two criteria to evaluate reliability. First, they consider the sign and statistical significance of the regression coefficients. For example, a reliable proxy is one that has a positive association with estimated CAPM beta. Second, they consider r-squares: higher r-squares imply greater reliability.

Our primary concern with BP's research design is that it is logically inconsistent with the underlying research question. In particular, as discussed in section 1.2, an implicit assumption (i.e., IA2) underlying the use of accounting-based estimates is that the risk factors are unknown and/or that they cannot be reliably estimated. The fact that BP evaluate the relation between accounting-based estimates and potential risk factors suggests that they believe IA2 is false. If they do not, the motivation for their tests and the interpretation of their results is unclear. If the factors that BP use to evaluate the reliability of accounting-based estimates are not the "true" risk factors, what exactly do we learn from BP's tests? Stated another way, it is illogical to evaluate the reliability of one proxy by comparing it to another set of proxies that may also be unreliable.

Although we believe our primary concern is quite valid, we anticipate at least two counter-arguments. We refer to these as: (1) the evidence by analogy argument and (2) the proof is in the pudding argument. In the following sub-sections, we elaborate on these arguments and we explain our thoughts regarding their merits. We anticipate these arguments because we have heard them during academic workshops and/or during private conversations. These arguments also serve as a rhetorical device: by discussing them, we are able to clarify our concerns about BP's approach.

3.2.1 The Evidence by Analogy Argument

A potential argument for BP's approach is: “We know the true factor model and estimation of the risk factors is feasible for many (i.e., normal) firms but not all firms.” For instance, some firms have short trading histories or have recently experienced major structural changes (e.g., large acquisitions). For these firms accounting-based estimates are the only alternative. However, since these accounting-based estimates may be unreliable, they must be evaluated. The researcher does this by analogy. The relation between an accounting-based estimate and the risk factors is evaluated for normal firms. If the relation between the accounting-based estimate and the risk factors accords with the theory then, by analogy, the accounting-based estimate is also assumed to be reliable for the sample of "abnormal" firms.

This argument is unconvincing for two reasons. First, IA2 is not controversial; rather, there is no consensus regarding the identity of the true factor model and estimates of the factors presently used in empirical finance are fraught with error. Second, accounting-based proxies are often used to evaluate samples of firms that are arguably "normal" and, thus, researchers are not acting as if they believe the evidence by analogy argument.

3.2.2 The Proof is in the Pudding Argument

Another argument for BP's approach is as follows: "Although the economic meaning of the firm-level variables used by BP is unclear, they work—i.e., the proof is in the pudding." In particular, some of the variables considered by BP (e.g., book-to-market and size) explain variation in average realized returns. Hence, they appear to explain variation in expected returns. Whether this variation is fully attributable to differences in risk is irrelevant. We believe this argument has some merit; it is, however, subject to several important caveats. We first describe the merits and then we provide caveats.

We agree with one part of the proof is in the pudding argument: whether an estimate of the ERR reflects risk or mis-pricing, is not the central issue. Rather, the central issue is whether a particular estimate is a reliable measure of expected return. In fact, an estimate that only reflects risk is imperfect if expected returns are also a function of non-risk factors. Why? If this is the case, we cannot use the estimate to draw unbiased inferences about the nature of expected returns. For example, tests based on it cannot reject a null hypothesis of market efficiency even if the null is false.

There are at least three important caveats regarding the proof is in the pudding argument. First, appearances can be deceiving. Lewellen, Nagel and Shanken (2010) provide evidence that the positive associations between a number of factors and portfolio-level realized returns are purely attributable to research design flaws and that once these flaws are eliminated, the associations disappear. In addition, as discussed in chapters five and six of Campbell, Lo and MacKinlay (1997), data-snooping bias and sample-selection bias are always potential concerns when testing factor models.

Second, whether some of the variables considered by BP "work" is debatable—i.e., i.e., the "proof in the pudding" is either weak or non-existent. BP include in their set of firm-level variables capital asset pricing model, CAPM, beta, a leverage measure, a measure of expected future earnings growth, and an information-risk measure. There is, however, little or no evidence that these variables are risk proxies. For example, there is little empirical evidence that supports a positive association between CAPM beta and returns. Bhandari (1988), Johnson (2004), Nielson (2006), George and Hwang (2010), Ipplolito, Steri and Tebaldi (2011), and Caskey, Hughes and Liu (2012) show that, despite well-known analytical results in Modigliani and Miller (1958), there is a *negative* relation between leverage and returns. LaPorta's (1996) evidence regarding earnings growth is based on a small sample and, thus, while interesting, it is not authoritative. Finally, there is an ongoing debate regarding the pricing of information risk and the evidence is mixed. For example, consider accruals quality. Francis, LaFond, Olsson and Schipper (2004) conclude that it is a priced factor. On the other hand, Core, Guay and Verdi (2008) provide evidence that this conclusion is unwarranted.[6]

Finally, the proof in the pudding argument is insular as it prevents us from "stepping outside the model." Dissatisfaction with factor models is one of the primary motivations for using accounting-based estimates. However, if we choose to evaluate accounting-based estimates by relating them to different factors, we cannot completely avoid the problems associated with using and testing factor models.

3.2.3 Summary

BP's approach is logically inconsistent with a key implicit assumption underlying the use of accounting-based proxies and it follows that their tests cannot generate meaningful inferences about the reliability of these proxies. Furthermore, because some of the factors that BP consider have little or no empirical or theoretical support, the potential for spurious inferences is considerable. This is not an idle concern. For example, consider the study by McInnis (2010). He shows that past evidence of a positive relation between earnings volatility and an estimate of the implied expected rate of return derived from Value Line data (i.e., the estimate BP refer to as r_{DIV}) is spurious. He demonstrates that earnings volatility and analyst optimism about long-term earnings growth are positively

[6] BPW conduct similar tests as BP; however, BPW also evaluate the: (1) risk-free rate, rf, (2) log of equity market value, LMKVL; and, (3) log of the book-to-price ratio, LBP. Including rf as an independent variable is unorthodox given that BPW claim to estimate cross-sectional regressions and for a true cross-section—i.e., a set of observations that are temporally aligned—there is no variation in the risk-free rate. BPW avoid this issue by estimating separate regressions for each year in which they pool observations from different months. Nonetheless, eight of the thirteen accounting-based proxies that BPW evaluate do not have a statistically significant association with rf. Regarding LMKVL and LBP, we know of no *equilibrium* model of agents' risk-return tradeoffs that implies that these two characteristics are risk factors. Hence, as discussed in section1.2, the interpretation of these two variables is unclear.

related. However, r_{DIV} is increasing in expected long-term earnings growth. Consequently, the positive relation between r_{DIV} and earnings volatility is mechanical and it does not imply that investors demand higher compensation for holding stocks with higher earnings volatility.

3.3. Discussion of Easton and Monahan (2005)

In this section we discuss the two-step approach developed by EM; and, we articulate some frequently asked questions about each step.

3.3.1 EM's First Step

In the first step of their analyses EM estimate regressions of realized returns on accounting-based estimates of the ERR and news proxies. A potential concern with this approach is that it appears logically inconsistent with IA1. This concern is unwarranted. Rather, EM developed their research design with the express purpose of dealing with the implications of IA1 head-on. In particular, EM develop measures of the information shocks (i.e., news proxies) that cause realized returns to differ from expected returns and they include these news proxies in their regressions as control variables.

It is important to note that EM's approach is not *ad hoc*. Rather, it is motivated by analytical results presented in Vuolteenaho (2002) who demonstrates that realized return can be decomposed in the following manner:

$$\begin{aligned} r_{i,t} &\approx E_{t-1}\left[r_{i,t}\right]+\left[\sum_{j=0}^{\infty}\rho^{j}\times\Delta E_{t}\left[roe_{i,t+j}\right]\right]-\left[\sum_{j=1}^{\infty}\rho^{j}\times\Delta E_{t}\left[r_{i,t+j}\right]\right] \\ &= E_{t-1}\left[r_{i,t}\right]+CN_{i,t}-RN_{i,t} \end{aligned} \tag{1}$$

In equation (1): $r_{i,t}$ is the natural log of one plus stock return for firm i at time t; $E_t[\cdot]$ is the expectation operator conditional on information available at time t; ρ is a positive number that is slightly smaller than one; $\Delta E_t[\cdot]$ equals $(E_t[\cdot] - E_{t-1}[\cdot])$; and, $roe_{i,t}$ is the natural log of one plus time t accounting return on equity for firm i. $CN_{i,t}$ and $RN_{i,t}$ are referred to as cash flow news and return news.

The interpretation of equation (1) is straightforward: realized return and expected return are equal when investors do not revise their expectations about future earnings or future discount rates. However, if investors' expectations change, realized and expected return are not equal. If investors become more optimistic (pessimistic) about future cash flows, time t realized return will be greater (less) than expected *ceteris paribus*. On the other hand, if future discount rates are revised upwards (downwards), time t realized return will be lower (higher) than expected *ceteris paribus*. These results follow directly from a present value model that Vuolteenaho (2002) derives from two tautologies.

EM exploit the fact that, as shown in equation (1), the coefficients on true expected return (i.e., $E_{t-1}[r_{i,t}]$), true cash flow news (i.e., $CN_{i,t}$), and the product of negative one and true return news (i.e., $-1 \times RN_{i,t}$) are all equal to one. Hence, for each accounting-based estimate that they evaluate, EM estimate the regression shown in equation (2) below and they compare the estimated coefficient on each accounting-based estimate.

$$r_{i,t} = \alpha_0 + \alpha_1 \times ERR\_P_{i,t} + \alpha_2 \times CN\_P_{i,t} + \alpha_3 \times \left(-1 \times RN\_P_{i,t}\right) + \varepsilon_{i,t} \qquad (2)$$

In equation (2): $ERR\_P_{i,t}$ is an accounting-based estimate of the expected rate of return; $CN\_P_{i,t}$ is a cash flow news proxy; $RN\_P_{i,t}$ is a return news proxy; α_0 through α_3 are estimated regression coefficients; and, $\varepsilon_{i,t}$ is an error term. $ERR\_P_{i,t}$ is calculated using data available at time $t-1$ whereas the news proxies are based on data available at time t. The reason for this is that $ERR\_P_{i,t}$ represents the time $t-1$ expectation whereas the news proxies relate to changes in expectations occurring during time t.

It is important to note that estimates of α_1 taken from equation (2) are affected by measurement error in the news proxies. Hence, EM do not base their conclusions about reliability solely on evidence take from the equation (2). Rather, in the second step of their approach EM develop a method that allows them to evaluate accounting-based estimates of ERR even when the news proxies are measured with error. Before discussing this section step, we address a frequently asked question about the first step of EM's approach: "*Why is it necessary to control for news? If the market is efficient, shouldn't the expected value of the news that arrives at time t+1 be zero on average and shouldn't the news arriving at time t+1 be uncorrelated with expectations formed at time t? Hence, isn't the inclusion of* $CN\_P_{i,t}$ *and* $RN\_P_{i,t}$ *unnecessary?*"

To understand why it is necessary to control for news, it is important to note that market efficiency is an *ex ante* concept with respect to information. It implies that the marginal investor is rational and, thus, at time *t*: (1) the expected value of news arriving at time *t+1* is zero and (2) the expected correlation between the news arriving at time *t+1* and expectations formed at time *t* is zero. However, market efficiency does not imply that there is no news or that *ex post* there is no correlation between the news arriving at time *t+1* and expectations formed at time *t*. In other words, market efficiency does not imply that the marginal investor is clairvoyant.

This argument for the inclusion of the news proxies often leads to a follow-up question: "*True, but for large panels of data, isn't the average value of the news equal to zero?*" The empirical evidence suggests that the answer to this question is, again, no. There is mounting evidence that, even with large panels of historical data, information shocks do not cancel out across sample observations. Furthermore, if the average news is zero for a particular sample, average realized returns are an acceptable proxy for expected returns and, thus, the reliability of accounting-based estimates of the implied expected rate of return is a moot point.

In addition, and more importantly, the evidence suggests that the average correlation between time *t+1* information shocks and time *t* expectations is also non-zero. For example, Fama and French (2002) provide evidence that persistent downward revisions in the expected market risk premium (i.e., discount rate shocks) occurred during the post-war era; and, this phenomenon caused the contemporaneous realized equity premium to exceed expectations (i.e., lower expected rates of return imply higher prices and, consequently, higher realized rates of return). These discount rate shocks did not affect all stocks equally. Rather, stocks with high loadings on the market risk factor exhibited both: (1) higher expected returns at time *t* and (2) the largest reaction to the discount rate shock occurring at time *t+1*. This implies a negative correlation between return news and expected returns. Hence, to avoid drawing spurious inferences attributable to correlated omitted variables bias, EM include a return news proxy in their regressions.

Second, it is also important to note that the return decomposition developed by Vuolteenaho (2002) and used by EM is based on two tautologies. This implies that EM do not assume, and do not need to assume, market efficiency. Unfortunately, there appears to be some confusion in the literature about this fact. For example, Lee (2010) writes the following on p. 746 of his review of Easton (2007).

> "*In the Vuolteenaho (2002) framework, which was adopted by Easton and Monahan (2005), stock returns are decomposed into innovations in cash flows or discount rates. But what if a substantial portion of each period's returns is due to ''exogenous liquidity shocks'' (or in the vernacular of behavioral finance,*

''changes in investor sentiment'') that represents neither cash flow news nor discount rate news? I think it is useful to consider a setting in which noise in price plays a more prominent role. In such a setting, the Easton and Monahan (2005) approach might not reduce measurement errors appreciably. Indeed, we would need to think more carefully about the proper benchmarks for evaluating the quality of ICC estimates."

The correct reply to Lee's comment is straightforward: Vuolteenaho's (2002) model does allow for "noise" in prices. Specifically, to derive the return decomposition, Vuolteenaho (2002) makes no assumptions about the manner in which investors form expectations, the nature of the information available, or the underlying market clearing process. Rather, the decomposition holds regardless of whether investors: (1) experience irrational mood swings in which they go from being wildly optimistic to being hopeless pessimistic; (2) throw caution to the wind on one day and scorn all types of risk the next; and/or (3) exhibit blissful ignorance on some days and are hyper vigilant on others; etc.

3.3.2 EM's Second Step

As discussed above, EM's approach is a logical extension of the implicit assumptions that motivate the use of accounting-based *ERR* proxies. In particular, EM model the news components that cause realized returns to differ from expected returns. Hence, their approach is designed with the express purpose of dealing with IA1. Moreover, their approach is based on analytical results that are derived from tautologies. Consequently, users of EM's approach are not put in the untenable position of having to defend *ad hoc* factors or unproven theories.

That said, like all empirical approaches, the first step in EM's approach has limitations. To understand these limitations and the importance of the second step of EM's approach it is important to note that expectations embedded in prices regarding future discount rates and future accounting numbers are unobservable. This implies that all accounting-based estimates of the ERR as well as all cash flow and return news proxies are measured with error. A well-known result in econometrics (e.g., Rao (1973), pp. 280-284 of chapter 9 of Greene (1993)) is that when all of the variables in a multiple regression are measured with error, the estimated coefficients are biased and the sign of the bias is unknown. This is true even if the measurement error in each variable is random (i.e., the measurement error is not correlated with the true values of the remaining variables or their measurement errors).

In light of the effect that measurement error has on the estimates of α_1 taken from equation (2), EM develop a second step in which they compare measurement error variances. To do this, EM rely on another, well-known result in econometrics (e.g., Garber and Klepper (1980) and Barth (1991)). Specifically, when the linear relation between the dependent variable (i.e., realized return) and the true independent variables (i.e., the true ERR, true cash flow news, and true return news) is known, we can infer the variance of the measurement error in each separate proxy variable. This result is quite pertinent in EM's research setting because, per equation (1), the coefficients on the true ERR, true cash flow news, and the product of negative one and true return news are equal to one. This is the motivation for the measurement error analyses, which are central to EM's approach (see pp. 506-507 and Appendix B). Since conversations with numerous colleagues lead us to believe that EM do not describe them well, we elaborate on them. We do this by posing and answering four frequently asked questions.

FAQ 1: why are measurement error ***variances*** pertinent?

Measurement error is less problematic if it is constant across observations. If a particular proxy contains the same amount of measurement error for every sample observation, the proxy variable will be an accurate measure of relative differences.

Moreover, relative differences are often the issue of concern (e.g., estimated regression slope coefficients relate purely to variation across observations). However, if the measurement error varies across observations, the proxy variable will not be a reliable indicator of relative differences; and, as the measurement error variance increases, the reliability of the proxy falls. Hence, measurement error variances are the relevant issue and, thus, EM compare measurement error variances.

FAQ 2: why are some of the modified noise variables estimated by EM negative?

EM use modified noise variables to infer measurement error variances. The estimated values of some of these noise variables are negative, which seems odd given variances cannot be less than zero. However, as shown in equation (5) of EM, these modified noise variables are equal to the measurement error variance less four, unobservable covariances. Hence, depending on the relative values of the measurement error variance and these covariances, the modified noise variable may be negative.

FAQ 3: are EM comparing variances or covariances?

The answer to FAQ 2 often raises a concern that differences in noise variables are attributable to differences in the covariances rather than the measurement error variances. This is unlikely for two reasons. First, two of the covariance terms are only a function of true values and, thus, these covariances do not lead to differences across estimates. Second, the remaining two covariances are a function of the true values of the news proxies and the measurement error in the ERR estimate. While these can vary across estimates of the ERR, it is difficult to believe that: (1) errors in the researcher's ability to measure expectations at time *t* are correlated with revisions in true expectations occurring during time *t+1* and (2) even if this correlation is non-zero, there is no reason to believe its magnitude differs across estimates.

<u>FAQ 4: do the noise variables provide information about reliability on an absolute scale?</u>

No, they do not. The noise variables only serve as relative rankings. However, given that many research questions relate to relative differences, this is not too disconcerting. Moreover, there are ways of ameliorating ambiguity associated with making relative comparisons. For example, if a researcher wants to avoid the problem of "picking the best of a bad lot" he can compare his estimate of the expected rate of return to one (or more) "straw men." For instance, EM use r_{pe}, which is based on restrictive assumptions about future earnings growth, as a straw man.

To summarize. The first step of EM's approach has limitations. These limitations are attributable to the fact that all of the proxy variables included in EM's regressions, which are shown in equation (2), are measured with error. In the second step of their approach EM circumvent these limitations by comparing measurement error variances. These comparisons allow EM to rank accounting-based estimates of the ERR in terms of their relative reliability: for a particular sample of firms, the most reliable proxy is the one with the lowest measurement error variance.

4. Criticisms Made by Botosan, Plumlee and Wen (2011) of Easton and Monahan (2005)

Before responding to BPW's criticisms, it is important to note that BPW do not criticize the use of equation (2) *per se*. Rather, they take issue with the news proxies used by EM and they are especially critical of EM's return news proxy (i.e., $RN\_P_{i,t}$). Hence, we begin by elaborating on EM's return news proxy and then we explain and respond to BPW's specific criticisms. We do not elaborate further on the second step of EM's approach because it is neither mentioned nor criticised by BPW.

4.1 EM's Return News Proxy

EM measure return news in the following manner (see pp. 512-513 of EM):

$$RN\_P_{i,t} = \frac{\rho}{1-\rho} \times \left(ERR\_P_{i,t+1} - ERR\_P_{i,t} \right) \tag{3}$$

Hence, EM's time *t* return news proxy is a function of the time *t+1* change in the accounting-based estimate of the ERR. This implies that there is a different return news measure for each accounting-based estimate, which makes sense: the same phenomena that determine risk levels also determine risk changes (i.e., levels and changes are inextricably linked).

In addition to being intuitive, EM's return news proxy follows directly from equation (1) and the nature of the accounting-based valuation models underlying the estimates of the ERR evaluated by EM (and BPW). To illustrate why this is true we state three facts. To our knowledge these facts are not in dispute.

<u>Fact 1: return news is a function of the change in the expected discount rate</u>

As shown in equation (1) above, $RN_{i,t}$ is a function of the difference between expectations formed at time *t* and expectations formed at time *t-1 (i.e.,* $\Delta E_t[\cdot]$

$= (E_t[\cdot] - E_{t-1}[\cdot]))$. Hence, $RN_{i,t}$ is a function of $(E_t[r_{i,t+1}] - E_{t-1}[r_{i,t+1}])$. However, EM's proxy relates to $(E_t[r_{i,t+1}] - E_{t-1}[r_{i,t}])$ not $(E_t[r_{i,t+1}] - E_{t-1}[r_{i,t+1}])$. In words, EM do **not** compare the time t expected return for year t+1 to the time t-1 expected return for **year t+1**. Rather, they compare the time t expected return for year t+1 to the time t-1 expected return for **year t**. This appears to be a mistake. However, it is not a mistake because of fact two.

Fact 2: for the accounting-based estimates evaluated by EM, $E_{t-1}[r_{i,t+1}] = E_{t-1}[r_{i,t}]$

The reason for this is that the accounting-based estimates of the ERR evaluated by EM (and BPW) are equivalent to internal rates of return. Consequently, EM (and BPW) are implicitly assuming that the expected rate of return is constant over the forecast horizon (i.e., $E_{t-1}[r_{i,t+j}] = E_{t-1}[r_{i,t}]$ for all j). Hence, the fact that EM use $(E_t[r_{i,t+1}] - E_{t-1}[r_{i,t}])$ instead of $(E_t[r_{i,t+1}] - E_{t-1}[r_{i,t+1}])$ is correct because, for the accounting-based estimates that EM (and BPW) evaluate, these two expressions are equivalent.

It is important to note that fact 2 does not imply that the expectation of $r_{i,t+j}$ formed in year $t-1$ equals the expectation of $r_{i,t+j}$ formed in year t (i.e., $E_{t-1}[r_{i,t+j}] \neq E_t[r_{i,t+j}]$). Investors can revise their expectations (e.g., they may become more risk averse, they can decide the firm has become riskier, etc.) but when they do they are assumed to revise the discount rate used for each period in the forecast horizon by the same amount. This leads to fact three:

Fact 3: for the accounting-based estimates evaluated by EM, $\Delta E_t[r_{i,t+1}] = \Delta E_t[r_{i,t+j}]$ for all j

This is equivalent to saying that the discount rate follows a random walk or that changes in the discount rate are permanent. When we combine fact three with facts one and two, we obtain at the following set of equalities:

$$
\begin{aligned}
RN_{i,t} &= \left[\sum_{j=1}^{\infty} \rho^{j} \times \Delta E_{t}\left[r_{i,t+1}\right]\right] = \left[\sum_{j=1}^{\infty} \rho^{j} \times \left(E_{t}\left[r_{i,t+1}\right] - E_{t-1}\left[r_{i,t}\right]\right)\right] \\
&= \rho \times \left(E_{t}\left[r_{i,t+1}\right] - E_{t-1}\left[r_{i,t}\right]\right) \times \left(1 + \sum_{j=1}^{\infty} \rho^{j}\right) = \frac{\rho}{1-\rho} \times \left(E_{t}\left[r_{i,t+1}\right] - E_{t-1}\left[r_{i,t}\right]\right)
\end{aligned} \tag{4}
$$

Ergo, EM's return news proxy.

The above is compelling. A critic arguing against EM's return news proxy must explain the problem with using a proxy that follows directly from equation (1), which is tautological, and the properties of the accounting-based estimates evaluated by EM. Colloquially speaking, the critic must argue with the math. In addition, the critic must derive a suitable substitute proxy that is not *ad hoc*.[7] Again, colloquially speaking, it takes a model to beat a model.

Second, we are not arguing that the manner in which EM measure return news is correct for all accounting-based estimates of the ERR. Fact 2 is true for all of the

[7] BPW suggest that researchers control for return news by including in equation (2) the contemporaneous change in: (1) the risk-free rate and (2) a firm-year-specific estimate of CAPM beta. We have a number of concerns about this approach. Including the change in the risk-free rate in equation (2) is odd for two reasons. First, by definition, the risk-free rate has nothing to do with risk. However, most of the researchers we are familiar with use accounting-based estimates of the implied expected rate of return to evaluate whether a particular phenomenon (e.g., disclosure quality) is a priced risk factor. Second, the change in the risk-free rate is a cross-sectional constant and the relation between realized returns and the change in the risk-free rate is constant (i.e., it is not a function of the factor loadings). Hence, a straightforward way of controlling for changes in the risk-free rate is to estimate *true* cross-sectional regressions and exclude the change in the risk-free rate from the model. Suggesting the use of the change in CAPM beta is also odd given that it requires BPW to make implicit assumptions that are dubious and inconsistent with some of their other assumptions. First, BPW are implicitly assuming that the return on the market portfolio is the only priced risk factor. There is, however, an ongoing debate regarding the nature of the "true" factor model. Moreover, the assumption that market risk is the only relevant factor is clearly inconsistent with other assumptions made by BPW. In particular, on p. 1088, BPW rely on Ross' (1976) arbitrage pricing theory to motivate use of other risk factors. Second, BPW are implicitly assuming that they can develop reliable, firm-year-specific measures of beta. Extant evidence suggests, however, that this is not possible. Third, BPW are implicitly assuming that market participants never revise their expectations of the equity premium. This is a strong assumption; and, even though it is a cross-sectional constant, the change in the expected equity premium leads to cross-sectional variation in realized returns. This is attributable to the fact that the relation between realized stock return and the change in the expected equity premium is a function of the firm-specific factor loading on the expected equity premium. Finally, BPW are implicitly assuming that accounting-based proxies are irrelevant, which is inconsistent with the basic motivation for their study. If the CAPM is descriptive and beta can be measured well, the reliability of accounting-based proxies is a moot issue. Rather, we can simply use estimates based on the CAPM.

accounting-based models analysed in BP, BPW, EM and most extant studies.[8] However, in a more general model, the discount rate may vary over the forecast horizon. Hence, if an empirical technique for imputing discount rates that vary over the forecast horizon is developed, EM's return news proxy will have to be modified. This does not imply that EM's approach is flawed. It is the correct approach for the accounting-based estimates they study.

4.2 Botosan, Plumlee and Wen's (2011) Criticisms of Easton and Monahan's (2005) Return News Proxy

Notwithstanding the compelling nature of the discussion above, we respond to BPW's specific concerns so that we may further clarify the issues and let the reader decide. In order to create a basis for discussion, we provide an excerpt from BPW. Please note that we modify their text in three ways. First, we substitute our notation for BPW's notation. We do this to avoid confusion. Second, we use the original equation numbers from BPW; however, to avoid confusion, we precede each equation number with the letters BPW—e.g., we refer to equation (6) of BPW as BPW6. Finally, we use bold font to highlight certain passages or equation numbers. We do this so that we can refer to these passages in our response—i.e., "regarding the second highlighted passage…" With these clarifications in mind, we restate the relevant passage of text, which is taken from pages 1116-1117 of BPW.

> *ERRs vary across approaches as different cash flow, CF, assumptions arise from different terminal-value assumptions. Nevertheless, by construction, all ERR ~ f(CF,P), and therefore, all ΔERR ~ f(ΔCF,ΔP).*

[8] Claus and Thomas (2001) is the exception that proves the rule in the sense that, while they allow the risk-free rate to vary over the forecast horizon, they maintain the assumption that equity premium is constant over the forecast horizon.

The theoretical specification of the realized return model (i.e., equation (2)) is shown below for convenience.

$$r_{i,t} = E_{t-1}[r_{i,t}] + CN_{i,t} - RN_{i,t} \qquad \textit{(BPW6)}$$

Empirically, $r_{i,t}$ ~ f(ΔP) and $CN_{i,t}$ ~ f(ΔCF). In EM's empirical specification $RN_{i,t}$ = ΔERR ~ f(ΔCF,ΔP). Consequently, the model EM estimate can be described by the following set of relationships:

$$f(\Delta P) = E_{t-1}[r_{i,t}] + f(\Delta CF) - f(\Delta CF, \Delta P) \qquad \textit{(BPW7)}$$

EM's proxy for expected return news (ΔERR) is by construction a function of ΔCF and ΔP, which are also included in the model as dependent and explanatory variables, respectively. Stated another way, solving (7) for $E_{t-1}[r_{i,t}]$ yields:

$$E_{t-1}[r_{i,t}] = f(\Delta CF) - f(\Delta CF) + f(\Delta P) - f(\Delta P) \qquad \textbf{\textit{(BPW8)}}$$

The right hand side of (BPW8) implies a product that is close to zero. *Expected return is not likely to explain realized returns under this empirical specification. Thus, while it is theoretically defensible to use the change in true $E_{t-1}[r_{i,t}]$ to capture expected return news, it is empirically problematic to use the change in an $E_{t-1}[r_{i,t}]$ proxy measured via an implied cost of capital approach for this purpose.* ***The resulting provoked circularity in the empirical model provides no role for $E_{t-1}[r_{i,t}]$ to contribute to the explanation of $r_{i,t}$, and as a result, any ICC estimate included in the model to proxy for $E_{t-1}[r_{i,t}]$ will be statistically insignificant, regardless of the validity, or lack thereof, of the ERR estimate employed.***

Frankly, it is not exactly clear to us what BPW are concerned about. Are they arguing that there is a mechanical relation between EM's return news proxy and the dependent variable; consequently, the remaining variables in the regression will have no explanatory power? Are they arguing that there is severe multicollinearity? Are they concerned that EM misinterpret the coefficient on the ERR proxy because the ERR proxy is also a component of the return news proxy? Is it some combination of these issues? Given the ambiguity, we suggest several different interpretations of BPW's statements, and then we explain why each of these are misplaced—i.e., there is no problem with the return news proxies used by EM.

4.2.1 Interpretation 1: Mechanical relation between the $R\_NEWS_{i,t}$ and $r_{i,t}$

In the first passage that we highlight BPW state "*EM's proxy for expected return news (ΔERR) is by construction a function of ΔCF and ΔP, which are also included in the model as dependent and explanatory variables, respectively*. *Stated another way, solving (7) for $E_{t-1}[r_{i,t}]$ yields:* $E_{t-1}\left[r_{i,t}\right] = f(\Delta CF) - f(\Delta CF) + f(\Delta P) - f(\Delta P)$ *(BPW8).*"

One interpretation of this passage is that BPW are concerned that there is a mechanical relation between EM's return news proxies and realized return, which is the dependent variable in (2). This, in turn, implies that the remaining regressors will have no relation with realized return.

Is the above concern valid? The short answer is no. There is no mechanical relation. Rather, EM's return news proxy measures the extent to which the valuation numerator (e.g., expected earnings) grew at a different rate than price. If expected earnings grew faster (slower) than price, $RN_{i,t}$ is positive (negative). This makes perfect sense. If investors become more optimistic (pessimistic) about future earnings but price decreases (increases), investors must be discounting future earnings at a higher (lower) rate - i.e., the discount rate must have increased (decreased).

To clarify this point, we assume, in the interest of simplicity, that the researcher is using an accounting-based model in which price equals expected forward earnings-per-share divided by the expected cost of capital. That is, $P_{t-1} = E_{t-1}[eps_t]/E_{t-1}[r_t]$, which implies $ERR\_P_t = E_{t-1}[eps_t]/P_{t-1}$. We do this for purposes of exposition but without loss of generality.

Recall that equation (1) relates to logged variables; hence, the ERR estimate based on price to expected forward earnings is defined as follows.

$$\begin{aligned} ERR\_P_t &= \ln\left(1 + \frac{E_{t-1}[eps_t]}{P_{t-1}}\right) = \ln\left(\frac{P_{t-1} + E_{t-1}[eps_t]}{P_{t-1}}\right) \\ &= \ln(P_{t-1} + E_{t-1}[eps_t]) - \ln(P_{t-1}) \end{aligned} \tag{5}$$

Combining equation (5) and equation (4) and ignoring the capitalization factor (i.e., $\rho/(1-\rho)$), we obtain the following return news proxy.

$$\begin{aligned} RN\_P_t &= ERR\_P_{t+1} - ERR\_P_t \\ &= \{\ln(P_t + E_t[eps_{t+1}]) - \ln(P_t)\} - \{\ln(P_{t-1} + E_{t-1}[eps_t]) - \ln(P_{t-1})\} \\ &= \{\ln(P_t + E_t[eps_{t+1}]) - \ln(P_{t-1} + E_{t-1}[eps_t])\} - \{\ln(P_t) - \ln(P_{t-1})\} \\ &= \ln\left(\frac{P_t + E_t[eps_{t+1}]}{P_{t-1} + E_{t-1}[eps_t]}\right) - \ln\left(\frac{P_t}{P_{t-1}}\right) \end{aligned} \tag{6}$$

Hence, the return news proxy equals the difference between two (continuously compounded) growth rates: (1) the growth rate in $(P_{t-1}+E_{t-1}(eps_t))$ and (2) the growth rate in P_{t-1}.

Equation (6) implies that if expected earnings grew at the same rate as price, the expected discount rate did not change (i.e., $RN_{i,t} = 0$). However, if expected earnings grew faster than price, the expected discount rate must have risen (i.e., $RN_{i,t} > 0$). On the other hand, if expected earnings grew slower than price, the expected discount rate must have fallen (i.e., $RN_{i,t} < 0$). It follows that the return news proxy captures the portion of the unexpected price change that is not attributable to changes in expectations about future earnings. This makes perfect sense: price in this model is a function of expected earnings and the expected discount rate and it follows that unexpected price changes are a function of changes in expectations about future earnings and future discount rates.

On inspection of equation (6) the question may come to mind: "*If you look at (6) you see that $ln(P_t/P_{t-1})$, which is essentially realized return at time t, shows up in the equation. Doesn't this lead to a mechanical bias?*" Again, the answer is no. Further inspection of equation (6) reveals that price-growth is, essentially, added and subtracted. What drives the equation is the extent to which earnings growth differs from price growth.

Regarding a potential mechanical relation, BPW state that "*[t]he right hand side of (BPW8) implies a product that is close to zero.*" This statement implies that $\{r_{i,t}-(CF_{i,t}-RN_{i,t})\}$, which we refer to as $RET\_LESS\_NEWS_{i,t}$, is approximately equal to zero. Hence, after controlling for the news proxies, there remains no variation in $r_{i,t}$ to explain. Consequently, the ERR proxy cannot have any explanatory power. Is this a valid point? Again, the short answer is no. The logic underlying our answer is provided below.

Equation (BPW8) does *not* follow from equation (BPW7). In particular, $RN_{i,t} \sim f(\Delta CF, \Delta P) \neq f(\Delta CF) - f(\Delta P)$. Rather, as shown in equation (6), $RN_{i,t}$ is a nonlinear function of ΔCF and ΔP. Hence, the conclusion that, after controlling for return news, $RET\_LESS\_NEWS_{i,t}$ is mathematically equal (or approximately equal) to zero is incorrect.

A potential rebuttal to the above is that "*Sure, $RET\_LESS\_NEWS_{i,t}$ isn't mathematically equal to zero but it is the empirical properties of $RET\_LESS\_NEWS_{i,t}$ that matter.*" This is a fair comment. However, descriptive statistics in Table 2 of EM show that, depending on the ERR proxy considered, the mean of $\mathrm{RET\_LESS\_NEWS}_{i,t}$ is between 0.046 and 0.176.[9] These are nontrivial amounts given that, as shown in Table 2 of EM, the mean of $r_{i,t}$ is 0.096.

4.2.2 Interpretation 2: Extreme multicollinearity

An alternative interpretation of the first two highlighted passages is that EM's results are attributable to extreme multicollinearity. Specifically, there may be an approximate linear relation between the ERR proxies EM evaluate and the variables EM use to measure news. Consequently, EM's regressions are inefficient and the standard errors are so large that it is impossible to reject the null of no association.

Although it is notoriously difficult to rule out multicollinearity when there are more than two regressors (Kennedy (1992)), we are sceptical that multicollinearity is an issue. We have three reasons. First, as shown in Table 3 of EM, the correlations between the three regressors in equation (2) are not high. Regarding the different ERR proxies and the cash flow news proxy, the correlation with the highest absolute value is 0.148. The highest absolute value of the correlations between the ERR proxies (cash flow news proxy) and the return news proxies is 0.414 (0.126). Moreover, the ERR proxies that have relatively high

---

[9] To make these calculations we refer to the means shown in Table 2 of EM. First, for each ERR proxy we subtract the mean of $\hat{rn}_{i,t}$ from the mean of $\hat{cn}_{i,t}$ to obtain the mean of total news. Next, to obtain the mean of $\mathrm{RET\_LESS\_NEWS}_{i,t}$, we subtract the mean of the total news from $r_{i,t}$. Note that because $\hat{rn}_{i,t}$ varies across ERR proxies, $\mathrm{RET\_LESS\_NEWS}_{i,t}$ also varies across ERR proxies. In particular, The mean of $\mathrm{RET\_LESS\_NEWS}_{i,t}$ for the different ERR proxies are: r_{pe} 0.176, r_{peg} 0.095, r_{mpeg} 0.100, r_{gm} 0.075, $r_{\Delta agr}$ 0.099, r_{ct} 0.046, and r_{gls} 0.129.

correlations with one news proxy do not have relatively high correlations with the other news proxy. For example, the ERR proxy with the highest correlation in absolute value with the return news proxy, $r_{\Delta agr}$: (1) has the third (out of seven) lowest correlation in absolute value with the cash flow news proxy (0.129) and (2) has an associated return news proxy which has the lowest correlation in absolute value with the cash flow news proxy (0.018).

Second, if there is an approximate linear relation between a particular ERR proxy and the cash flow news and return news proxies, each of the news proxies has an approximate linear relation with the ERR proxy and the remaining news proxy. Consequently, all of the estimated coefficients in equation (2) will be insignificant—i.e., they will each be affected by multicollinearity. As shown in Table 4 of EM, this is not true. Rather, all seven of the estimated coefficients on $CN\_P_{i,t}$ and $RN\_P_{i,t}$ have the predicted sign and are significantly different from zero.

Finally, multicollinearity is a data problem and a well-known solution is to obtain more data (e.g., Kennedy (1992)). Consequently, for a particular sample of data, multicollinearity will be more severe for regressions estimated on partitions of the sample because these partitions contain less observations. However, as shown in EM, the opposite is true. In particular, EM partition their sample into thirds, and then estimate equation (2) on each separate partition. As shown in Panel C of Table 9 of EM, the estimated coefficients on two of the ERR proxies that EM consider (r_{mpeg} and r_{gm}) are positive for one of these partitions; and, the estimated coefficient on the ERR proxy r_{ct} is positive for two of these partitions.

4.2.3 Misinterpretation of the Estimated Coefficient on ERR Proxies

A final possibility is that because the return news proxy is a function of the ERR proxy, EM misinterpret the coefficient on $ERR\_P_{i,t}$ in equation (2). Specifically, because equation (2) can be rearranged to arrive at equation (7), which is shown below, the correct test for determining the reliability of a particular expected return proxy is to compare A_1 instead of α_1 to one. This is incorrect, however.

$$\begin{aligned} r_{i,t} &= \alpha_0 + \left(\alpha_1 + \alpha_3 \times \frac{\rho}{1-\rho}\right) \times ERR\_P_{i,t} + \alpha_2 \times CN\_P_{i,t} \\ &\quad + \alpha_3 \times \left(-\frac{\rho}{1-\rho} \times ERR\_P_{i,t+1}\right) + \varepsilon_{i,t} \\ &= \alpha_0 + A_1 \times ERR\_P_{i,t} + \alpha_2 \times CN\_P_{i,t} + \alpha_3 \times \left(-\frac{\rho}{1-\rho} \times ERR\_P_{i,t+1}\right) + \varepsilon_{i,t} \end{aligned} \tag{7}$$

To understand why A_1 should not be compared to one (which is the correct benchmark for α_1 as well as α_2 and α_3) it is important to note that A_1 equals $\{\alpha_1 + \alpha_3 \times \rho/(1-\rho)\}$. Hence, assuming ρ equals 0.95, the correct benchmark for A_1 is $\{1 + 1 \times \rho/(1-\rho)\} = 1/(1-\rho) = 20$.

Another way of saying this is that, if a researcher estimates equation (2), the correct benchmark for α_1 is 1. However, if the researcher wants to fully isolate the relation between $r_{i,t}$ and $ERR\_P_{i,t}$ she can rearrange equation (2) and arrive at equation (7). Doing so is mathematically equivalent to estimating equation (7), not equation (2). Hence, she needs to use the benchmark for A_1 that is implied by equation (7). This benchmark is 20 (assuming $\rho = 0.95$) not one.

In the table below, we show the values of A_1 implied by the estimates of α_1 and α_3 taken from Table 10 of BPW. When solving for A_1 we assume ρ equals 0.95; hence, $A1 = \{\alpha_1 + \alpha_3 \times (0.95/(1-0.95)\}$. All of the estimates of A_1 are much less than 20.

	α_1	α_3	A_1	Implied ρ*
r_{DIV}	-0.29	0.04	0.47	3.91
r_{PEG}	-0.43	0.07	0.90	2.86
r_{MPEG}	-0.31	0.03	0.26	3.85
r_{GM}	-0.39	0.03	0.18	3.31
r_{CT}	0.76	0.11	2.85	-0.37
r_{GLS}	0.46	0.18	3.88	-1.93

There are two potential criticisms of the above. First, the benchmark of 20 is ambiguous because ρ is not known with certainty. One way to determine whether this criticism is valid is to take the values of α_1 and α_3 pertaining to a particular proxy and solve for the value of $\rho*$ that sets $\{1 + 1 \times \rho*/(1-\rho*)\}$ equal to $\{\alpha_1 + \alpha_3 \times \rho*/(1-\rho*)\}$ (i.e., $\rho* = (1-\alpha_1)/(\alpha_3-\alpha_1)$).[10] If the implied $\rho*$ for a particular proxy is plausible, we can argue that the proxy is reliable. However, as shown above, all of the implied values of $\rho*$ are outside the interval containing zero and one, which is the interval that the true value of ρ must fall within. Hence, all of the implied values of $\rho*$ are implausible.

The second criticism is that A_1 and the implied value of ρ* are functions of both α_1 and α_3. Hence, the fact that they take on implausible values may be attributable to measurement error in the return news proxy not the ERR proxy. For example, BPW's estimate of α_1 on r_{CT} is a statistically significant 0.76. However, the estimate of α_3 taken

[10] This approach has a clear limitation: as α_1 and α_3 approach one, ρ* approaches $\pm\infty$.

from the same regression is 0.11. This is a fair comment and it further reinforces the problem with using evidence from equation (2) as the sole basis for evaluating the ERR proxies. In particular, as discussed in section three, if one or more of the regressors in equation (2) are measured with error, all the estimated coefficients obtained from equation (2) are biased—i.e., the bias is interdependent. Hence, the second step of EM's approach is key as it allows the researcher to isolate and evaluate the measurement error in the ERR proxies.

4.2.4 Additional Comments

Finally, it is important to note that the empirical evidence in EM (and BPW) is also inconsistent with BPW's argument. If there is a mechanical relation between $RN_{i,t}$ and $r_{i,t}$, we should observe three empirical results. First, as BPW point out in the second passage of text that we highlight: "*The resulting provoked circularity in the empirical model provides no role for $E_{t-1}[r_{i,t}]$ to contribute to the explanation of $r_{i,t}$, and as a result, any ICC estimate included in the model to proxy for $E_{t-1}[r_{i,t}]$ will be statistically insignificant, regardless of the validity, or lack thereof, of the ERR estimate employed.*" We agree with this statement in the sense that if there is a mechanical relation, it should be ever-present. Consequently, neither EM nor BPW should ever document a positive relation between realized return and any ERR proxy after controlling for news in the manner prescribed by EM. They do, however. As discussed above, in Panel C of their Table 9, EM show that two (i.e., r_{MPEG} and r_{GM}) of the accounting-based estimates they evaluate are reliable for one-third of the sample and one (i.e., r_{CT}) is reliable for two-thirds of the sample. Moreover, as discussed above, in their Table 10, BPW document a significant, positive relation between realized return and r_{CT} even after they use the news proxies suggested by EM.

Second, a mechanical relation between $RN_{i,t}$ and $r_{i,t}$ (severe multicollinearity between the regressors) should also destroy the relation between $CN\_P_{i,t}$ (both news proxies) and $r_{i,t}$. This does not happen, however. Rather, as shown in Table 4 of EM, there is a consistent, positive, statistically significant relation between EM's news proxies and realized returns. Finally, if there is a mechanical relation between $RN\_P_{i,t}$ and $r_{i,t}$, the r-squares taken from EM's regressions should be high. They are not. As shown in Table 4 of EM, the highest r-square is 0.30, which is much lower than what we expect to observe if one of the independent variables is "*by construction*" a function of the dependent variable.

4.3 Summary

EM's approach is a logical extension of the implicit assumptions that motivate the use of accounting-based ERR proxies. In particular, EM model the news components that cause realized returns to differ from expected returns; hence, their approach is designed with the express purpose of dealing with IA1. Moreover, their approach is based on analytical results that are derived from tautologies. Hence, users of EM's approach are not put in the untenable position of having to defend ad hoc factors or unproven theories. Finally, the news proxies EM use follow directly from the underlying analytical model and the properties of the accounting-based proxies EM evaluate; hence, criticisms made by BPW are baseless.

5. Conclusion

The expected rate of return on equity capital is a key construct. It is, however, unobservable. Hence, practitioners, policy-makers, and academics often use accounting-based proxies. This choice is made because: (1) the true factors that drive expected returns are unknown and/or cannot be measured reliably and (2) realized returns are biased and noisy. As a result, there has been considerable effort focussed on developing accounting based estimates of the expected rate of return and, more recently, on improving the forecasts of accounting earnings on which these estimates are based.

While accounting-based estimates of the expected rate of return are potentially useful, users cannot simply assume they are reliable. Rather, before an estimate is used its reliability/validity must be evaluated. The results of this evaluation will be more persuasive when the methodology is logically consistent with the reasons underlying the use of the accounting-based estimates. EM develops such a methodology. First, they base their analyses on a rigorous analytical model of the bias and noise in realized returns. Second, their news proxies follow directly from this model and the nature of the accounting-based estimates they evaluate. Finally, they exploit the properties of the analytical model to derive an econometric approach for comparing the measurement error variances of different accounting-based estimates.

References

Banz, R. 1981. “The relation between return and the market value of common stocks.”*Journal of Financial Economics* 9: 3-18.

Barth, M. 1991. “Relative measurement errors among alternative pension asset and liability measures.” *The Accounting Review* 66: 433-463.

Black, F., M. Jensen, and M. Scholes. 1972. “The Capital Asset Pricing Model: Some empirical tests.” *Studies in the Theory of Capital Markets*. Praeger.

Botosan, C. and M. Plumlee. 2005. “Assessing alternative proxies for the expected risk premium.” *The Accounting Review* 80: 21-53.

Botosan, C., M. Plumlee, and J. Wen. 2011. "The Relation Between Expected Returns, Realized Returns, and Firm Risk Characteristics." *Contemporary Accounting Research.* Vol 28, No 4 (Winter 2011) pp. 1085-1122.

Campbell, C., A. Lo and A. McKinley. 1997. “The econometrics of financial markets.” Princeton University Press.

Carhart, M. 1997. “On Persistence in Mutual Fund Performance” *Journal of Finance* 52: 57-82.

Claus, J. and J. Thomas. 2001. “Equity risk premium as low as three percent? Evidence from analysts’ earnings forecasts for domestic and international stocks.” *Journal of Finance* 56: 1629-1666.

Cochrane, J. 2001. “Asset pricing.” Princeton University Press.

Core, J., W. Guay, and R. Verdi. 2008. "Is accruals quality a priced risk factor?" *Journal of Accounting and Economics* 46: 2-22."

Eaton, P. 2004. “PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital.” The Accounting Review 79:1 (2004) 73-95.

Easton, P. 2007. "Estimating the cost of capital implied by market prices and accounting data." *Foundations and Trends® in Accounting*. Now Publishers Inc.

Easton, P. and S. Monahan. 2005. "An evaluating accounting-based measures of expected returns." *The Accounting Review* 80: 501-538.

Elton, J. 1999. “Expected return, realized return, and asset pricing tests.” *Journal of Finance* 54: 1199-1220.

Fama, E. and K. French. 1992. “The cross-section of expected returns.” *Journal of Finance* 47: 427-465.

Fama, E. and K. French. 1993. “Common risk factors in the returns on stocks and bonds.” *Journal of Financial Economics* 33: 3-56.

Fama, E. and K. French. 1997. “Industry costs of equity.” *Journal of Financial Economics* 43: 154-194.

Fama, E. and K. French. 2002. "The equity premium." *Journal of Finance* 57: 637–659.

Fama, E. and J. MacBeth. 1973. “Risk return and equilibrium: Empirical tests.” *Journal of Political Economy* 81: 607-636.

Francis, J., R. LaFond, P. Olsson, and K. Schipper. 2004. “Costs of capital and earnings attributes.” *The Accounting Review* 79: 967-1010.

Garber, S. and S. Klepper. 1980. “’Administrative pricing’ or competition coupled with errors of measurement?” *International Economic Review* 21: 413-435.

Gebhardt, W., C. Lee, and B. Swaminathan, 2001. “Toward an Implied Cost of Capital.” Journal of Accounting Research 39:1 (2001) 135-176.

Gerakos, J. and R. Gramacy, 2013. “Regression-based Earnings Forecasts.” Chicago Booth Research Paper No. 12-26.

Gode, D., P. Mohanram, 2003. “Inferring the Cost of Capital Using the Ohlson-Juettner Model.” *Review of Accounting Studies* 8 (2003) 399-431.

Greene, W. 1993. "Econometric analysis." Prentice Hall Inc.

Hou,K., M. van Dijk and Y. Zhang, 2012. “The implied cost of capital: A new approach.” *Journal of Accounting and Economics* 53 (2012) 504-526.

Jegadeesh, N. and S. Titman. 1993. “Returns to buying winners and selling losers: Implications for stock market efficiency.” *Journal of Finance* 48: 65-91.

Lakonishok, J., A. Shleifer, and R. Vishny. 1992. “Contrarian investment extrapolation and risk.” *Journal of Finance* 49: 1541-1578.

LaPorta, R. 1996. “Expectations and the Cross-section of Stock Returns.” *Journal of Finance* 51: 1715-1742.

LaPorta, R., J. Lakonishok, A. Shleifer, and R. Vishny. 1997. “Good News for Value Stocks: Further evidence on market efficiency.” *Journal of Finance* 52: 859- 874.

Larocque, S. 2013. “Analysts’ Earnings Forecast Errors and Cost of Equity Capital Estimates.” Review of Accounting Studies (2013) 18:135-166.

Lee, C. 2010. "Book Review of 'Estimating the cost of capital implied by market prices and accounting data'." *The Accounting Review* 85: 745-748.

Lewellen, J., S. Nagel, and J. Shanken. 2010. "A skeptical appraisal of asset pricing tests." *Journal of Financial Economics*: Forthcoming.

McInnis, J. 2010. "Earnings smoothness, average returns, and implied cost of equity capital." *The Accounting Review*: Forthcoming."

Mohanram, P. and D. Gode, 2013. “Removing predictable analyst forecast errors to improve implied cost of equity estimates.” *Review of Accounting Studies* (2013) 18:443-478.

Mohanram, P. and K. Li, 2014. “Evaluating Cross-Sectional Forecasting Models for Implied Cost of Capital.” *Review of Accounting Studies* (2014) 19:1152-1185.

Piotroski, J. 2000. "Value investing: The use of historical financial information to separate winners from losers." *Journal of Accounting Research* 38: 1-41."

Rao, P. 1973. “Some notes on the errors-in-variables model.” *The American Statistician* 7: 217-218.

Rosenberg, B., K. Reid, and R. Lanstein. 1985. "Persuasive evidence of market inefficiency." *Journal of Portfolio Management* 11: 9-17.

Ross, S., 1976. “The arbitrage theory of capital asset pricing.” *Journal of Economic Theory* 13 (3):341.60.

Vuolteenaho, T. 2002. “What drives firm-level stock returns?” *Journal of Finance* 57: 233-264.

Effect of analysts' optimism on estimates of the expected rate of return implied by earnings forecasts

Peter D. Easton
University of Notre Dame

and

Gregory A. Sommers
Southern Methodist University

August 2006

The comments of Ashiq Ali, Robert Battalio, Sung Chung, Somnath Das, Gus DeFranco, John Lyon, Hai Lu, Paul Healy, Rick Mendenhall, Krishna Palepu, Gord Richardson, Scott Richardson, Steven Rock, Cathy Schrand, Lisa Sedor, Margaret Shackell-Dowel, Pervin Shroff, Philip Stocken, Phil Shane, Tom Stober, Rex Thompson, Jenny Tucker, Kent Womack, Tzachi Zach, Paul Zarowin, and workshop participants at the 2006 American Accounting Association annual meeting in Washington, DC, Dartmouth College, Drexel University, the 2006 London Business School Accounting Symposium, the 2006 Lone Star Accounting Research Conference, Harvard University, New York University, Pennsylvania State University, Southern Methodist University, Tilburg University, the University of Colorado, the University of Illinois, the University of Melbourne, the University of Minnesota, the University of Notre Dame, and the University of Toronto are greatly appreciated. The paper reflects many long conversations with Mark Zmijewski. We thank Lorie Marsh for her assistance with the preparation of this paper.

Abstract

Recent literature has used analysts' earnings forecasts, which are known to be optimistic, to estimate expected rates of return; yielding upwardly biased estimates. We find a bias of 2.84 percent computed as the difference between the estimates of the expected rate of return based on analysts' earnings forecasts and estimates based on current earnings realizations. The importance of this bias is illustrated by the fact that studies using the biased estimates of the expected rate of return suggest an equity premium in the vicinity of 3 percent. Further analyses show that use of value-weighted, rather than equally-weighted, estimates reduces the bias and yields more reasonable estimates of the equity premium. We also show that analysts recommend "buy" ("sell") when they expect the future return to be high (low) regardless of market expectations and that bias is present for all recommendation types.

1. Introduction

A large and expanding body of literature uses analysts' forecasts of earnings to determine the expected rate of return implied by these forecasts, current book values, and current prices. These implied expected rates of return are often used as estimates of the market's expected rate of return and/or as estimates of the cost of capital.[1] Yet the earnings forecasts are optimistic; and they are made by sell-side analysts who are in the business of making buy/hold/sell recommendations which are, presumably, based on the difference between their expectation of the future rate of return and the market expectation of this rate of return. If these earnings forecasts are optimistically biased, the expected rates of return implied by these forecasts will be upward biased. We estimate the extent of this bias.[2]

We show that, consistent with the extant evidence that forecasts (particularly longer-run forecasts) are optimistic, the difference between the expected rate of return implied by analysts' earnings forecasts and the expected rate of return implied by current earnings is statistically and economically significantly positive. In other words, *ceteris paribus*, studies that use the expected rate of return implied by current prices and these forecasts of earnings have estimates of the cost of capital that may be too high.[3]

The extant literature on analysts' optimism/pessimism generally compares forecasts of earnings with realizations of the earnings that are forecasted. This is an ex post measure of optimism and one that pervades the extant literature. Most of our analysis is a comparison of the expected rate of return implied by analysts' earnings forecasts and the expected rate of return

[1] Cost of capital is an equilibrium concept that relies on the no arbitrage assumption. In the absence of arbitrage opportunities, the markets expected rate of return is equal to the cost of capital.
[2] Claus and Thomas (2001) observe that the optimistic bias in analysts' forecasts will bias their estimate of the equity premium upward.
[3] Examples include Gebhardt, Lee, and Swaminathan (2001), Claus and Thomas (2001), and Easton, Taylor, Shroff, and Sougiannis (2002).

implied by current earnings. This is an ex ante measure of optimism/pessimism. We are primarily interested in this ex ante comparison for two reasons. First, our goal is to determine the bias in estimates of expected rates of return implied by analysts' forecasts at the time that these forecasts are made. Second, this comparison provides an indication of optimism/pessimism that is not affected by events that occur between the forecast date and the time of the earnings realization.[4]

All of our analyses are based on two methods for simultaneously estimating the expected rate of return and the expected growth rate for a portfolio/group of stocks. The estimate of the expected growth rate is not important in and of itself in our study; but estimating it simultaneously with the estimation of the expected rate of return avoids the introduction of error which will almost inevitably arise when the expected growth rate is assumed. Any assumed growth rate will almost invariably differ from the growth rate implied by the data.[5]

The method we use for estimating the expected rate of return that is implied by prices and current accounting data is an adaptation of the method that O'Hanlon and Steele (2000) use to estimate the expected market equity premium for the U.K. The method we use for estimating the expected rate of return that is implied by prices, current book values, and forecasts of earnings is an adaptation of the method that Easton, Taylor, Shroff, and Sougiannis (2002) use to estimate the equity premium in the U.S.

Literature that reverse-engineers valuation models to obtain estimates of the expected rate of return on equity investment is very new. These models include the dividend capitalization model in Botosan (1997); the residual income valuation model in O'Hanlon and Steele (2000),

[4] An obvious recent example of such an event is the tragedy of the terrorist attack of September 11, 2001. This event, which was not foreseen by analysts, would almost certainly have made their forecasts overly optimistic with the benefit of hindsight. We will return to this example.
[5] See Easton (2005) for a detailed discussion of this source of error.

Gebhardt, Lee, and Swaminathan (2001), Claus and Thomas (2001), Easton, Taylor, Shroff, and Sougiannis (2002), and Baginski and Wahlen (2003); and the abnormal growth in earnings model in Gode and Mohanram (2003) and Easton (2004). Literature using these estimates to test hypotheses regarding factors that may affect the expected rate of return developed almost simultaneously; for example, see Daske (2006); Dhaliwal, Krull, Li, and Moser (2005); Francis, Khurana, and Periera (2005); Francis, LaFond, Olsson, and Schipper (2004); Hail and Leuz (2006); Hribar and Jenkins (2004); and Lee, Myers, and Swaminathan (1999). This development took place despite the fact that (1) some of these methods were not designed to provide firm-specific estimates; see, in particular, Claus and Thomas (2001), Easton, Taylor, Shroff, and Sougiannis (2002), and Easton (2004); and (2) there is very little evidence regarding the empirical validity of these methods.

The conclusion from the very recent studies that examine the validity of firm-specific estimates of expected rate of return derived from these reverse-engineering exercises (see, Botosan and Plumlee, 2005; Guay, Kothari and Shu, 2005; and Easton and Monahan, 2005), is that these estimates are poor, indeed. None of these studies addressed the issue of the difference between the market expectation of the rate of return, which these studies purport to measure, and rates implied by analysts' forecasts. Nevertheless, it is possible that the difference is a correlated omitted variable, which could affect the results in studies comparing estimates of the implied expected rate of return on equity capital. For example, it is possible that analysts' forecasts for firms under one accounting regime (say, accounting based on international accounting standards) may be more optimistic than analysts' forecasts for firms under a different accounting regime (say, accounting based on domestic standards). These optimistic forecasts will bias the estimate of the expected rate of return upward, potentially leading to the (possibly erroneous) conclusion that the cost of capital is higher for these firms.

In light of analysts' tendency to be optimistic, estimates of the expected rate of return based on analysts' forecasts are likely to be higher than the cost of capital. Williams (2004) makes this point in his discussion of Botosan, Plumlee, and Xie (2004). This effect of analysts' optimism is exacerbated by the fact that all studies using analysts' forecasts to calculate an implied expected rate of return are based on forecasts made well in advance (usually at least a year ahead) of the earnings announcement. These forecasts tend to be much more optimistic than those made closer to the earnings announcement; see Richardson, Teoh, and Wysocki (2004).

All of our analyses are based on I/B/E/S forecasts of earnings and recommendations for the years 1993 to 2004 and actual prices and accounting data for 1992 to 2004. Consistent with the extant literature, the forecasts tend to be optimistic. We show that, on average, the estimate of the expected rate of return based on analysts' forecasts is 2.84 percent higher than the estimate that is based on current accounting data. An implication of the observation that analysts tend to make optimistic forecasts is that caution should be taken when interpreting the meaning of the expected rate of return implied by analysts' earnings forecasts; it may not be, as the literature generally claims, an estimate of the cost of capital.

The observation that the optimism bias in analysts' forecasts may imply a 2.84 percent upward bias in the estimate of the implied expected rate of return is troublesome. Comparing this bias with the estimates of the expected equity premium based on these data (3 percent or less in Claus and Thomas (2001); between 2 and 3 percent in Gebhardt, Lee, and Swaminathan (1999); and 4.8 percent in Easton, Taylor, Shroff, and Sougiannis (2002)) suggests that there may be no premium at all! It is important to note, however, that each of these papers attributes equal weight to all stocks that are used in the calculation of the mean or median estimate of the market expected rate of return in Claus and Thomas (2001) and Gebhardt, Lee, and Swaminathan (1999), and in the regression in Easton, Taylor, Shroff, and Sougiannis (2002).

This equal-weighting has two potential effects. First, small stocks have an undue effect on the estimate of the market return. Second, stocks with low or negative earnings, which are somewhat meaningless as summary valuation metrics, potentially have an influence that is similar to the influence of large stable firms where earnings are a much more meaningful valuation metric. In order to avoid these undue influences, we repeat all of the analyses weighting each of the observations by market capitalization.

Our estimate of the implied expected rate of return on the market from the value-weighted regression, after removing the effect of bias in analysts' forecasts, is 9.67 percent with an implied equity premium of 4.43 percent. Of course, this estimate of the equity premium is more reasonable than that obtained when all observations have equal weight. We also find that the extent of analysts' optimism decreases as firm size increases. The effect of analysts' bias on the estimate of the implied expected rate of return on the market that is based on the value-weighted regression is lower than the estimate from the equally-weighted regression; 1.60 percent compared with 2.84 percent.

Studies such as Michaely and Womack (1999); Boni and Womack (2002); Eames, Glover, and Kennedy (2002); and Bradshaw (2004) show that analysts generally make "strong buy" and "buy" recommendations. They sometimes recommend "hold", and rarely recommend "sell". It seems reasonable to expect that buy recommendations will be associated with ex ante optimistic forecasts. In other words, the pervasiveness of buy recommendations may explain the optimistic bias in forecasts and in expected rates of return based on analysts' forecasts.

To examine this issue further, we repeat the analyses for sub-samples formed on the basis of number of analysts comprising the consensus who recommend "buy". Contrary to our expectations, we show that the consensus analyst forecast is optimistic even when less than 30 percent of analysts' comprising the consensus recommend "buy".[6] Estimates of the implied expected rate of return are biased upward even for these sub-samples. Interestingly, we show that the implied expected rate of return declines monotonically as the percentage of analysts recommending "buy" declines. In other words, analysts' recommendations appear to be based on expected rates of return rather than the difference between the analysts' expectations and the market expectation. This evidence is consistent with the observation in Groysberg, Healy, Chapman, and Gui (2006) that analysts' salary increases and bonuses are based on stock returns subsequent to their recommendations adjusted for the return on the S&P 500 index.

The remainder of the paper proceeds as follows. In section 2, we outline the methods used in estimating the expected rate of return implied by market prices, current book value of equity, and current and forecasted accounting earnings. Section 3 describes the data used in our analyses. In section 4, we document the ex post and the ex ante bias in consensus analysts' forecasts and discuss the implications for cost of capital estimates in extant accounting research, which are generally based on equal weighting of observations from the entire sample of firms followed by analysts. In section 5, we repeat the analyses using value-weighting of firms to show that the estimate of the bias is lower and the estimate of the expected equity risk premium is more reasonable than that obtained in extant studies. Sub-samples based on percentage of

[6] While it is reasonable to expect that the level of the analyst's recommendation should be associated with *expected* abnormal returns, it should be noted that Bradshaw (2004) finds analysts' recommendations uncorrelated with future *realized* abnormal returns.

analysts recommending buy are analyzed in section 6. Section 7 concludes with a summary of implications for future research.

2. Methods of estimating the implied expected rate of return

We develop three methods for estimating the implied expected rate of return. These estimates, which are based on (1) I/B/E/S earnings forecasts, (2) realized earnings, and (3) perfect foresight forecasts of earnings, lead to two determinations of the bias when estimates of the market expected rate of return are based on analysts' forecasts of earnings. Each of these methods determines bias as the difference between estimates based on forecasts of earnings and estimates based on earnings realizations.

We refer to the primary measure as the *ex ante* measure of bias because it relies on information available at the time of the earnings forecast. This measure compares the estimates of the implied expected rate of return based on analysts' forecasts with estimates based on current earnings realizations. The other measure compares estimates formed using analysts' forecasts with estimates based on perfect foresight of next-period earnings realizations. We refer to this as the *ex post* measure. We note there may be factors other than analysts' optimism affecting each of these measures of bias; but, since other factors affecting the ex ante measure would not affect the ex post measure (and vice-versa), obtaining similar results based on both measures suggests that the effect of other factors is minimal. We elaborate on this point in section 2.3.

2.1. Ex ante determination of the effect of bias

Each of the methods for estimating the implied expected rate of return are derived from the residual income valuation model which may be written as follows:

$$v_{jt} \equiv bps_{jt} + \sum_{\tau=1}^{\infty} \frac{eps_{jt+\tau} - r_j \times bps_{jt+\tau-1}}{\left(1+r_j\right)^{\tau}} \tag{1}$$

where v_{jt} is the intrinsic value per share of firm j at time t, bps_{jt} is the book value per share of common equity of firm j at time t, eps_{jt} is the earnings per share of firm j at time t and r_j is the cost of capital for firm j.[7] Easton, Taylor, Shroff, and Sougiannis (2002) rely on the following finite horizon version of this model:

$$p_{jt} \equiv bps_{jt} + \frac{eps_{jt+1}^{IBES} - r_j \times bps_{jt}}{\left(r_j - g_j\right)} \tag{2}$$

where p_{jt} is price per share for firm j at time t, eps_{jt+1}^{IBES} is an I/B/E/S forecast of earnings for period t+1, and g_j is the expected rate of growth in residual income beyond period $t+1$ required to equate ($p_{jt} - bps_{jt}$) and the present value of an infinite residual income stream.[8, 9]

Easton, Taylor, Shroff, and Sougiannis (2002), like many other studies, implicitly use analysts' forecasts of earnings as a proxy for market expectations of next period earnings. Optimistic bias in analysts' forecasts implies a bias in this proxy. In this paper we use a modification of the method in O'Hanlon and Steele (2000) to determine, ex ante, the effect of the forecast error on the estimate of the expected rate of return. This method provides an estimate of the expected rate of return implied by current realized accounting earnings; we compare this with

[7] Derivation of this model requires the no arbitrage assumption, which is necessary to derive the dividend capitalization formula, and that earnings are comprehensive – in other words, the articulation of earnings and book value is clean surplus.

[8] Price in this relation replaces intrinsic value. This form of the residual income model does not rely on the no-arbitrage assumption – rather it is simply based on the definition of the expected rate of return (the difference between current price and expected cum-dividend end-of-year price divided by current price).

[9] In Easton, Taylor, Shroff, and Sougiannis (2002) the period t to t+1 is 4 years so that eps_{jt+1} is aggregate expected cum-dividend earnings for the four years after date t. We use a one-year forecast horizon instead of four years in order to facilitate more effective use of the data on analysts' recommendations. Easton, Taylor, Shroff, and Sougiannis (2002) note that estimates of the expected rate of return based on just one year of forecasts are very similar to those based on four years of forecasts.

the estimate implied by analysts' earnings forecasts from Easton, Taylor, Shroff, and Sougiannis (2002).

The method adapted from O'Hanlon and Steele (2000) is based on the following form of the residual income valuation model:

$$p_{jt} \equiv bps_{jt} + \frac{\left(eps_{jt} - r_j \times bps_{jt-1}\right)\left(1 + g'_j\right)}{\left(r_j - g'_j\right)} \tag{3}$$

The difference between this form of the model and the form used by Easton, Taylor, Shroff, and Sougiannis (2002) is that g'_j is the perpetual growth rate starting from *current residual income* (that is, at time t) that implies a residual income stream such that the present value of this stream is equal to the difference between price and book value; in Easton, Taylor, Shroff, and Sougiannis (2002), g_j is the perpetual growth rate starting from *next-period residual income* (that is, time $t+1$). Since eps_{jt} (that is, realized earnings) is the only pay-off used in estimating the implied expected rate of return based on equation (3), this estimate is not affected by analysts' optimism unless that optimism is shared by the market and captured in p_{jt}.[10] Therefore, the estimate based on current accounting data can serve as an estimate of market expectations. It follows that the difference between the estimate of the expected rate of return based on analysts' forecasts in equation (2) and the estimate based on current earnings in equation (3) is an ex ante estimate of bias introduced when analysts' forecasts are used to estimate the markets' expected rate of return.

2.2. Ex post determination of the effect of bias

Optimistic bias in analysts' earnings forecasts is well-established in the literature; see, for example, O'Brien (1988); Mendenhall (1991); Brown (1993); Dugar and Nathan (1995); and

[10] Our empirical evidence is consistent with the maintained hypothesis that the analysts' optimism is not shared by the market.

Das, Levine, and Sivaramakrishnan (1998). Each of these studies estimates the ex post bias by comparing earnings forecasts with realizations of these forecasted earnings. We obtain an ex post measure of the bias in the estimate of the expected rate of return by comparing the estimate of the expected rate of return based on I/B/E/S analysts' forecasts using the method in Easton, Taylor, Shroff, and Sougiannis (2002) with the expected rate of return based on (perfect foresight forecasts of) earnings realizations; that is, we replace eps_{jt+1}^{IBES} in equation (2) with earnings realizations for period t+1, denoted eps_{jt+1}^{PF}. Of course, this ex post comparison, like the studies of bias in analysts' forecasts, will be affected by events having an effect on earnings, which happen between the time of the forecast and the date of the earnings announcement.

2.3. Ex ante and ex post comparisons

In the ex post comparison of expected rates of return, unforeseen events are *omitted* from the market price, which is used as the basis for estimating the expected rate of return. On the other hand, in the ex ante comparison, expectations of future events impounded in market expectations of earnings are not included in the current accounting earnings but are implicitly *included* in the market price, which is used as the basis for estimating the expected rate of return. Since there is no obvious reason to expect a correlation between the information omitted from price in the analyses based on equation (2) and the information included in price but excluded from earnings in the analyses based on equation (3), we use the results from both methods to gain alternative, independent estimates of the bias. As expected our results are similar using either method.

Our maintained hypothesis in the ex ante comparison of implied expected rates of return is that the market at time t sees through (un-does) the optimistic bias in the analysts' forecasts.

The observation that the implied expected rates of return based on current earnings and on realized future earnings are the same, suggests that this maintained hypothesis is reasonable.

2.4. Estimation based on prices, book value, and earnings forecasts

Easton, Taylor, Shroff, and Sougiannis (2002) transform equation (2) to form the following regression relation:

$$\frac{eps_{jt+1}}{bps_{jt}} = \gamma_0 + \gamma_1 \frac{p_{jt}}{bps_{jt}} + \mu_{jt} \tag{4}$$

where $\gamma_0 = g$, $\gamma_1 = r - g$.[11] This regression may be estimated for any group/portfolio of stocks to obtain an estimate of the implied expected rate of return, *r,* and the implied expected growth rate, *g*, for the portfolio. Easton, Taylor, Shroff, and Sougiannis (2002) run this regression for a sample of U.S. stocks to obtain an estimate of the expected rate of return on the U.S. equity market and hence an estimate of the equity premium for that market. In the empirical implementation of this model, eps_{jt+1} is the I/B/E/S forecast of earnings. Since this is the only pay-off which is used in the estimation of implied expected rate of return, any bias in the forecast will lead to a bias in the estimate of the expected rate of return.

[11] At the firm-specific level, the following relation between the regression variables: $\frac{eps_{jt+1}}{bps_{jt}} = \gamma_{0j} + \gamma_{1j} \frac{p_{jt}}{bps_{jt}}$, is readily obtained by rearranging the identity shown in equation (2). In the re-expression of this relation for a group of observations (as in equation (4)) as a regression relation, the coefficients γ_0 and γ_1 represent an average of the firm-specific γ_{0j} and γ_{1j} coefficients and the cross-sectional variation in these coefficients creates the regression residual. Easton, Taylor, Shroff, and Sougiannis (2002) describe this regression in more detail pointing out that it involves the implicit assumption that it has the properties of a random coefficient regression. It is, of course, possible that the γ_{0j} and γ_{1j} are correlated in cross-section with either (or both) the dependent or the independent variable and this correlation may introduce bias into the estimates of the regression coefficients (and, hence, into the estimates of the implied expected rates of return). It seems reasonable to assume, however, that this bias will be very similar for the regressions based on analysts' earnings forecasts (eps_{jt+1}^{IBES}) and for those based on perfect foresight forecast of earnings (eps_{jt+1}^{PF}). Also, we can think of no reason why the effect of the bias in the analyses based regression (4) will be the same as the effect for the analyses based on current accounting earnings (regression (5)). In other words, similar results from the analysis based on perfect foresight forecasts and from the analyses based on current accounting data support the conclusion that this bias does not unduly affect our estimates.

2.5. Estimation based on current accounting data

The analyses in O'Hanlon and Steele (2000) are based on realized earnings rather than earnings forecasts. Following the essence of the idea in O'Hanlon and Steele (2000), which is summarized in equation (3), we transform this equation to form the following regression relation:[12]

$$\frac{eps_{jt}}{bps_{jt-1}} = \delta_0 + \delta_1 \frac{p_{jt} - bps_{jt}}{bps_{jt-1}} + \zeta_{jt} \tag{5}$$

where $\delta_0 = r$, $\delta_1 = (r - g')/(1 + g')$. This regression may be estimated for any group/portfolio of stocks to obtain an estimate of the expected rate of return, *r,* and the expected growth rate, g', for the portfolio. O'Hanlon and Steele (2000) run a regression similar to (5) for a sample of U.K. stocks to obtain an estimate of the expected rate of return on the U.K. equity market; and hence an estimate of the equity premium for that market. In the empirical implementation of regression (5), eps_{jt} is realized earnings. Since this is the only pay-off used in estimating the implied expected rate of return, this estimate is not affected by analysts' optimism unless that optimism is shared by the market and captured in p_{jt}. It follows that the difference between the estimate of the expected rate of return obtained via regression (4) and the estimate based on regression (5) is an ex ante estimate of the bias when analysts' forecasts are used to estimate expected rates of return.

---

[12] We attribute this model to O'Hanlon and Steele (2000) because they capture its essential elements. The similarity to their model may not, however, be immediately apparent. Since the derivation in O'Hanlon and Steele (2000) is based on Ohlson (1989), the observation that the regression intercept is an estimate of the implied expected rate of return is not evident and O'Hanlon and Steele (2000) do not use it in this way. Rather, they estimate the implied expected rate of return at the firm-specific level by applying their model to time-series data and then measuring the risk premium as the slope of the Securities Market Line estimated from a regression of these firm-specific rates of return on corresponding beta estimates. Notice that, in addition to requiring earnings to be clean surplus in all future periods, this form of the residual income model also requires that the relation between earnings for period *t* and book value for periods *t* and *t*-1 follows the clean surplus relation.

2.6. The relation between prices, actual earnings, and forecasts of earnings

In order to ensure that we obtain an estimate of the expected rate of return implied by analysts' forecasts we must use prices in regression (4) that reflect analysts' forecasts. Similarly, in regression (5) we must use prices that reflect earnings realizations to obtain an estimate of the markets' expected rate of return. The alignment of price-dates, earnings announcement dates, and analysts' forecast-dates is described in this sub-section and summarized in figure 1.

We choose the first consensus forecast announced at least 14 days after the date of the earnings announcement.[13] In the analyses based on these forecasts, we use the price at the close of trade one day after the earnings announcement. Consistent with numerous studies of the information content of earnings, it seems reasonable to assume that this price incorporates the information in realized earnings. Further, we implicitly assume that this price was known to analysts at the time they formed their earnings forecasts. In view of the fact that the forecasts comprising the consensus are formed at various points in time, this assumption may be invalid; some of the forecasts comprising the consensus may precede the earnings announcement date or they may have been issued a considerable time after this date. We examine the sensitivity of the results to this assumption by varying the price-date from the day after the earnings announcement to one day after the consensus forecast is measured. This latter measurement date for price allows for the incorporation of the information in the analysts' forecasts in price. The results are not sensitive to this choice. We will return to this point.

The residual income valuation model underlying regressions (4) and (5) describes the value of a stock at the fiscal period end-date. Our analyses are based on prices after this date. To accommodate this difference, we replace price (p_{jt}) in equations (4) and (5) with price at the

[13] Use of the first forecast made after the earnings announcement from the I/B/E/S Detail History database does not alter any results.

dates described above discounted by the expected rate of return ($\hat{r}$) back to the fiscal year end; that is, $p_{jt+\tau} \big/ (1+\hat{r})^{\tau/365}$, where τ is the number of days between the fiscal year end and the price-date. Since the discounting of price requires the expected rate of return we are attempting to estimate in equations (4) and (5), we use an iterative method as used in Easton, Taylor, Shroff, and Sougiannis (2002). We begin these iterations by assuming a discount rate for prices of 12 percent. We run each regression and obtain estimates of the expected rate of return which we then use as the new rate for discounting prices. We then re-run the regressions to re-estimate equation (4) and/or equation (5) and provide another estimate of expected return. This procedure is repeated until the estimate of the expected return and the rate used in discounting price converge.[14]

3. Description of the data

All earnings forecast and recommendation data are obtained from the I/B/E/S unadjusted research databases. We use the first median consensus forecast of earnings for year t+1 released 14 days or more after the announcement of earnings for year t. This forecast is released on the third Thursday of each month. These data are obtained from the I/B/E/S Summary database. "Actual" earnings are also obtained from this database. The first year of our analyses uses forecasts and recommendations for 1993 in order to ensure the dates of the individual analysts' forecasts are reliable.[15] Book value of common equity and common shares outstanding are

[14] This iterative process is repeated until none of the annual estimates changes by more than 0.00001%. In our samples, the annual estimates usually converged in 5-6 iterations. This iterative procedure is not sensitive to choices of beginning discount rates between five and 20 percent.

[15] Zitzewitz [2002, p. 16] describes the importance of not relying on forecast dates in the I/B/E/S database prior to 1993 as follows:

> "I/B/E/S dates forecasts using the date it was entered into the I/B/E/S system. It has been well documented (e.g., by O'Brien, 1988) that the lags between a forecast becoming public and its entry into the I/B/E/S system were substantial in the 1980s (i.e., up to a month). In the 1980s, analysts mailed their forecasts,

obtained from the CRSP/COMPUSTAT annual merged database.[16] Prices are obtained from the CRSP daily price file.

We delete firms with non-December fiscal-year end so that the market implied discount rate and growth rate are estimated at the same point in time for each firm-year observation. For each set of tests, firms with any of the dependent or independent variables for that year in the top or bottom two percent of observations are removed to reduce the effects of outliers. Dropping between one and five percent of observations does not affect the conclusions of the study. For December 1999, in particular, removal of only one percent of observations has a large effect on that year's results in the value-weighted analyses; this is due to the extremely high price-to-book ratios of some internet firms prior to the market crash in 2000.

4. Ex post and ex ante bias in analysts' consensus forecasts

We begin by documenting the accuracy (that is, the mean/median *absolute* earnings forecast error) and the ex post bias (that is, the mean/median earnings forecast error) in the earnings forecasts for the entire sample of stocks. We then compare the estimate of the expected rate of return implied by prices, book values, and analysts' forecasts of earnings with the

---

often in monthly batches, to I/B/E/S where they were hand entered into the system. Since 1991-92, however, almost all analysts have entered their forecasts directly into the I/B/E/S system on the day they wish to make their forecast widely available (Kutsoati and Bernhardt, 1999). Current practice for analysts is now usually to publicly release forecasts within 24 hours of providing them to clients. I/B/E/S analysts have real-time access to each other's forecasts through this system, so an analyst entering a forecast into the system on Wednesday knows about forecasts entered on Tuesday and could potentially revise her forecast to incorporate their information. An additional advantage of the post-92 data is the shift from retrospective data entry by a specialist to real-time data entry by either the analyst or her employee should have considerably reduced data-entry related measurement error."

[16] In order to ensure that the clean-surplus assumption required for the derivation of the residual income valuation model holds in the data for fiscal year *t*, contemporaneous book value in regression (5) – that is, b_{jt} – is calculated as Compustat book value of common equity minus Compustat net income plus I/B/E/S actual income. That is, we use the book value number that would have been reported if the (corresponding) income statement had been based on I/B/E/S actual earnings. We also remove year *t* dirty surplus items from Compustat book value. These adjustments are unnecessary for the book value variable in regression (4) because the clean-surplus assumption only refers to future income statements and balance sheets.

estimate obtained from prices, book values, and actual current earnings. This is an estimate of ex ante bias in the estimates of the expected rate of return reported in the extant literature.

4.1. Accuracy and bias in the analysts' forecasts of earnings

Table 1 summarizes the accuracy and the ex post measure of bias in the I/B/E/S consensus forecast of earnings at the end of each of the years 1992 to 2003. We use the mean and the median absolute forecast error as the measure of accuracy. The mean absolute forecast error ranges from \$0.427 in 1994 to \$1.394 in 2000; the median absolute forecast error ranges from \$0.160 in 2002 to \$0.310 in 2000. We also present the mean and the median absolute forecast error deflated by end-of-year price in order to give an indication of the scale of these errors. The mean absolute price-deflated forecast error ranges from 0.019 in 2003 to 0.052 in 2000; the median absolute price-deflated forecast error ranges from 0.008 in 2003 to 0.018 in 2000.

We use the mean (median) forecast error as the measure of the ex post bias in the analysts' forecasts. The mean forecast error ranges from -\$1.257 in 2000 to \$0.119 in 2002. The median forecast error ranges from -\$0.240 in 2000 to -\$0.010 in 2003. The mean price-deflated forecast error ranges from -0.041 in 2000 to -0.003 in 2003. The median price-deflated forecast error ranges from -0.012 in 2000 to 0.000 in 2003.

These predominantly negative forecast errors are consistent with the prior literature, which concludes that analysts' forecasts, particularly long-run forecasts, tend to be optimistic; see, for example, O'Brien (1993); Lin (1994); and Richardson, Teoh, and Wysocki (2004). As noted earlier, these forecast errors compare forecasts with ex post realizations.

4.2 Description of regression variables

The number of observations we use to estimate the annual regressions ranges from 1,418 at December 1992 to 2,137 at December 1997. As shown in table 2, the mean price-to-book ratio, which is the independent variable in regression (4), ranges from 1.945 at December 2002 to 3.398 at December 1999; the median price-to-book ratio ranges from 1.625 at December 2002 to 2.409 at December 1997. Regression (4) is run with the forecasted return-on-equity based on the I/B/E/S consensus forecast as the dependent variable. The mean forecasted return-on-equity ranges from 0.079 at December 2001 to 0.146 at December 1994; the median forecasted return-on-equity ranges from 0.111 at December 2001 to 0.145 at December 1994.

The annual mean and median current return-on-equity, which is the dependent variable in regression (5), is generally a little less than the corresponding mean and median forecasted return-on-equity. The mean current return-on-equity ranges from 0.077 at December 2001 to 0.122 at December 1995; the median current return-on-equity ranges from 0.010 at December 2001 to 0.132 at December 1995. The mean of the independent variable in this regression, the difference between price and current book value deflated by lagged book value, ranges from 1.007 at December 2002 to 2.699 at December 1999; the median ranges from 0.662 at December 2002 to 1.491 at December 1997.

4.3. Comparison of implied expected rates of return based on I/B/E/S forecasts of earnings with implied expected rate of return based on current accounting data

In this section, we compare the estimates of the implied expected rates of return based on the method in Easton, Taylor, Shroff, and Sougiannis (2002), which uses one-year ahead I/B/E/S consensus forecasts of earnings in regression (4), with the estimates obtained from the method adapted from O'Hanlon and Steele (2000), which uses current earnings and current and lagged book value in regression (5). We also compare the estimates based on analysts' forecasts to those implied by future earnings realizations; that is, by perfect foresight forecasts.

4.3.1. The expected rate of return implied by analysts' earnings forecasts

The summary statistics from regression (4), where the dependent variable is I/B/E/S forecasted return-on-equity, are included in panel A of table 3. We provide year-by-year estimates of the regression coefficients and t-statistics for tests of their difference from zero. These t-statistics may be over-stated due to the possibility of correlated residuals; so we present the mean coefficient estimates and the related Fama and MacBeth (1973) t-statistics. The regression adjusted r-square ranges from 0.73 percent at December 1999, to 36.60 percent at December 1992.[17] The mean estimate of the intercept coefficient γ_0, an estimate of the implied growth in residual income beyond the one-year forecast horizon, is 0.074 with a t-statistic of 8.50. The mean estimate of the slope coefficient γ_1, an estimate of the difference between the implied expected rate of return and the implied growth in residual income beyond the one-year forecast horizon, is 0.020 with a t-statistic of 5.86.

The estimates of the implied expected rate of return obtained from the estimates of the regression (4) coefficients, where the dependent variable is analysts' forecasts of return-on-equity, are in panel A of table 3. These estimates range from 4.93 percent at December 2001, to 13.29 percent at December 1999; with a mean (t-statistic) of 9.43 percent (14.16).

[17] We note the very low r-square in some of these regressions. As a result we performed several analyses of the effects of outliers including more severe outlier removal – for example, removing up to the top and bottom 20 percent of observations or by eliminating all observations with an R-student statistic greater than 2 -- the regression r-square increases but none of our inferences based on the resulting estimates of the implied expected rate of return change. We also perform all analyses on the sub-set of observations for which analysts forecast positive earnings. Again we obtain much higher r-squares but inferences remain unchanged. These further analyses of outliers are also performed on all subsequent regressions and, in all cases, our inferences are unchanged.

4.3.2. The expected rate of return implied by current accounting data

The summary statistics from regression (5) are included in panel A of table 3. The regression adjusted r-square ranges from 0.34 percent at December 1999 to 27.09 percent at December 1992. The mean estimate of the intercept coefficient $\delta_{0,}$ which is an estimate of the implied expected rate of return, is 0.066 (t-statistic of 10.50); and the mean estimate of the slope coefficient $\delta_{1,}$ which is a function of the expected rate of return and the expected growth in residual income, is 0.022 (t-statistic of 5.51). The estimates of the implied expected rate of return are also included in panel A of table 3. These estimates range from 2.82 percent at December 2001 to 9.97 percent at December 1999; with a mean (t-statistic) of 6.59 percent (10.50).

4.3.3. The ex ante difference between the estimate of the expected rate of return based on analysts' earnings forecasts and the estimate of the expected rate of return based on current accounting data

Differences between the estimates of expected rate of return based on regressions (4) and (5) are included in the last column of panel A of table 3. On average, the difference between the estimate of the expected rate of return based on analysts' earnings forecasts and the estimate of the expected rate of return based on earnings realizations is 2.84 percent (t-statistic of 12.33). There are some years when the difference is quite large; for example, for the sample of stocks at December 1994, the difference is 3.83 percent. These results are not surprising in view of the fact that analysts' forecasts are known to be optimistic.

An implication of the observation that expected rates of return based on analysts' forecasts tend to be higher is that caution should be taken when interpreting the meaning of the rate of return that is implied by analysts' earnings forecasts; if, as is often the case in the extant literature, it is used as an estimate of the cost of capital, it is likely upward biased.

4.3.4. Estimates of the expected rate of return based on perfect foresight forecasts

The results in section 4.3.3 are roughly consistent with the results in Table 1. For example, we saw, in Table 1 that the mean deflated forecast error is -0.020. A crude PE valuation model, which relies on full payout and earnings following a random walk, suggests that the price-to-forward-earnings ratio is equal to the inverse of the expected rate of return. Thus a deflated forecast error of -0.020 implies an error in the expected rate of return of 2 percent. Allowing for the conservative nature of accounting, as in the models used in the ex ante indicators of optimism in panel A of table 3, leads to the conclusion that these estimates are at least "in the same ball-park".

Alternatively, the ex post forecast error can be re-parameterized as an error in the implied expected rate of return. This error may be estimated as the difference between the implied expected rate of return based on regression (4) where expected earnings are I/B/E/S forecasts (as in panel A of table 3) and the implied expected rate of return when these expected earnings are replaced in this regression with realized earnings for year t+1. The results of estimating the implied expected rate of return using realized earnings as "perfect foresight" forecasts are reported in panel B of table 3. Using perfect foresight earnings, the estimates of expected rate of return range from 3.13 percent at December 2001 to 9.79 percent at December 1999; with a mean (t-statistic) of 6.68 percent (10.79). Comparing the perfect foresight forecast to the consensus forecasts, the mean bias is 2.75 percent (t-statistic of 7.13).

4.3.5. Comparison of the estimates of the expected rate of return

The two estimates of expected rate of return that are not expected to contain bias, that is, those based on perfect foresight earnings and those based on current accounting data are very similar. The difference of -0.09 percent between these estimates is not significantly different from zero with a t-statistic of -0.19. It follows that our estimates of the bias are similar using either method. That is, both methods yield alternative, independent estimates of the bias that do not differ significantly; this observation supports the maintained hypothesis that the market sees through the optimistic bias in the analysts' forecasts.

Further evidence consistent with the notion that the market sees through the optimistic bias is the fact that, consistent with Richardson, Teoh, and Wysocki (2004), the forecast error declines almost monotonically as the forecast horizon decreases from approximately 12 months as in the analyses in panel C of table 3 to shortly before the earnings announcement date for year t+1. The un-tabulated associated implied expected rate of return based on these forecast and prices immediately following these forecasts also decreases almost monotonically to 6.47 percent for the consensus forecasts (of t+1 earnings) made in January of year t+1. That is, the expected rate of return implied by analysts' forecasts declines to the expected rate of return implied by the ex ante estimate of the expected rate of return implied by accounting earnings at date t. Again these results suggest that the market at date t sees through the optimistic bias in the analysts' forecasts of earnings for period t+1.

4.3.6. Effects of altering the timing of price measurement

As mentioned in section 2.3, we use price measured after the release of the prior year earnings but before analysts' forecast revisions in our primary analyses. Panel C of table 3 summarizes the results of the analysis summarized in panels A and B of table 3, but using prices measured at close of trade on the day after the consensus forecast is measured. This price is at least 14 days and could be a month and a half after the price used in panels A and B. We assume that this price reflects the information in the analysts' forecasts. Comparison of panels A and C reveals that the measurement of price at differing points; and, therefore, differing periods for discounting of price back to fiscal year-end; has no statistically or economically significant effect. The primary result from panel A of table 3 of an average 2.84 percent difference between the analysts' and market's expected rate of return is virtually unchanged at 2.93, with an un-tabulated t-statistic of 14.69, when price is measured at the day after the consensus forecast is measured.[18]

5. Value-weighted estimates of the implied expected rate of return

The analyses in section 4 examine the average effect of bias in analysts' forecasts of earnings on estimates of the implied expected rate of return. All observations are given equal weight in the analyses. Such weighting will be appropriate in some studies. Easton, Sommers, and Zmijewski (2006), for example, compare the difference between the expected rate of return implied by analysts' forecasts and the expected rate of return implied by current earnings for firms subject to litigation under section 10b-5.[19] Since the focus of their study is on average differences, they give each observation equal weight; value-weighting would lead to results that were dominated by cases associated with WorldCom and Enron.

Value-weighting will be more appropriate in many studies. Perhaps the best example is the estimation of the equity risk premium, which is a central part of three well-known studies based on analysts' earnings forecasts by Gebhardt, Lee, and Swaminathan (2001); Claus and Thomas (2001); and Easton, Taylor, Shroff, and Sougiannis (2002). These studies give equal weighting to all stocks. Yet, estimating the risk premium from investing in the equity market is more meaningful if stocks are weighted by their market capitalization. In the equally-weighted

[18] The results are virtually identical if we use prices taken from any date ranging from one day after the earnings announcement date to one day after the forecast announcement date (the set of *s* price-dates shown in Figure 1).
[19] Under Rule 10b-5, a firm and its officials can be held liable for damages to investors who bought and sold the firm's securities if the damages are attributable to investors' reliance on misleading statements or omission of material facts.

analyses in the papers referred to above, small stocks will have an undue effect on the estimate of the market return. Further, stocks with low or negative earnings, which are somewhat meaningless as summary valuation metrics, potentially have an influence that is similar to the influence of large stable firms where earnings are a much more meaningful valuation metric. In order to avoid these undue influences, and to provide an estimate of the equity risk premium that is (1) not affected by analysts' optimism; and (2) more representative of the risk premium for the market portfolio; we repeat all of the analyses weighting each of the observations by market capitalization.

In order to provide a sense of the likely effect of value weighting, we begin by describing the way that analysts' optimism differs with firm size. We also document the relation between firm size and the variables used in regressions (4) and (5). Central to our analyses is the observation, documented in panel A of table 4, that the mean scaled absolute forecast error declines in a monotonic manner from 0.102 for the decile of smallest firms to 0.012 for the decile of largest firms. Similarly, the median absolute scaled forecast error declines in a monotonic manner from 0.042 to 0.006.

Analysts' optimism, measured by the mean (median) forecast error, declines almost monotonically from -0.116 (-0.023) for the decile of smallest firms to -0.086 (-0.002) for the decile of largest firms. The differences in optimistic bias across these size deciles illustrate the point that difference in bias across samples of observations may explain a significant portion of the difference in the implied expected rates of return across these samples; in other words, differences in bias across samples may lead to spurious inferences.

Consistent with prior literature, see, for example, Fama and French (1992), the price-to-book ratio increases with firm size from a mean of 1.707 for the decile of smallest firms to a mean of 3.593 for the decile of largest firms. The forecasted and the realized return-on-equity also increase with firm size, suggesting that the smaller firms tend to be firms with higher expected earnings growth.[20]

The results from the estimation of value-weighted regressions (4) and (5) are summarized in panel B of table 4. A notable difference between these value-weighted regression results and the results for equally-weighted regressions (see panels A and B of table 3) is the higher adjusted r-square for the value-weighted regressions. For example, the average adjusted r-square for regression (4) based on analysts' consensus forecasts is 47.16 percent for the value-weighted regression; whereas it is 9.58 percent for the equally-weighted regression. As expected, t-statistics on the coefficient estimates in these value-weighted regressions are also higher.

The mean estimates (t-statistic) of the expected rate of return, also reported in panel B of table 4, are 11.27 percent (21.20) using analysts' forecasts and 9.67 percent (13.90) using current accounting data.[21] The un-tabulated minimum expected rate of return estimated using current accounting data is 6.22 percent at December 1992. The average of 9.67 percent yields a more reasonable estimate of the risk premium than the equal-weighted sample; 4.43 percent using 5-year treasuries as a proxy for the risk free rate. Differences between the estimates are also reported in panel B of table 4. The difference, though smaller in the value-weighted analyses than in the equally-weighted analyses, 1.60 percent compared with 2.84 percent, is still significantly positive (t-statistic of 4.90).

---

[20] The firms in the deciles of smaller firms also tend to have a much greater proportion of losses (the proportion of losses decreases monotonically from 17.64 percent for the decile of smallest firms to 1.65 percent for the decile of largest firms).

[21] The mean estimate (t-statistic) of the expected rate of return based on perfect foresight forecasts is 10.63 percent (14.35).

6. Variation in the implied expected rate of return with changes in the percentage of analysts making “buy” recommendations

Having documented a bias in the estimates of the expected rate of return based on analysts’ forecasts of earnings, we now examine how the bias varies across analysts’ recommendations. It is well-known that analysts seldom issue “sell” recommendations. To the extent that our samples examined thus far contain a majority of firms with “buy” recommendations, the observed positive bias in the expected rate of return using analysts’ forecasts may be capturing the analysts’ expectation of the abnormal returns, which can be earned from these stocks. To examine this notion, we compare estimates of the expected rates of return for stocks where the consensus forecast is comprised of analysts with varying recommendation types.

6.1 Sample description

I/B/E/S provides data on the percentage of analysts whose forecasts comprise the consensus who also make either a “strong buy” or a “buy” recommendation. We repeat the analyses in section 4.3 for sub-samples with various percentages of these types of recommendations. Descriptive statistics are provided in table 5, panel A. The choice of the five partitions of the data is based on a desire to maintain a sufficient number of observations to provide reasonable confidence in the regression output in each year. We restrict the sample to those consensus forecasts which are comprised of at least 5 analysts so that it is possible for a firm to appear in any of the partitions.[22]

The mean and median forecast error is always negative; that is, analysts are optimistic, regardless of the percentage of “buy” recommendations in the consensus. For example, the median deflated forecast error is -0.004 when the percentage of buy recommendations is greater

[22] Our findings and conclusions are unchanged when firms with consensus forecasts comprised of less than 5 analysts are included.

than 90 percent, between 30 and 50 percent, and when the percentage of “buy” recommendations is less than 30 percent.

Both the return-on-equity and the price-to-book ratio tend to be higher for the observations where there are more “buy” recommendations comprising the consensus. For example, the median forecasted return-on-equity for the sub-samples where greater than 90 percent of the analysts recommend “buy” and where between 70 and 90 percent recommend “buy” is 0.157 and 0.162 while median forecasted return-on-equity for the sub-sample where less than 30 percent of the analysts recommend “buy” is 0.112. The median price-to-book ratio for the sub-samples where greater that 90 percent of the analysts recommend “buy” and where between 70 and 90 percent recommend “buy” is 3.011 and 2.686 while median price-to-book ratio for the sub-samples where less than 30 percent of the analysts recommend “buy” is 1.649.

6.2. Estimates of implied expected rates of return

The results from the estimation of regression (4) based on price, I/B/E/S forecasts of earnings, and current book value and from the estimation of regression (5) based on price and current accounting data and are summarized in table 5, panel B. We focus our discussion on the estimates of the implied expected rates of return obtained from these regression parameters. These estimates are also included in panel B.

The estimates of the expected rates of return implied by I/B/E/S analysts’ forecasts decline almost monotonically with the percentage of “buy” recommendations associated with the forecasts of earnings comprising the consensus; the means of these estimates are 11.20 percent, 11.84 percent, 10.82 percent, 9.18 percent, and 6.86 percent, suggesting that analysts’ recommendations are, indeed, consistent with the implied expectations of rates of return. The estimates of the expected rates of return based on prices and current accounting data show a pattern that is very similar to that of those based on analysts' forecasts. The mean estimates of the expected rate of return for each of the groups of data decline monotonically with the percentage of "buy" recommendations associated with the forecasts of earnings comprising the consensus; the means of these estimates are 10.94 percent, 10.22 percent, 8.90 percent, 7.23 percent, and 4.60 percent.

Differences between the estimates of expected rate of return based on percentage of "buy" recommendations are included in table 5, panel C. Comparing the expected rates of return based on prices and current accounting data with the estimates based on analysts' forecasts reveals that even when the analysts are not to recommending "buy" their forecasts imply a rate of return that is higher than expectations based on current accounting data; these mean differences between the estimates based on analysts' forecasts and estimates based on current accounting data are 0.26 percent, 1.61 percent, 1.92 percent, 1.95 percent, and 2.27 percent. Four of these differences are significant. This pervasive optimism in the expected return measured by comparing analysts' return expectations with return expectations based on current accounting data is, interestingly, quite similar to the pervasive optimism observed when comparing expectations of future earnings with actual realizations of earnings; see table 5, panel A.

6.3. Summary

To summarize the analyses in this section, we observe that analysts' recommendations are consistent with their expectations of returns; that is, there is a monotonic decrease in expected rate of return as the percentage of "buy" recommendations declines.[23] Analysts' expected rates of return are higher than expectations based on current accounting data regardless of their recommendation. An interpretation of this result is that analysts are always optimistic;

[23] Our findings and conclusions are unchanged when the analysis is repeated using a value-weighted analysis similar to section 5.

even when they are not issuing “buy” recommendations.[24] The bias in expected rates of return based on analysts’ forecasts is not the result of analysts’ expectations of positive abnormal returns isolated in firms with “buy” or “strong buy” recommendations.

7. Summary and conclusions

We show that, on average, the difference between the estimate of the expected rate of return based on analysts’ earnings forecasts and the estimate of based on current earnings realizations is 2.84 percent. An implication of the observation that rates of return based on analysts’ forecasts are higher than market expectations is that caution should be taken when interpreting the meaning of the rate of return that is implied by analysts’ earnings forecasts; it may not be, as the literature generally claims, an estimate of the cost of capital.

When estimates of the expected rate of return in the extant literature are adjusted to remove the effect of optimism bias in analysts’ forecasts, the estimate of the equity risk premium appears to be approximately zero. We show, however, when estimates are based on value-weighted analyses, the bias in the estimate of the expected rate of return is lower and the estimate of the expected equity premium is more reasonable; 4.43 percent.

Results from sub-samples formed on the basis of percentage of analysts comprising the consensus recommending “buy” show that the estimate of the expected rate of return, based on both analysts’ forecasts of earnings and on current earnings, declines in a monotonic manner as the percentage of analysts recommending “buy” declines. A comparison of the estimates of the expected rate of return based on the analysts’ forecasts, with estimates based on earnings realizations, suggests that analysts tend to be more optimistic than the market even when they are

[24] This result is consistent with Barber, Lehavy, McNicholls, and Trueman (2001) who show that analysts’ recommendations (in their case, those summarized in the Zach’s database) can not be used to form profitable trading strategies.

not making "buy" recommendations. That is, analysts recommend "buy" when they expect the future return to be high and "sell" when they expect the return to be low regardless of market expectations.

Our paper has two key implications for future research which uses market price, book value of equity, and accounting earnings to obtain estimates of the implied expected rate of return for a portfolio of stocks. First, since analysts' forecasts are pervasively optimistic, estimates of the implied expected rate of return formed using forecasts will be pervasively and significantly upward biased. This bias may be avoided by estimating the rate of return implied by price, book values, and *realized* earnings rather than biased earnings *forecasts*. Second, value-weighted analyses may be more appropriate in addressing certain issues such as estimating the equity premium, than equal-weighted analyses. The value-weighted analyses may provide more realistic estimates of the expected rate of return than are implied by equally-weighted analyses; which may be unnecessarily affected by less representative observations, such as penny stocks, and stocks making losses.

When coupled with results from the papers that demonstrate the troublesome effects of measurement error in firm-specific estimates of the expected rate of return, the results in this study suggest that the extant measures of implied expected rate of return should be used with considerable caution. The challenge is to find means of reducing the measurement error and to mitigate the effects of bias. Easton and Monahan (2005) suggest focusing on sub-samples where the measurement error is likely to be small. Our paper suggests that methods based on realized earnings rather than earnings forecasts may be a possible means of avoiding the effects of bias in analysts' forecasts. Another possible avenue might be to attempt to un-do the bias; following, for example, the ideas in Frankel and Lee (1998).

Figure 1: Alignment of Price-Dates, Earnings Announcement Dates, and Analysts' Forecast-Dates

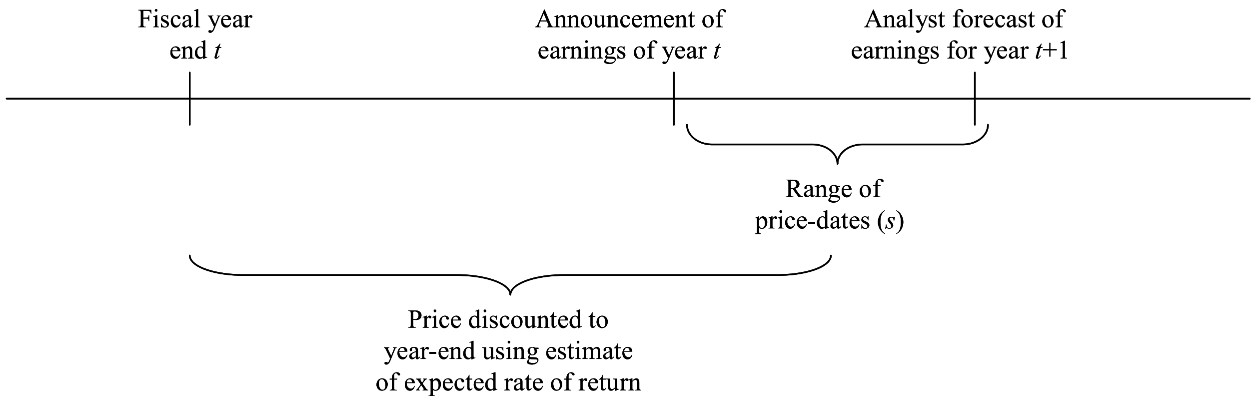


Table 1: Descriptive statistics on forecast errors for the consensus sample

		Accuracy of forecasts				Bias in forecasts			
		$\lvert FE_{jt+1}\rvert$		$\lvert FE_{jt+1}\rvert / p_{jt}$		FE_{jt+1}		FE_{jt+1} / p_{jt}	
t	N	Mean	Median	Mean	Median	Mean	Median	Mean	Median
12/92	1,418	0.594	0.280	0.030	0.014	-0.241	-0.150	-0.017	-0.007
12/93	1,544	0.461	0.190	0.028	0.009	-0.228	-0.070	-0.019	-0.003
12/94	1,781	0.427	0.220	0.030	0.012	-0.206	-0.080	-0.019	-0.004
12/95	1,939	0.451	0.210	0.028	0.011	-0.261	-0.070	-0.019	-0.004
12/96	2,006	0.518	0.210	0.027	0.010	-0.187	-0.100	-0.018	-0.005
12/97	2,137	0.606	0.270	0.031	0.013	-0.376	-0.200	-0.024	-0.009
12/98	2,044	0.718	0.215	0.040	0.012	-0.515	-0.080	-0.025	-0.004
12/99	1,854	0.668	0.230	0.046	0.012	-0.399	-0.090	-0.028	-0.004
12/00	1,729	1.394	0.310	0.052	0.018	-1.257	-0.240	-0.041	-0.012
12/01	1,809	0.705	0.200	0.033	0.011	0.063	-0.060	-0.018	-0.003
12/02	1,825	0.570	0.160	0.031	0.011	0.119	-0.030	-0.012	-0.002
12/03	2,000	0.650	0.170	0.019	0.008	-0.251	-0.010	-0.003	0.000
Means	1,841	0.647	0.222	0.033	0.012	-0.312	-0.098	-0.020	-0.005

Notes to Table 1:

FE_{jt+1} is actual earnings per share for year *t*+1 as reported by I/B/E/S less the first median consensus forecast of earnings per share for year t+1 released at least 14 days after the announcement of year *t* earnings

p_{jt} is price per share as of the end of fiscal year *t*

Table 2: Summary statistics for regression variables

		$\frac{eps_{jt+1}^{Cons}}{bps_{jt}}$ Equation (4) dependent variable		$\frac{eps_{jt}}{bps_{jt-1}}$ Equation (5) dependent variable		$\frac{p'_{jt}}{bps_{jt}}$ Equation (4) independent variable		$\frac{p'_{jt}-bps^{*}_{jt}}{bps_{jt-1}}$ Equation (5) independent variable	
t	N	Mean	Median	Mean	Median	Mean	Median	Mean	Median
12/92	1,418	0.138	0.132	0.104	0.110	2.193	1.792	1.265	0.854
12/93	1,544	0.138	0.138	0.113	0.122	2.374	1.929	1.505	0.994
12/94	1,781	0.146	0.145	0.121	0.126	2.114	1.706	1.334	0.834
12/95	1,939	0.145	0.142	0.122	0.132	2.454	1.906	1.679	1.060
12/96	2,006	0.135	0.139	0.108	0.126	2.654	2.114	1.851	1.228
12/97	2,137	0.125	0.140	0.102	0.125	2.998	2.409	2.132	1.491
12/98	2,044	0.118	0.134	0.093	0.116	2.728	1.974	1.810	0.959
12/99	1,854	0.126	0.141	0.094	0.124	3.398	1.883	2.699	0.996
12/00	1,729	0.116	0.136	0.100	0.130	2.749	1.964	2.022	1.109
12/01	1,809	0.079	0.111	0.068	0.100	2.457	1.928	1.548	0.989
12/02	1,825	0.093	0.117	0.077	0.102	1.945	1.625	1.007	0.662
12/03	2,000	0.106	0.121	0.090	0.111	2.883	2.314	2.198	1.450
Means	1,841	0.122	0.133	0.099	0.119	2.579	1.962	1.754	1.052

Notes to Table 2:

eps_{jt+1}^{Cons} is the first median consensus forecast of earnings per share for firm *j* for year *t*+1 released at least 14 days after the announcement of year *t* earnings

eps_{jt} is the I/B/E/S actual earnings per share for firm *j* for year *t*

bps_{jt} is common book value of equity per share for firm *j* at time *t*

$p'_{jt} = \frac{p_{jt+\tau}}{(1+\hat{r})^{\tau/365}}$ is the price per share for firm *j* at time t+τ (one day after the earnings announcement date), $p_{jt+\tau}$, adjusted for stock splits and stock dividends since the end of the fiscal year, discounted to year end using the estimated discount rate

bps^{*}_{jt} is the common book value of equity per share for firm *j* at time *t* less net income for firm *j* for year *t* plus I/B/E/S actual earnings per share for firm *j* for year *t*

Table 3: Comparison of implied expected rates of return based on I/B/E/S forecasts of earnings with implied expected rate of return based on current accounting data

Panel A: Estimates of expected rate of return based on analysts' forecasts and current accounting data

$$\frac{eps_{jt+1}^{Cons}}{bps_{jt}} = \gamma_0 + \gamma_1 \frac{p'_{jt}}{bps_{jt}} + \mu_{jt} \quad (4) \qquad \frac{eps_{jt}}{bps_{jt-1}} = \delta_0 + \delta_1 \frac{p'_{jt} - bps^*_{jt}}{bps_{jt-1}} + \zeta_{jt} \quad (5)$$

		Analysts' consensus earnings forecasts				**Current accounting data**				**Difference in expected rate of return**
T	N	γ_0	γ_1	Adj R^2	$\hat{r} = \gamma_0 + \gamma_1$	δ_0	δ_1	Adj R^2	$\hat{r} = \delta_0$	
12/92	1,418	0.057	0.037	36.60%	9.39%	0.057	0.037	27.09%	5.67%	3.72%
		(17.71)	(28.62)			(18.96)	(22.97)			
12/93	1,544	0.073	0.027	15.59%	10.08%	0.068	0.030	15.32%	6.83%	3.25%
		(16.53)	(16.91)			(18.37)	(16.74)			
12/94	1,781	0.073	0.035	16.81%	10.73%	0.069	0.039	24.00%	6.90%	3.83%
		(16.25)	(18.99)			(21.01)	(23.73)			
12/95	1,939	0.095	0.021	10.83%	11.53%	0.092	0.018	6.55%	9.22%	2.31%
		(23.47)	(15.38)			(23.40)	(11.70)			
12/96	2,006	0.089	0.018	6.66%	10.61%	0.073	0.019	6.77%	7.26%	3.35%
		(18.91)	(12.00)			(16.79)	(12.11)			
12/97	2,137	0.082	0.014	3.71%	9.64%	0.066	0.017	5.60%	6.62%	3.02%
		(14.64)	(9.13)			(14.61)	(11.30)			
12/98	2,044	0.082	0.013	3.50%	9.50%	0.065	0.016	6.43%	6.49%	3.01%
		(15.23)	(8.67)			(15.86)	(11.89)			
12/99	1,854	0.136	-0.003	0.73%	13.29%	0.100	-0.002	0.34%	9.97%	3.32%
		(32.67)	(-3.83)			(22.54)	(-2.71)			
12/00	1,729	0.084	0.012	3.38%	9.57%	0.086	0.007	1.00%	8.61%	0.96%
		(15.42)	(7.84)			(16.02)	(4.30)			
12/01	1,809	0.029	0.020	4.63%	4.93%	0.028	0.026	9.99%	2.82%	2.11%
		(4.64)	(9.42)			(6.30)	(14.20)			
12/02	1,825	0.019	0.038	9.83%	5.70%	0.030	0.047	21.13%	2.96%	2.74%
		(3.12)	(14.14)			(7.98)	(22.13)			
12/03	2,000	0.069	0.013	2.72%	8.18%	0.057	0.015	4.35%	5.74%	2.44%
		(11.65)	(7.55)			(11.55)	(9.59)			
Means	1,841	0.074	0.020	9.58%	9.43%	0.066	0.022	10.71%	6.59%	2.84%
t-Statistics		(8.50)	(5.86)		(14.16)	(10.50)	(5.51)		(10.50)	(12.33)

Table 3: Continued

Panel B: Estimates of expected rate of return based on future realized earnings

	$\frac{eps_{jt+1}}{bps_{jt}} = \gamma_0 + \gamma_1 \frac{p'_{jt}}{bps_{jt}} + \mu_{jt}$ (4) **Perfect foresight earnings forecasts**				**Analysts' Forecasts Less Perfect Foresight**	**Current Accounting Data Less Perfect Foresight**
t	γ_0	γ_1	Adj R^2	$\hat{r} = \gamma_0 + \gamma_1$		
12/92	0.037	0.031	14.10%	6.77%	2.62%	-1.10%
	(7.09)	(15.31)				
12/93	0.049	0.026	7.97%	7.45%	2.63%	-0.62%
	(8.10)	(11.61)				
12/94	0.046	0.031	8.33%	7.71%	3.02%	-0.81%
	(7.56)	(12.77)				
12/95	0.076	0.013	2.22%	8.87%	2.66%	0.35%
	(13.29)	(6.69)				
12/96	0.082	0.004	0.12%	8.56%	2.05%	-1.30%
	(12.01)	(1.83)				
12/97	0.040	0.009	0.77%	4.89%	4.75%	1.73%
	(5.14)	(4.18)				
12/98	0.057	0.006	0.44%	6.27%	3.23%	0.22%
	(8.28)	(3.15)				
12/99	0.105	-0.007	1.87%	9.79%	3.50%	0.18%
	(17.73)	(-6.01)				
12/00	0.043	0.004	0.18%	4.70%	4.87%	3.91%
	(6.16)	(2.05)				
12/01	0.018	0.013	1.40%	3.13%	1.80%	-0.31%
	(2.47)	(5.16)				
12/02	-0.003	0.041	9.16%	3.77%	1.93%	-0.81%
	(-0.48)	(13.60)				
12/03	0.075	0.007	0.64%	8.28%	-0.10%	-2.54%
	(11.02)	(3.71)				
Means	0.052	0.015	3.93%	6.68%	2.75%	-0.09%
t-Statistics	(6.12)	(3.63)		(10.79)	(7.13)	(-0.19)

Table 3: Continued

Panel C: Comparison of implied expected rates of return based on I/B/E/S forecasts of earnings with implied expected rate of return based on current accounting data and on future realized earnings using prices measured the day after the consensus forecast

$$\frac{eps^{Cons}_{jt+1}}{bps_{jt}} = \gamma_0 + \gamma_1 \frac{p'_{jt}}{bps_{jt}} + \mu_{jt} \quad (4)$$

Analysts' consensus earnings forecasts

	N	γ_0	γ_1	Adj R^2	$\hat{r} = \gamma_0 + \gamma_1$
Means	1,841	0.072	0.021	10.07%	9.34%
t-Statistics		(8.04)	(5.93)		(13.68)

$$\frac{eps_{jt}}{bps_{jt-1}} = \delta_0 + \delta_1 \frac{p'_{jt} - bps^*_{jt}}{bps_{jt-1}} + \zeta_{jt} \quad (5)$$

Current accounting data

	N	δ_0	δ_1	Adj R^2	$\hat{r} = \delta_0$
Means	1,841	0.064	0.023	11.36%	6.41%
t-Statistics		(10.13)	(5.86)		(10.13)

$$\frac{eps^{PF}_{jt+1}}{bps_{jt}} = \gamma_0 + \gamma_1 \frac{p'_{jt}}{bps_{jt}} + \mu_{jt} \quad (4)$$

Perfect foresight earnings forecasts

	N	γ_0	γ_1	Adj R^2	$\hat{r} = \gamma_0 + \gamma_1$
Means	1,841	0.049	0.016	4.42%	6.50%
t-Statistics		(5.36)	(3.84)		(9.72)

Notes to Table 3:

Panel A of the table reports the results of estimating regression (4) using I/B/E/S consensus forecasts and regression (5) using current accounting data cross-sectionally using all available observations. Panel B reports the results of estimating regression (4) using subsequent earnings realizations as perfect foresight forecasts. Observations with any of the dependent or independent variables in the top and bottom two percent observations are removed to reduce the effects of outliers. The variables are as defined in the notes to Tables 1 and 2. Summary means across the annual regressions and the related Fama and MacBeth (1973) t-statistics are provided. The last column of Panel A contains the difference between estimates of expected return from the estimation of regression (4) using I/B/E/S consensus forecasts and regression (5) using current accounting data. The last two columns of Panel B contain the differences between perfect foresight estimates and the estimates of expected return from the estimation of regression (4) using I/B/E/S consensus forecasts and regression (5) using current accounting data. Panel C repeats the analysis performed in Panels A and B using an alternative definition of price. Instead of measuring price at trade close the day after the earnings announcement, price is measured at trade close the day following the consensus forecast. This results in a price variable measured 14 days to a month and a half later. All other variables remain unchanged.

Table 4: Value-weighting observations, results of comparison of implied expected rates of return based on I/B/E/S forecasts of earnings, based on current accounting data and based on future realizations of earnings

Panel A: Descriptive statistics

	Decile of market capitalization at time t									
Mean of annual means	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
$\lvert FE_{jt+1}\rvert$	0.419	0.397	0.398	0.443	0.428	0.455	0.466	0.488	0.579	2.369
$\lvert FE_{jt+1}\rvert / p_{jt}$	0.102	0.053	0.040	0.034	0.026	0.023	0.018	0.017	0.015	0.012
FE_{jt+1}	-0.284	-0.235	-0.242	-0.266	-0.233	-0.237	-0.214	-0.246	-0.273	-0.890
FE_{jt+1}/p_{jt}	-0.075	-0.033	-0.025	-0.021	-0.015	-0.013	-0.009	-0.009	-0.007	-0.005
$eps_{jt+1}^{Cons} / bps_{jt}$	0.065	0.081	0.093	0.095	0.113	0.128	0.140	0.149	0.160	0.186
eps_{jt} / bps_{jt-1}	0.002	0.050	0.066	0.075	0.095	0.113	0.126	0.134	0.145	0.168
p'_{jt} / bps_{jt}	1.707	1.954	2.188	2.362	2.482	2.676	2.794	2.895	2.941	3.593
$\left(p'_{jt} - bps^{*}_{jt}\right) / bps_{jt-1}$	0.641	1.000	1.275	1.533	1.752	1.958	2.083	2.142	2.146	2.732

	Decile of market capitalization at time t									
Mean of annual medians	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
$\lvert FE_{jt+1}\rvert$	0.218	0.200	0.211	0.225	0.225	0.221	0.238	0.223	0.242	0.246
$\lvert FE_{jt+1}\rvert / p_{jt}$	0.042	0.024	0.018	0.016	0.012	0.010	0.009	0.008	0.007	0.006
FE_{jt+1}	-0.116	-0.106	-0.108	-0.116	-0.098	-0.092	-0.092	-0.090	-0.075	-0.086
FE_{jt+1}/p_{jt}	-0.023	-0.012	-0.009	-0.007	-0.005	-0.004	-0.004	-0.003	-0.002	-0.002
$eps_{jt+1}^{Cons} / bps_{jt}$	0.095	0.110	0.115	0.118	0.126	0.134	0.143	0.148	0.155	0.176
eps_{jt} / bps_{jt-1}	0.052	0.086	0.097	0.104	0.114	0.125	0.131	0.136	0.142	0.160
p'_{jt} / bps_{jt}	1.316	1.577	1.748	1.836	1.926	2.060	2.183	2.221	2.304	2.829
$\left(p'_{jt} - bps^{*}_{jt}\right) / bps_{jt-1}$	0.259	0.605	0.818	0.944	1.017	1.220	1.327	1.313	1.439	1.934

Table 4: Continued

Panel B: Value-weighted estimates of expected rate of return based on analysts' forecasts and current accounting data

$$\frac{eps_{jt+1}^{Cons}}{bps_{jt}} = \gamma_0 + \gamma_1 \frac{p'_{jt}}{bps_{jt}} + \mu_{jt} \quad (4)$$

$$\frac{eps_{jt}}{bps_{jt-1}} = \delta_0 + \delta_1 \frac{p'_{jt} - bps^*_{jt}}{bps_{jt-1}} + \zeta_{jt} \quad (5)$$

		Analysts' consensus earnings forecasts				**Current accounting data**				**Difference in expected rate of return**
T	N	γ_0	γ_1	Adj R^2	$\hat{r} = \gamma_0 + \gamma_1$	δ_0	δ_1	Adj R^2	$\hat{r} = \delta_0$	
12/92	1,418	0.047	0.047	57.76%	9.35%	0.062	0.044	46.89%	6.22%	3.13%
		(14.73)	(44.03)			(23.49)	(35.38)			
12/93	1,544	0.052	0.047	51.76%	9.82%	0.079	0.042	46.23%	7.87%	1.95%
		(14.70)	(40.70)			(29.00)	(36.43)			
12/94	1,781	0.072	0.049	52.03%	12.15%	0.084	0.050	57.05%	8.39%	3.76%
		(22.46)	(43.95)			(34.82)	(48.64)			
12/95	1,938	0.092	0.036	46.89%	12.76%	0.127	0.028	32.37%	12.65%	0.11%
		(26.96)	(41.36)			(41.25)	(30.46)			
12/96	2,006	0.081	0.034	51.09%	11.53%	0.106	0.029	44.72%	10.64%	0.89%
		(25.50)	(45.77)			(38.36)	(40.29)			
12/97	2,137	0.094	0.026	44.60%	12.01%	0.106	0.023	39.89%	10.58%	1.43%
		(28.17)	(41.48)			(41.10)	(37.67)			
12/98	2,044	0.093	0.022	47.17%	11.49%	0.090	0.022	49.99%	8.97%	2.52%
		(28.30)	(42.72)			(33.70)	(45.20)			
12/99	1,855	0.147	0.010	23.55%	15.69%	0.147	0.004	4.00%	14.66%	1.03%
		(35.74)	(23.92)			(36.07)	(8.85)			
12/00	1,729	0.091	0.022	43.02%	11.26%	0.110	0.021	33.61%	11.04%	0.22%
		(22.09)	(36.13)			(28.77)	(29.60)			
12/01	1,808	0.059	0.031	44.84%	8.98%	0.070	0.030	47.31%	6.98%	2.00%
		(15.74)	(38.34)			(22.45)	(40.29)			
12/02	1,825	0.055	0.043	59.95%	9.76%	0.083	0.041	61.56%	8.26%	1.50%
		(18.77)	(52.26)			(34.75)	(54.05)			
12/03	2,000	0.072	0.032	43.22%	10.41%	0.098	0.031	40.17%	9.76%	0.65%
		(21.58)	(39.02)			(27.36)	(36.65)			
Means	1,841	0.079	0.033	47.16%	11.27%	0.097	0.030	41.98%	9.67%	1.60%
t-Statistics		(10.09)	(9.62)		(21.20)	(13.90)	(8.38)		(13.90)	(4.91)

Notes to Table 4:

Panel A of the table reports the summary statistics from repeating the analysis performed in Tables 1 and 2 by annual decile of market capitalization at time *t*. Panel B repeats the analysis in Table 3 using weighted least squares regression with regression weights equal to market capitalization at time *t*.

Table 5: Variation in the implied expected rate of return with changes in the percentage of analysts' making "buy" recommendation – minimum of five analysts following firm

Panel A: Descriptive statistics by percent of buy recommendations

	90 ≤ % Buy ≤ 100		70 ≤ % Buy ≤ 90		50 ≤ % Buy < 70		30 ≤ % Buy < 50		0 ≤ % Buy < 30	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
$\lvert FE_{jt+1}\rvert$	0.437	0.218	0.932	0.232	0.497	0.220	0.540	0.235	0.536	0.229
$\lvert FE_{jt+1}\rvert / p_{jt}$	0.017	0.008	0.017	0.008	0.019	0.008	0.026	0.010	0.041	0.011
FE_{jt+1}	-0.268	-0.101	-0.725	-0.103	-0.251	-0.083	-0.271	-0.089	-0.287	-0.082
FE_{jt+1}/p_{jt}	-0.010	-0.004	-0.009	-0.003	-0.010	-0.003	-0.016	-0.004	-0.027	-0.004
$eps^{Cons}_{jt+1} / bps_{jt}$	0.140	0.157	0.164	0.162	0.159	0.153	0.134	0.131	0.108	0.112
eps_{jt} / bps_{jt-1}	0.125	0.150	0.152	0.151	0.143	0.140	0.120	0.120	0.091	0.101
p'_{jt} / bps_{jt}	3.860	3.011	3.435	2.686	2.848	2.305	2.371	1.921	2.029	1.649
$(p'_{jt} - bps^{*}_{jt}) / bps_{jt-1}$	3.649	2.313	2.844	1.948	2.005	1.438	1.485	1.016	1.032	0.704
# of observations	135		227		263		176		154	

Table 5: Continued

Panel B: Summary of results of estimation by percent of buy recommendations

$$\frac{eps^{Cons}_{jt+1}}{bps_{jt}} = \gamma_0 + \gamma_1 \frac{p'_{jt}}{bps_{jt}} + \mu_{jt} \quad (4)$$

$$\frac{eps_{jt}}{bps_{jt-1}} = \delta_0 + \delta_1 \frac{p'_{jt} - bps^*_{jt}}{bps_{jt-1}} + \zeta_{jt} \quad (5)$$

		Analysts' consensus earnings forecasts				**Current accounting data**			
Recommendation	N	γ_0	γ_1	Adj R^2	$\hat{r} = \gamma_0 + \gamma_1$	δ_0	δ_1	Adj R^2	$\hat{r} = \delta_0$
90 ≤ % Buy ≤ 100	135	0.100	0.012	7.90%	11.20%	0.109	0.012	18.18%	10.94%
		(7.93)	(3.32)		(9.93)	(5.12)	(1.46)		(5.12)
70 ≤ % Buy ≤ 90	227	0.098	0.021	16.82%	11.84%	0.102	0.020	17.42%	10.22%
		(9.87)	(7.73)		(14.29)	(10.23)	(5.88)		(10.23)
50 ≤ % Buy < 70	263	0.080	0.029	34.28%	10.82%	0.089	0.028	30.29%	8.90%
		(13.67)	(12.69)		(20.84)	(18.09)	(10.96)		(18.09)
30 ≤ % Buy < 50	176	0.060	0.031	28.31%	9.18%	0.072	0.033	26.85%	7.23%
		(7.04)	(6.80)		(16.25)	(13.25)	(8.38)		(13.25)
0 ≤ % Buy < 30	154	0.032	0.037	32.00%	6.86%	0.046	0.044	30.09%	4.60%
		(3.13)	(9.60)		(8.85)	(5.60)	(9.67)		(5.60)

Table 5: Continued

Panel C: Mean differences in (t-statistics for) estimates of expected rate of return

		Analysts' expected rate of return					**Expected rate of return based on current accounting data**			
		90 ≤ % ≤ 100	70 ≤ % ≤ 90	50 ≤ % < 70	30 ≤ % < 50	0 ≤ % < 30	90 ≤ % ≤ 100	70 ≤ % ≤ 90	50 ≤ % < 70	30 ≤ % < 50
Analysts' expected rate of return	70 ≤ % ≤ 90	-0.64% (-0.79)								
	50 ≤ % < 70	0.38% (0.50)	1.02% (2.11)							
	30 ≤ % < 50	2.02% (2.50)	2.66% (4.76)	1.64% (3.96)						
	0 ≤ % < 30	4.34% (5.46)	4.97% (9.01)	3.96% (8.90)	2.31% (5.04)					
Expected rate of return based on current accounting data	90 ≤ % ≤ 100	0.26% (0.15)								
	70 ≤ % ≤ 90		1.61% (3.14)				0.72% (0.30)			
	50 ≤ % < 70			1.92% (5.04)			2.04% (1.03)	1.32% (1.81)		
	30 ≤ % < 50				1.95% (6.38)		3.72% (1.82)	3.00% (4.77)	1.68% (3.96)	
	0 ≤ % < 30					2.27% (7.15)	6.35% (3.15)	5.63% (8.25)	4.31% (7.40)	2.63% (5.29)

Table 5: Continued

Notes to Table 5:

Using the median consensus analysts' forecast and the percent of buy recommendations from the summary I/B/E/S database, we estimate expected rate of return by percentage of buy recommendations for all firms with at least five analysts included in the consensus. Panel A reports descriptive statistics by percentage of buy recommendations. The variables are as defined in the notes to Tables 1 and 2. Panel B reports the results of estimating regression (4) using I/B/E/S consensus forecasts and regression (5) using current accounting data cross-sectionally using all available observations of that percentage of buy recommendations. Within the percentage of buy recommendations, observations with any of the dependent or independent variables in the top and bottom two percent observations are removed to reduce the effects of outliers. The reported numbers are the summary means across the annual regressions and the related Fama and Macbeth (1973) t-statistics. The last column for each regression in Panel B reports the annual estimates of expected rate of return by percentage of buy recommendations. Panel C reports summary means of the differences in estimates across the annual regressions and the related Fama and Macbeth (1973) t-statistics.

REFERENCES

Baginski, S. and J. Wahlen. 2003. Residual income, risk, intrinsic values, and share prices. *The Accounting Review* 78: 327-351.

Barber, B, R. Lehavy, M. McNicholls, and B. Trueman. 2001. Can investors profit from the prophets? Security analyst recommendations and stock returns. *Journal of Finance* 56: 531-563.

Boni, L. and K. Womack. 2002. Wall Street's credibility problem: misaligned incentive and dubious fixes. *Wharton Financial Institutions Center.*

Botosan, C. 1997. Disclosure level and the cost of equity capital. *The Accounting Review* 72: 323-349.

Botosan, C. and M. Plumlee. 2005. Assessing alternative proxies for the expected risk premium. The Accounting Review 80: 21-54.

Botosan, C., M. Plumlee, Y. Xie. 2004. The role of information precision in determining the cost of equity capital. *Review of Accounting Studies* 9: 233-259.

Bradshaw, M. 2004. How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review* 79: 25-50.

Brown, L. 1993. Earnings forecasting research: Its implications for capital markets research. *International Journal of Forecasting* 9: 295-320.

Claus, J. and J. Thomas. 2001. Equity risk premium as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stocks. *Journal of Finance* 56: 1629-1666.

Das, S., B. Levine, and K. Sivaramankrishnan. 1998. Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review* 73: 277-294.

Daske, H. 2006. Economic benefits of adopting IFRS or US-GAAP – Have the expected costs of equity capital really decreased? *Journal of Business, Finance, and Accounting.* Forthcoming.

Dhaliwal, D., L. Krull, O. Li, and W. Moser. 2005. Dividend taxes and implied cost of capital. *Journal of Accounting Research.* 43: 675-715.

Dugar, A., and S. Nathan. 1995. The effect of banking relationships on financial analysts' earnings investment recommendations. *Contemporary Accounting Research* 12: 131-160.

Eames, M., S. Glover, and J. Kennedy. 2002. The association between broker recommendations and broker-analysts' earnings forecasts. *Journal of Accounting Research* 40: 85-104.

Easton, P. 2004. PE ratios, PEG ratios, and estimating the implied expected rate of return on equity capital. *The Accounting Review* 79: 73-96.

Easton, P. 2006. Use of forecasts of earnings to estimate and compare cost of capital across regimes. *Journal of Business, Finance, and Accounting* 33: 374-394.

Easton, P, and S. Monahan. 2005. An evaluation of accounting-based measures of expected returns. *The Accounting Review* 80: 501-538.

Easton, P., G. Sommers, and M. Zmijewski. 2006. Effects of fraud on expectations of returns on equity investment and growth. University of Notre Dame. Working paper.

Easton, P., G. Taylor, P. Shroff, and T. Sougiannis. 2002. Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. *Journal of Accounting Research* 40: 657-676.

Fama, E. and K, French. 1992. The cross-section of expected returns. *Journal of Finance* 47: 427-465.

Fama, E. and J. MacBeth. 1973. Risk, return and equilibrium – Empirical tests. *The Journal of Political Economy* 81: 607-636.

Francis, J., I. Khurana, and R. Periera. 2005. Disclosure incentives and effects on cost of capital around the world. *Accounting Review* 80: 1125-1163.

Francis, J., R. LaFond, P. Olsson and K. Schipper. 2004. Costs of equity and earnings attributes. *Accounting Review* 79: 967-1011.

Frankel, R. and C. Lee. 1998. Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics* 25: 283-319.

Gebhardt, W., C. Lee, and B. Swaminathan. 2001. Towards an ex-ante cost of capital. *Journal of Accounting Research* 39: 135-176.

Gode, D. and P. Mohanram. 2003. Inferring the cost of capital using the Ohlson-Juettner model. *Review of Accounting Studies* 8: 399-431.

Groysberg, B., P. Healy, C. Chapman, and Y. Gui. 2006. Do buy-side analysts out-perform the sell-side. *Journal of Accounting and Economics*. Forthcoming.

Guay, W., S. Kothari, S. Shu. 2005. Properties of implied cost of capital using analysts' forecasts. University of Pennsylvania, Wharton School. Working paper.

Hail, L. and C. Leuz. 2006. International differences in the cost of equity capital: Do legal institutions and securities regulation matter? *Journal of Accounting Research*. Forthcoming.

Hribar, P. and N. Jenkins. 2004. The effect of accounting restatements on earnings revisions and the estimated cost of capital. *Review of Accounting Studies* 9: 337-356.

Kutsoati, E. and D. Bernhardt. 1999. Can relative performance evaluation explain analysts' forecasts of earnings. Tufts University. Working paper.

Lee, C., J. Myers, and B. Swaminathan. 1999. What is the intrinsic value of the Dow? *Journal of Finance* Volume 54 Issue 5: 1693-1741.

Lin, H. 1994. Security analysts' investment recommendations. Stanford University. Ph.D. dissertation.

Mendenhall, R. 1991. Evidence on the possible underweighting of earnings information. *Journal of Accounting Research* 29: 170-179.

Michaely, R and K. Womack. 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12: 653-686.

O'Brien, P. 1988. Analysts' forecasts as earnings recommendations. *Journal of Accounting and Economics* 10: 53-83.

O'Brien, P. 1993. Earnings forecasts and the macroeconomy. University of Michigan. Working Paper.

O'Hanlon, J. and A. Steele. 2000. Estimating the equity risk premium using accounting fundamentals. *Journal of Business Finance and Accounting* 27: 1051-1084.

Ohlson, J. 1989. Accounting earnings, book value and dividends: the theory of the clean surplus equation (part I). Reproduced in R. Brief and K. Peasnell (eds.) *Clean Surplus: A Link Between Accounting and Finance* (Garland), 167-227.

Richardson, S., S. Teoh, and P. Wysocki. 2004. The walk-down to beatable analyst forecasts: the role of equity issuances and insider trading incentives. *Contemporary Accounting Research*, 21, 885-924.

Williams, M. 2004. Discussion of "The role of information precision in determining the cost of equity capital." *Review of Accounting Studies* 9: 261-264.

Zitzewitz, E. 2002. Regulation fair disclosure and the private information of analysts. Stanford University. Working paper.

The Economist

Buttonwood

Hung, drawn and first-quartered

The trend in corporate profits in America is worrying

Apr 4th 2015 | From the print edition

Economist.com

Correction to this article (#correction)

IF ONLY America could abolish the first quarter, its economy would look so much better. In 2014 a cold snap triggered by the "polar vortex" caused GDP to fall by 2.1% at an annualised rate. This time round, more cold weather, a decline in oil drilling and a labour dispute at west-coast ports is causing growth estimates to be revised down once more.

Figures for manufacturing output, durable-goods orders, housing starts and retail sales have all been weaker than expected. The consensus forecast for growth in the first quarter is 1.4% at an annualised rate. But a nimbler model created by the Atlanta Federal Reserve points to just 0.2%—barely any growth at all.

A weak first-quarter number will make life even harder for the Federal Reserve, which has hinted that it might push up interest rates later this year. Inflation is running at zero, so the justification for higher rates would look very flimsy if the growth outlook was faltering too.

However, as in 2014, most economists expect the first-quarter figures to be a blip, with activity rebounding in the rest of the year. Low oil prices should be a boost to spending; consumer-confidence figures released on March 31st showed an upturn. The employment figures for March, which are due to be published on April 4th, will be the next big test of the economy's strength. The Fed has indicated that the labour market may trigger a decision to raise rates; if unemployment falls much below the current rate of 5.5%, wage pressures might start to appear. Strong figures on job creation have generally belied the weak tone of numbers on durable-goods orders and retail sales.

Some investors may be inclined to take a relaxed view of the Fed's dilemma. After all, if the economy is strong enough to allow the central bank to raise interest rates, that would be good news; and if the economy isn't strong enough, then investors will continue to enjoy the benefit of low rates. However, that rosy view is being somewhat undermined by the recent weakness in corporate profits.

After plunging in 2008, profits rebounded strongly, hitting their highest levels as a proportion of GDP since the second world war. That trend may be coming to an end. Corporate profits in America fell by 1.6% in the fourth quarter of 2014, according to the Bureau of Economic Analysis, and were 6.4% lower than in the same quarter of 2013. Those figures do not translate directly into the profits of S&P 500 companies, many of which are multinationals: their earnings per share rose at an annual rate of 7.8% in the fourth quarter, with the help of buy-backs, which spread profits among a smaller number of shares (see chart).

However, the dollar's surge in 2015 is dragging down earnings forecasts for the current year: earnings per share for S&P 500 firms are now expected to rise by only 2.6%. Three factors are at work. First, the strong dollar is reducing the value of profits earned in other currencies. Second, those foreign profits are being squeezed by a slowdown in developing economies. And third, the fall in the oil price is battering the profits of the energy sector.

Wall Street analysts tend to be optimistic when it comes to medium-term profit projections. After a sluggish 2015, they think 2016 will be a bumper year, with earnings per share rising by 12.9%. That allows them to claim that the market looks cheap when future earnings growth is taken into account: using their 2016 forecasts, the market is on a prospective price-earnings ratio of 15.3.

But if the market is compared with past earnings numbers, the picture looks rather different. The cyclically-adjusted price-earnings ratio (which averages profits over ten years) is currently 27.9, according to Robert Shiller of Yale University. The long-term average is 16.6. The sluggish performance of profits may explain why the American stockmarket has struggled to make progress so far this year.

Investors in the rest of the world should also be concerned about weak economic data. There have been 29 instances of monetary easing by central banks around the world in the past five months, an indication that monetary authorities are worried about growth. Low government-bond yields and falling commodity prices are further signals of poor economic momentum.

Although there have been tentative signs that the euro-zone economy is recovering, the world has been very reliant on China and America in recent years. China's growth rate has slowed to 7% or so from the double-digit rates regularly seen in the past decade. If America's growth slows as well, the global economy may find itself becalmed.

Economist.com/blogs/buttonwood (http://www.economist.com/blogs/buttonwood)

Correction: This piece originally said that GDP fell by 2.9% in the first quarter of 2014 rather than 2.1%. We regret the error.

From the print edition: Finance and economics

THE WALL STREET JOURNAL.

Study Suggests Bias in Analysts' Rosy Forecasts

By ANDREW EDWARDS
March 21, 2008; Page C6

Despite an economy teetering on the brink of a recession -- if not already in one -- analysts are still painting a rosy picture of earnings growth, according to a study done by Penn State's Smeal College of Business.

The report questions analysts' impartiality five years after then-New York Attorney General Eliot Spitzer forced analysts to pay $1.5 billion in damages after finding evidence of bias.

"Wall Street analysts basically do two things: recommend stocks to buy and forecast earnings," said J. Randall Woolridge, professor of finance. "Previous studies suggest their stock recommendations do not perform well, and now we show that their long-term earnings-per-share growth-rate forecasts are excessive and upwardly biased."

The report, which examined analysts' long-term (three to five years) and one-year per-share earnings expectations from 1984 through 2006 found that companies' long-term earnings growth surpassed analysts' expectations in only two instances, and those came right after recessions.

Over the entire time period, analysts' long-term forecast earnings-per-share growth averaged 14.7%, compared with actual growth of 9.1%. One-year per-share earnings expectations were slightly more accurate: The average forecast was for 13.8% growth and the average actual growth rate was 9.8%.

"A significant factor in the upward bias in long-term earnings-rate forecasts is the reluctance of analysts to forecast" profit declines, Mr. Woolridge said. The study found that nearly one-third of all companies experienced profit drops over successive three-to-five-year periods, but analysts projected drops less than 1% of the time.

The study's authors said, "Analysts are rewarded for biased forecasts by their employers, who want them to hype stocks so that the brokerage house can garner trading commissions and win underwriting deals."

They also concluded that analysts are under pressure to hype stocks to generate trading commissions, and they often don't follow stocks they don't like.

Write to Andrew Edwards at andrew.edwards@dowjones.com

EQUITY RISK PREMIUM FORUM

NOVEMBER 8, 2001

Forum Participants

Martin L. Leibowitz (*Forum Chair*)

Robert D. Arnott
Clifford S. Asness
Ravi Bansal
John Y. Campbell
Peng Chen, CFA
Bradford Cornell
William N. Goetzmann
Brett Hammond
Campbell R. Harvey
Roger G. Ibbotson
Rajnish Mehra
Thomas K. Philips
William Reichenstein, CFA
Stephen Ross
Robert J. Shiller
Jeremy Siegel
Kevin Terhaar, CFA
Richard H. Thaler
J. Peter Williamson

Editorial Staff	
Kathryn Dixon Jost, CFA *Editor*	**Jaynee M. Dudley** *Production Manager*
Bette Collins *Book Editor*	**Kathryn L. Dagostino** *Production Coordinator*
Christine E. Kemper *Assistant Editor*	**Kelly T. Bruton** **Lois A. Carrier** *Composition*

Biographies

ROBERT D. ARNOTT is managing partner and CEO at First Quadrant, LP. Previously, he served as equity strategist at Salomon Brothers and as president and chief investment officer at TSA Capital Management. Mr. Arnott co-edited the first and second editions of *Active Asset Allocation* and *Style Management* and is the author of numerous articles. He has received four Graham and Dodd Awards of Excellence for articles published in the *Financial Analysts Journal* and is a member of the editorial boards of the *Journal of Portfolio Management, Journal of Investing,* and *Journal of Wealth Management*. Mr. Arnott holds degrees in economics, applied mathematics, and computer science from the University of California.

CLIFFORD S. ASNESS is president and managing principal at AQR Capital Management, LLC. Previously, he was a managing director and director of quantitative research at Goldman, Sachs & Company. Mr. Asness's articles on a variety of financial topics have appeared in many journals, including the *Journal of Portfolio Management*, *Financial Analysts Journal*, and *Journal of Banking & Finance*. In 2000, he received a Graham and Dodd Award of Excellence from the *Financial Analysts Journal*. Mr. Asness serves on the editorial board of the *Journal of Portfolio Management* and on the governing board of the Courant Institute of Mathematical Finance. He holds a B.S. in economics from The Wharton School at the University of Pennsylvania, a B.S. in engineering from the Moore School of Electrical Engineering, and an M.B.A. and a Ph.D. from the University of Chicago.

RAVI BANSAL is associate professor of finance at Fuqua School of Business at Duke University. Professor Bansal has published numerous articles in such finance journals as the *Journal of Finance*, *Review of Financial Studies*, and *Journal of Econometrics*. He is associate editor of the *Journal of Financial Econometrics* and *Journal of Securities Markets*. He holds an M.S and a Ph.D from the Carnegie Mellon University Graduate School of Industrial Administration and a B.A. and an M.A. in economics from the Delhi School of Economics at the University of Delhi, India.

JOHN Y. CAMPBELL is Otto Eckstein Professor of Applied Economics at Harvard University. Previously, he was Class of 1926 Professor of Economics and Public Affairs at Princeton University. He is the author of numerous articles in financial journals and was awarded the Smith Breeden Prize from the *Journal of Finance*, the Roger F. Murray Prize from the Institute for Quantitative Research in Finance and the Paul A. Samuelson Certificate of Excellence from TIAA-CREF. He holds a B.A. from Corpus Christi College at the University of Oxford and a Ph.D. and an M.Phil. from Yale University.

PENG CHEN, CFA, is vice president and director of research at Ibbotson Associates. He conducts research projects on asset allocation, portfolio risk measurement, nontraditional assets, and global financial markets. His writings have appeared in various journals, including the *Journal of Portfolio Management*, *Journal of Association of American Individual Investors*, and *Consumer Interest Annual*. Mr. Chen holds a bachelor's degree in industrial management engineering from Harbin Institute of Technology and master's and doctorate degrees in consumer economics from Ohio State University.

BRADFORD CORNELL is professor of financial economics at the Anderson Graduate School of Management at the University of California at Los Angeles and president of FinEcon. Previously, Professor Cornell was vice president and director of the Securities Litigation Group at Economic Analysis Corporation. He has also taught at the California Institute of Technology, the University of Southern California, and the University of Arizona. Professor Cornell has served as an associate editor for a variety of scholarly and business journals and is the author of several books and numerous articles published in various finance journals. Professor Cornell holds an A.B. in physics, philosophy, and psychology, an M.S. in statistics, and a Ph.D. in financial economics from Stanford University.

WILLIAM N. GOETZMANN is Edwin J. Beinecke Professor of Finance and Management Studies and director of the International Center for Finance at the Yale School of Management. Previously, he was associate professor of finance at Yale School of Management, assistant professor of finance at Columbia Business School, and a lecturer at Ibbotson Associates. Professor Goetzmann has also served as director of the Museum of Western Art and as a writer and producer of PBS documentaries. He is the author or co-author of numerous articles on finance and real estate and is co-editor or editorial board member for several journals. Professor Goetzmann holds a B.A., an M.A., and an M.Phil. from Yale College, an M.P.P.M. from Yale School of Management, and a Ph.D. in operations research with a specialty in finance from Yale University.

BRETT HAMMOND is director of corporate projects at TIAA-CREF Investment Management, LLC.

CAMPBELL R. HARVEY is J. Paul Sticht Professor of International Business at the Fuqua School of Business at Duke University and a research associate for the National Bureau of Economic Research. Professor Harvey is an internationally recognized expert in portfolio management and global risk management. His work on the implications of changing risk and the dynamics of risk premiums for tactical asset allocation has been published in the top academic and practitioner journals. He is editor of the *Review of Financial Studies*, co-editor of the *Emerging Markets Review*, and an associate editor for a number of journals. Professor Harvey received the 1993–1994 Batterymarch Fellowship and has received four Graham and Dodd Awards for Excellence in financial writing from the *Financial Analysts Journal*. He holds a B.A. from Trinity College at the University of Toronto, an M.B.A. from York University, and a Ph.D. from the University of Chicago.

ROGER G. IBBOTSON is chair and founder of Ibbotson Associates and the author of numerous books and articles, including the annual *Stocks, Bonds, Bills, and Inflation Yearbook*. Mr. Ibbotson is a frequent lecturer at universities, academic conferences, business conferences, and other forums. He is associate editor of the *Journal of Applied Corporate Finance* and a member of the editorial board of the *Financial Analysts Journal*. He received a *Financial Analysts Journal* Graham and Dodd Award of Excellence in 1980, 1982, 1984 and 2000. Mr. Ibbotson holds a bachelor's degree in mathematics from Purdue University, a master's degree in finance from Indiana University, and a Ph.D. in finance and economics from the University of Chicago Graduate School of Business, where he taught for more than 10 years and served as executive director of the Center for Research in Security Prices.

MARTIN L. LEIBOWITZ is vice chair and chief investment officer at TIAA-CREF Investment Management, LLC, where he is responsible for the overall management of all TIAA-CREF investments. Previously, he was a managing director, director of research, and member of the executive committee with Salomon Brothers. Mr. Leibowitz has authored several books and more than 130 articles, 9 of which have received a *Financial Analysts Journal* Graham and Dodd Award of Excellence. He was the recipient of AIMR's Nicholas Molodovsky Award in 1995 and the James R. Vertin Award in 1998. In 1995, he received the Distinguished Public Service Award from the Public Securities Association and became the first inductee into the Fixed Income Analysts Society's Hall of Fame. Mr. Leibowitz is a trustee of the Carnegie Corporation, the Institute for Advanced Study at Princeton, and the Research Foundation of AIMR. He holds a B.A. and an M.S. from the University of Chicago and a Ph.D. in mathematics from the Courant Institute of New York University.

RAJNISH MEHRA is professor of finance at the University of California at Santa Barbara and visiting professor of finance at the University of Chicago Graduate School of Business. Previously, he was visiting professor of finance at the Sloan School of Management at the Massachusetts Institute of Technology and assistant and associate professor at the Graduate School of Business at Columbia University. He is the author of numerous articles published in such journals as *Econometrica*, the *Review of Economic Studies*, and the *Journal of Monetary Economics*, and his writings have been included in several books. He holds a B.Tech. in electrical engineering from the Indian Institute of Technology in Kanpur, India, an M.S. in computer science from Rice University, and an M.S. in industrial administration and a Ph.D. in finance from the Graduate School of Industrial Administration at Carnegie-Mellon University.

THOMAS K. PHILIPS is chief investment officer at Paradigm Asset Management Company, where he is responsible for all aspects of the investment process, including the development of new products and the enhancement of existing ones. Previously, he was managing director at RogersCasey and worked at IBM Corporation in research and active equity management for the IBM Thomas J. Watson Research Center and the IBM Retirement Fund. Mr. Philips is the author of several articles and book chapters on topics in finance, engineering, and mathematics. He holds an M.S. and a Ph.D. in electrical and computer engineering from the University of Massachusetts at Amherst.

WILLIAM REICHENSTEIN, CFA, holds the Pat and Thomas R. Powers Chair in Investment Management at Baylor University and is a consultant to TIAA-CREF. He is an associate editor of the *Journal of Investing*, on the editorial board of *Journal of Financial Education*, on the editorial review board of the *Journal of Financial Planning*, and on the editorial advisory board of the *Journal of Applied Business*

Research. Professor Reichenstein has written more than 50 articles for professional and academic journals and is a frequent contributor to the *Financial Analysts Journal*, *Journal of Portfolio Management*, and *Journal of Investing*. His work has been discussed in the *Wall Street Journal, Barron's, Forbes, Smart Money*, and elsewhere, and he is frequently quoted in newspapers throughout the United States. Professor Reichenstein holds a B.A. in mathematics from St. Edward's University and a Ph.D. in economics from the University of Notre Dame.

STEPHEN ROSS is Franco Modigliani Professor of Financial Economics at the Sloan School of Management at the Massachusetts Institute of Technology. He is also a principal of Roll and Ross Asset Management Corporation. Professor Ross, a widely published author in finance and economics, is best known as the inventor of Arbitrage Pricing Theory and the Economic Theory of Agency and as the co-discoverer of risk-neutral pricing and the binomial model for pricing derivatives. He is the author of *Corporate Finance*, which is in its fourth edition.

ROBERT J. SHILLER is Stanley B. Resor Professor of Economics at the Cowles Foundation for Research in Economics at Yale University. He is co-founder of Case Shiller Weiss, an economics research and information firm, and also a co-founder of Macro Securities Research, a firm that promotes securitization of unusual risks. Professor Shiller is the author of numerous articles on financial markets, behavioral economics, macroeconomics, real estate, statistical methods, and public attitudes, opinions, and moral judgments regarding markets. He is the author of *Market Volatility*, a mathematical and behavioral analysis of price fluctuations in speculative markets, and *Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks*, which won the 1996 Paul A. Samuelson Award from TIAA-CREF. Professor Shiller's recent book, *Irrational Exuberance*, won the Commonfund Prize for 2000. He holds a Ph.D. in economics from the Massachusetts Institute of Technology.

JEREMY SIEGEL is Russell E. Palmer Professor of Finance at The Wharton School of the University of Pennsylvania, where he has been a member of the faculty since 1976. Previously, he taught at the University of Chicago Graduate School of Business. Professor Siegel has received many awards and citations for his research and excellence in teaching, including a Graham and Dodd Award of Excellence from the *Financial Analysts Journal* and the 2000 Peter Bernstein and Frank Fabozzi Award from the *Journal of Portfolio Management*. In 1994, he received the highest teaching rating in a worldwide ranking of business school professors conducted by *BusinessWeek*. Professor Siegel has published articles in numerous finance journals and is the author of the book *Stocks for the Long Run* and co-author of *Revolution on Wall Street*. He holds a B.A. from Columbia University and a Ph.D. from the Massachusetts Institute of Technology.

KEVIN TERHAAR, CFA, is director of risk management and specialized investments at Brinson Partners. Previously, Mr. Terhaar was manager of investments for a trust office, where his responsibilities included asset allocation and investment analysis. He has published articles on a number of investment topics and is the co-author of "Maintaining Consistent Global Asset Views," which appeared in the *Financial Analysts Journal* in 1998. Mr. Terhaar holds a B.A. and an M.A. from the University of Virginia.

RICHARD H. THALER is Robert P. Gwinn Professor of Behavioral Science and Economics and director of the Center for Decision Research at the University of Chicago Graduate School of Business. He is also a research associate at the National Bureau of Economic Research and co-director, with Robert Shiller of the Behavioral Economics Project, funded by the Russell Sage Foundation. He has published articles on finance in such journals as the *Journal of Economic Perspectives*, *Journal of Finance*, *Journal of Portfolio Management*, and *American Economic Review*. Professor Thaler holds a B.A. from Case Western Reserve University and an M.A. and a Ph.D. from the University of Rochester.

J. PETER WILLIAMSON is the Laurence F. Whittemore Professor of Finance, Emeritus, at the Amos Tuck School of Business Administration at Dartmouth College, where he has been a member of the faculty since 1961. He is the author of more than a dozen books and monographs and numerous articles. Professor Williamson has served as a consultant and as an expert witness on cost of capital and capital structure in approximately 100 federal and state cases. He has prepared for publication summaries of all the presentations made at the semi-annual seminars of the Institute for Quantitative Research in Finance for 26 years. Professor Williamson holds a B.A. from the University of Toronto and an M.B.A., an LL.B., and a D.B.A. from Harvard University.

INTRODUCTION

Martin L. Leibowitz (*Forum Chair*)
TIAA-CREF
New York City

Our goal here today is to foster a very candid discussion of the many facets of the equity risk premium. Generally, the risk premium is thought of as the incremental return of certain equity market components relative to certain fixed-income components. Even when these two measures are clarified, however, which they often are not, considerable ambiguity can remain as to just what we're talking about when we talk about the risk premium. Are we talking about a premium that has been historically achieved, a premium that is the ongoing expectation of market participants, an analytically determined forecast for the market, or a threshold measure of required return to compensate for a perceived level of risk? All of these measures can be further parsed out as reflections of the broad market consensus, the opinions of a particular individual or institution, or the views of various market cohorts looking at specific and very different time horizons.

As for the issue of the risk premium as uncertainty, we often see the risk premium defined as an extrapolation of historical volatility and then treated as some sort of stable parameter over time. A more comprehensive (and more difficult) approach might be to view the risk premium as a sufficient statistic unto itself, a central value that is tightly embedded in an overall distribution of incremental returns. From this vantage point, we would then look at the entire risk premium distribution as an integrated dynamic, one that continually reshapes itself as the market evolves.

With the enormous variety of definitions and interpretations, the risk premium may seem to be the ultimate "multicultural" parameter and our forum today may have the character of a masked ball within the Tower of Babel. However, every one of us here does know and understand the particular aspect of the risk premium that we are addressing in our work. And I hope that we can communicate that clarity even as we tackle the many thorny questions that surround this subject. The risk premium is a concept that is so central to our field of endeavor that it might properly be called the financial equivalent of a cosmological concept.

Theoretical Foundations I

Richard H. Thaler

Graduate School of Business
University of Chicago
Chicago

One of the puzzles about the equity risk premium is that in the U.S. market, the premium has historically been much greater than standard finance theory would predict. The cause may lie in the mismatch between the actual asset allocation decisions of investors and their forecasts for the equity risk premium. In this review of the theoretical explanations for this puzzle, two questions are paramount: (1) How well does the explanatory theory explain the data? (2) Are the behavioral assumptions consistent with experimental and other evidence about actual behavior? The answers to both questions support the theory of "myopic loss aversion"—in which investors are excessively concerned about short-term losses and exhibit willingness to bear risk based on their most recent market experiences.

A good place to start consideration of what the equity risk premium should theoretically be is a discussion of the risk premium puzzle: The equity risk premium in the U.S. market has historically been much bigger than standard finance theory would predict. Based on the familiar Ibbotson Associates (2001) data of the long-term historical return to U.S. stocks, T-bonds, and T-bills, if you had invested $1 in the stock market at the end of 1925 (with dividends reinvested), you would now have more than $2,500; if you had put $1 in T-bonds, you would have about $49; and if you had put $1 in T-bills, you would have only $17. These differences are much too large to be explained by any reasonable level of risk aversion.

The Puzzle

The formal puzzle, which was posed by Mehra and Prescott (1985), is that, on the one hand, if you ask, "How big a risk premium should we expect?" the standard economic model (assuming expected-utility-maximizing investors with standard additively separable preferences and constant relative risk aversion, A) provides a much smaller number than is historically true, but if you ask, "How risk averse would investors have to be to demand the equity risk premium we have seen?" (that is, how large does A have to be to explain the historical equity premium), the answer is a very large number—about 30. Mehra and Prescott's response was that 30 is too large a number to be plausible.

Why? What does a coefficient of relative risk aversion of 30 mean? If I proposed to you a gamble in which you have a 50 percent chance that your wealth will double and a 50 percent chance that your wealth will fall by half, how much would you pay to avoid the chance that you will lose half your wealth? If you have a coefficient of relative risk aversion equal to 30, you would pay 49 percent of your wealth to avoid a chance of losing half your wealth, which is ridiculous. And that is why I believe that investors do not have a coefficient of relative risk aversion of 30.

Another way to think about this puzzle is that for reasonable parameters (and theorists argue about what those are), we would expect an annual risk premium for stocks over bonds of 0.1 percent (10 basis points).

In the Mehra–Prescott model, the coefficient of relative risk aversion, A, is also the inverse of the elasticity of intertemporal substitution, so a high value of A implies an extreme unwillingness to substitute consumption tomorrow for consumption today, which implies a long regime of high interest rates. We have not, however, observed high interest rates for extended periods of time. Historically, the risk-free rate has been low, barely positive for much of the 20th century. Therefore, part of the risk premium puzzle is the "risk-free-rate puzzle": Why do we not see very high interest rates if investors are so risk averse?

How do we resolve these puzzles? One answer is to "blame the data"—for example, survivorship bias. The returns in the U.S. equity market have been particularly favorable, which may be simply the product of good luck. In other words, some markets have collapsed and disappeared. So, we should not focus all our attention on one market in one period; one market can go awry.

My view is that if we can worry about stock markets going awry, we had better also worry about bond markets going awry. For example, over the long run, bond investors have experienced bad periods of hyperinflation. Bond investors have been wiped out by hyperinflation just as stock investors have been wiped out by crashes. So, if we are going to consider the effect of survivorship bias on the data, we need to look at both sides of the equation—stock and bond returns—which brings us back to a puzzle. If you adjust *both* returns for risk, you still end up with a puzzle.

The part of the puzzle that I want to stress is the contrast between investor investments and investor expectations. I am a behaviorist, and the behavior I find puzzling is how investor expectations fit with their investments.

Throughout the 1980s and 1990s, investors had expectations of a big equity premium, typically in the range of 4 percent to 7 percent. **Table 1** provides the results of a survey of fund managers on their forecasts for U.S. security returns at two points in time almost 10 years apart. Note that investor estimates of the equity risk premium fall into the 4–6 percent range in both years.

Other evidence comes from surveys of forecasts of the 10-year equity risk premium over the last decade (for example, Welch's 2000 survey of economists); again, the estimates are substantial. A problem with such surveys, of course, is that we never know the question the people were really answering. For example, most respondents, including economists, do not know the difference between the arithmetic and the geometric return, and this confusion can skew the results. So, we cannot know precisely what such surveys show, but we can know that the estimates of the equity risk premium are big numbers compared with an estimate of 0.1 percent.

Table 1. Forecasted Returns: Survey of Fund Managers (N = 395)

Fund/Premium	1989	1997
90-day T-bills	7.4%	4.7%
Bonds	9.2	6.9
S&P 500	11.5	10.4
S&P 500 – T-bills	4.1	5.7

Source: Greenwich Associates.

Thaler's Equity Premium Puzzle

The real puzzle is a mismatch between the allocations of investors and their forecasts for the equity risk premium. Many long-term investors—individuals saving for retirement, endowments, and pension fund managers—think the long-term equity risk premium is 4–5 percent or higher yet still invest 40 percent of their wealth in bonds. This phenomenon is the real puzzle.

One version of this puzzle is "Leibowitz's Lament." In a former life, Marty Leibowitz was a bond guy at Salomon Brothers. As a bond guy, his job was to give investors a reason to buy bonds. The numbers Marty was crunching in 1989 for the wealth produced by $1 in stocks versus the wealth produced by $1 invested in bonds could have been those from the Ibbotson Associates studies. The historical risk premium was 6.8 percent, which made the return numbers ridiculous. Marty's analysis showed that if we assume investors may lever, the correct asset allocation at that time would have been at least 150 percent in equities. The puzzle is that investors did not invest this way then and do not do so now.

Theoretical Explanations

Many explanations for the puzzle have been offered, and all the theoretical explanations so far proposed are behavioral—in the sense that they build on the Mehra–Prescott model and then make some inference about investor preferences. In most of these models, the investors make rational choices but their preferences are still slightly different from ones traditionally considered normal.

Epstein and Zin (1989) broke the link that A is equal to the coefficient of relative risk aversion and the elasticity of intertemporal substitution. With their approach, the standard assumptions of expected utility maximization are destroyed.

Constantinides (1990) introduced the theory of habit formation based on the following postulate: If I'm rich today, then I'm more miserable being poor tomorrow than if I'd always been poor. A similar theory of habit formation, the approach of Abel (1990), is based on the concept of "keeping up with the Joneses." Perhaps the leading model at the moment, however, is that of Campbell and Cochrane (1995, 1999), which combines the idea of habit formation with high levels of risk aversion. Together, these behavioral theories appear to explain some, but not all, of the data—including the risk-free-rate puzzle.

Benartzi and I (1995) suggested the theory of loss aversion, which is the idea that investors are more sensitive to market changes that are negative than to those that are positive, and the idea of mental accounting, which adds that investors are more sensitive when they are given frequent market evaluations. Combined, loss aversion and mental accounting produce what we called "myopic loss aversion." We explicitly modeled investors as being myopic, in that they think about and care most strongly about the market changes that occur over short periods, such as a year.

Barberis, Huang, and Santos (1996) used the myopic loss aversion model and added another behavioral phenomenon, the "house money effect" (that is, loss aversion is reduced following recent gains), in an equilibrium model. When people are ahead in whatever game they are playing, they seem to be more willing to take risks. I also documented this effect in some experimental work about 10 years ago. I discovered this phenomenon playing poker. If you're playing with people who have won a lot of money earlier in the game, there is no point in trying to bluff them. They are in that hand to stay.

So, we have a long list of possible behavioral explanations for the equity risk premium. How do we choose from them? We should concentrate on two factors. The first factor is how well the models fit the data. The second factor, and it is a little unusual in economics, is evidence that investors actually behave the way the modeler claims they are behaving. On both counts, the myopic loss aversion arguments that Benartzi and I (1995) proposed do well.

First, all the consumption-based models have trouble explaining the behavior of two important groups of investors, namely, pension funds and endowments. And these two groups hold a huge amount of the equity market in the United States.

Second, I do not understand why habit formation would apply to a pension-fund manager or the manager of the Rockefeller Foundation.

Third, explanations based on high levels of risk aversion do not fit the following situation: Consider these gambles. Gamble 1: You have a 50 percent chance to win \$110 and a 50 percent chance to lose \$100. Gamble 2: You have a 50 percent chance to win \$20 million and a 50 percent chance to lose \$10,000. Most people reject Gamble 1 and accept Gamble 2. Now, those two preferences are not consistent with expected utility theory. To be consistent with expected utility theory, if you reject the first gamble, you must also reject the second gamble. This inconsistency between behavior and utility theory is a problem for all the models except those that incorporate loss aversion and "narrow framing." In narrow framing, people treat gambles one at a time.

In Thaler, Tversky, Kahneman, and Schwartz (1997), we reported on some experiments to determine whether investors actually behave the way our myopic loss aversion model says they do. In the first experiment, we sat participants down at a terminal and told them, "You are a portfolio manager, and you get to choose between two investments, A and B." One choice was stocks, and the other was bonds, but they were not told that. They were simply shown each investment's returns for the investment period just completed. At the end of every period, the pseudo portfolio managers were instructed to invest their money for the upcoming period based only on the prior-period returns for A and for B. So, they made an asset allocation decision every period. The participants were paid based on the amount of wealth their portfolio had earned at the end of the experiment.

To test the effect of how often investors receive feedback, in various runs of the experiment, we manipulated "how often" the participants were able to look at the return data. In the learning period, the participants learned about the risk and returns of the investments over time. One group of participants received feedback the equivalent of every six weeks, which led to a lot of decision making. Another group made decisions only once a year. So, the first group was working in a condition of frequent evaluation, whereas the second group was receiving exactly the same random feedback as the first one but the returns for the first eight periods were collapsed into a single return. A third group was given a five-year condition. We also had an "inflated monthly" condition in which we increased returns by a constant over the 25-year period that was sufficient to create periods with never any losses. Over the 25 years, 200 decisions were being made in the most frequent condition and 5 in the least frequent condition.

When that part of the experiment was completed and the participants had enjoyed plenty of opportunity to learn the distribution patterns, we instructed them to make one final decision for the next 40 years. Outcomes were "yoked" to assure that all manipulations had the same investment experience.

Our hypotheses were, first, that more frequent reports would induce more risk aversion, resulting in an increased allocation to bonds and, second, that shifting the returns of both assets up to eliminate losses would make stocks relatively more attractive. **Table 2** presents the results.

Table 2. Effect of Frequency of Feedback: Allocation to Bonds

Feedback Group	Number	Mean
A. Final decisions		
Monthly	21	59.1%
Yearly	22	30.4
Five year	22	33.8
Inflated monthly	21	27.6
B. Decisions during the last five "years"		
Monthly	840	55.0%
Yearly	110	30.7
Five year	22	28.6
Inflated monthly	840	39.9

As you can see, participants involved in the monthly condition (the most frequent decision-making condition), on average, chose to invest 60 percent of their money in bonds. Participants in the yearly condition chose to invest only 30 percent in bonds. The participants made the most money if they chose 100 percent stocks every period.

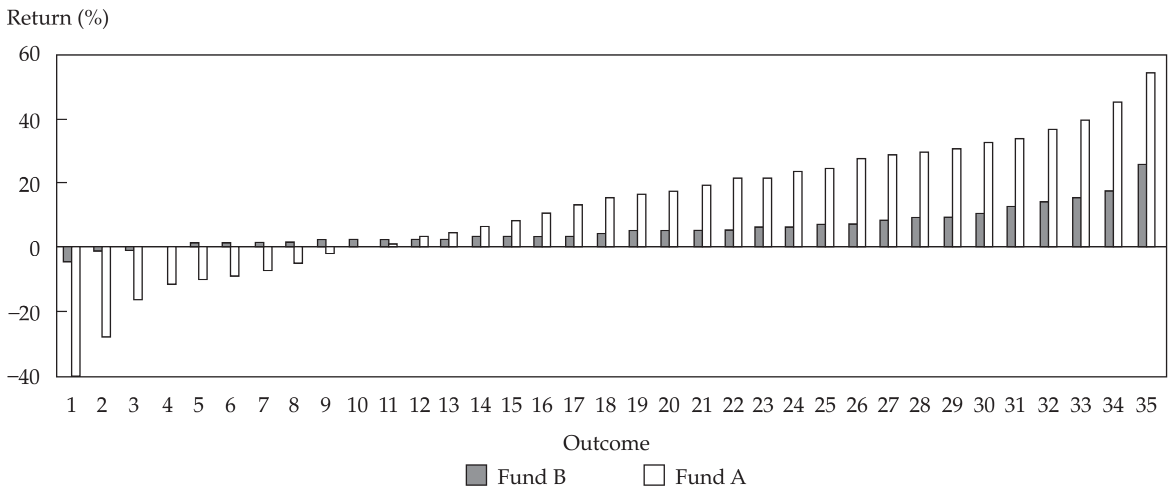
We concluded that the more often investors look at the market, the more risk averse they become, which is exactly what our theory suggests. Loss aversion can be mitigated by forced aggregation (to avoid narrow framing), and learning may be improved by less frequent feedback.

Another set of experiments on myopic loss aversion involved 401(k) participants—specifically, staff among University of Southern California employees who had become eligible for the program in the past year. They were shown return data for Fund A (providing higher returns than Fund B but riskier, equivalent to stocks) and Fund B (equivalent to bonds) and then asked how they would allocate their money. One group was given one-year returns and one group was given 30-year returns. **Figure 1** contains the charts presented in which the historical equity risk premium was used. The figure shows the distribution of periodic rates of return that were drawn from the full sample. That is, if this is the distribution you're picking from, what allocations would you make? Possible outcomes are ranked from worst on the left to best on the right. When we showed the participants the distribution of 1-year rates of return for each asset category (Panel A), the average choice was to invest about 40 percent in stocks. Stocks seemed a bit risky to participants under this scenario. When we showed exactly the same data as compounded annual rates of return for a 30-year investment (Panel B), the participants chose to put 90 percent of their money in stocks. The data are the same in both charts, but the information is presented in a different way. Again, we concluded that the amount investors are willing to invest in stocks depends on how often they look at periodic performance.

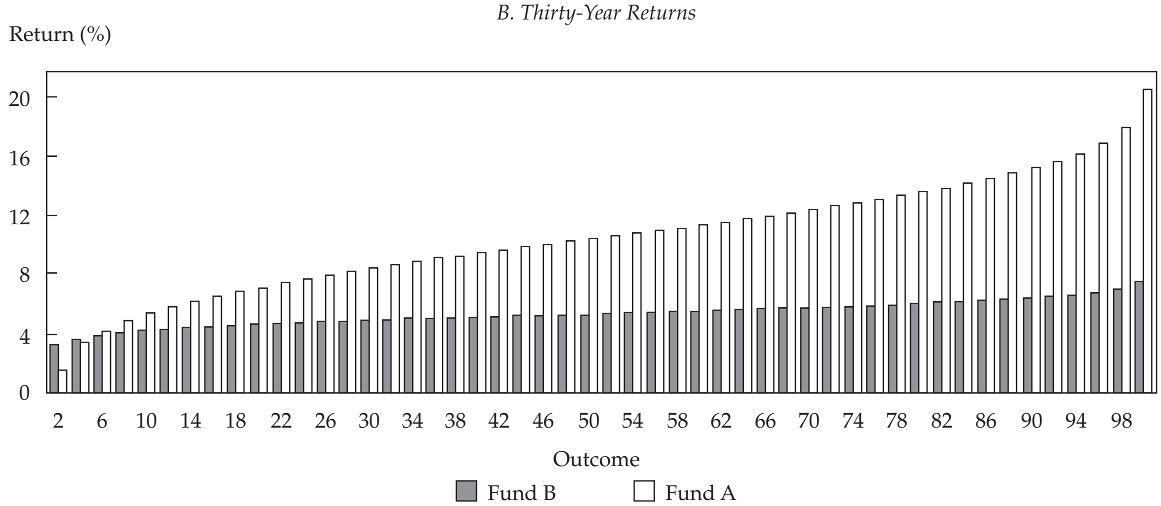
Finally, we showed participants the data with a lower risk premium. As **Figure 2** shows, we divided the equity premium in half. Again, Panel A shows the revised return data for the 1-year periods, and Panel B shows the revised return data for the 30-year period. In this experiment, the participants liked stocks equally well either way they viewed the data. They chose to put about 70 percent of their money in stocks in either scenario. We call this situation a "framing equilibrium." If the equity premium were a number such as 3 percent, investors would put about the same amount of money into the stock market whether they had a long-term perspective or not.

Figure 1. Charts Constructed with Historical Risk Premium of Equity over Five-Year T-Bonds

A. One-Year Returns

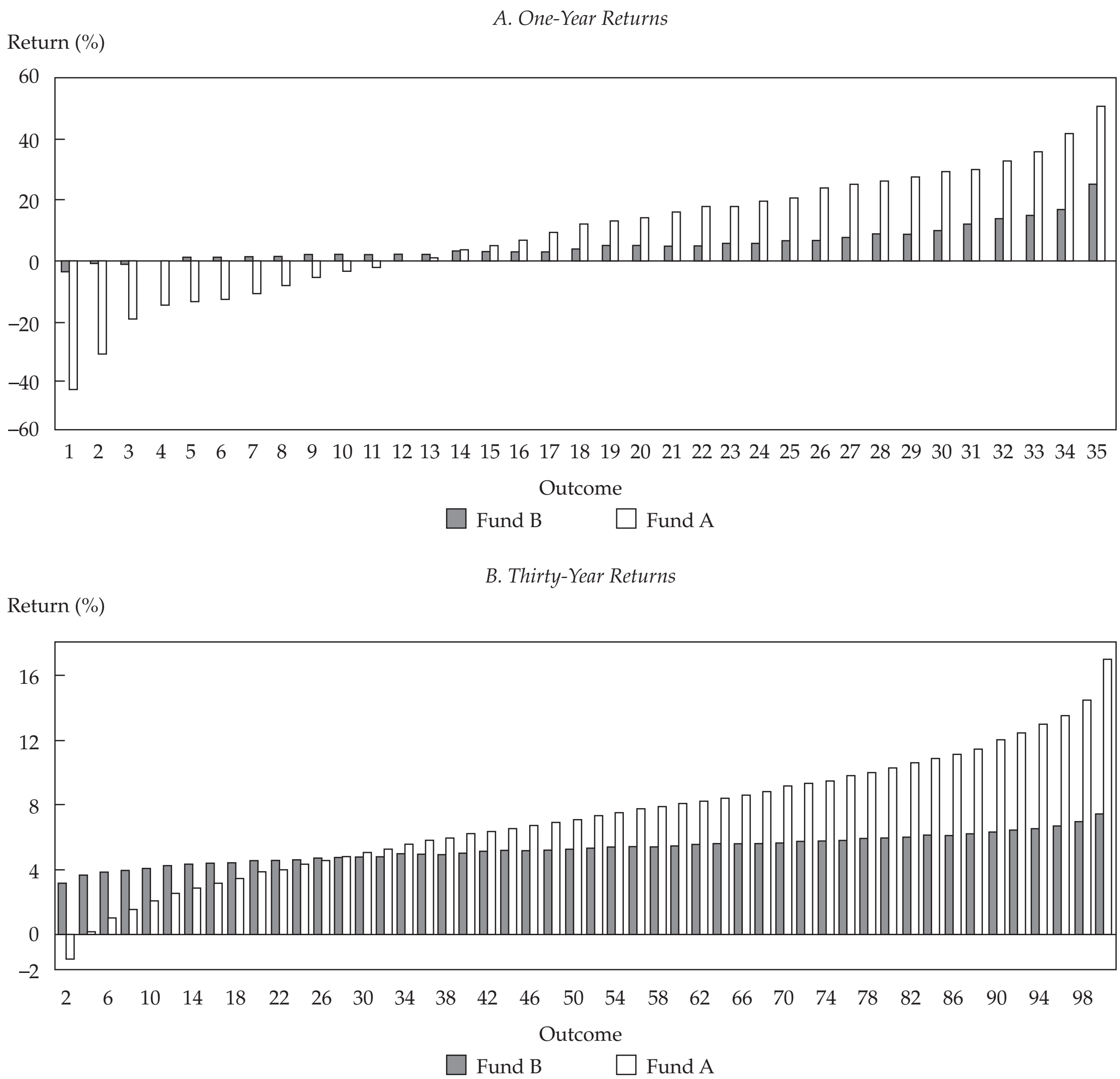


B. Thirty-Year Returns



Notes: Fund A was constructed from the historical returns on the NYSE value-weighted index, and Fund B was constructed from the historical returns on five-year U.S. T-bonds.

Figure 2. Charts Constructed with Half the Historical Risk Premium of Equity over Five-Year T-Bonds



Notes: Fund A was constructed from the historical returns on the NYSE value-weighted index, but 3 percentage points were deducted from the historical annual rates of return on stocks. Fund B was constructed from the historical returns on five-year U.S. T-bonds.

EQUITY RISK PREMIUM FORUM, NOVEMBER 8, 2001

Theoretical Foundations I

Richard H. Thaler

Graduate School of Business
University of Chicago
Chicago

SUMMARY

by Peter Williamson

Amos Tuck School of Business Administration
Dartmouth College, Hanover, New Hampshire

Richard Thaler was the first to speak to the group and the only one dealing essentially with behavioral finance aspects of the equity risk premium puzzle.

He started by discussing the now familiar Ibbotson Associates data from the *2000 Yearbook*,[1] showing the cumulative value of a dollar invested at the end of 1925 in U.S. stocks, T-bonds, and T-bills, with the stock investment (with reinvested dividends) growing to more than $2,500 while a dollar invested in T-bonds grew to about $49 and one invested in T-bills to only $17 by the year 2000. The difference, he said, is much too large to be explained by any reasonable level of risk aversion. Thaler described analysis showing that a 0.1 percent (10 basis point) per year premium for stocks over bonds would be a reasonable equilibrium risk premium; the actual excess return, however, has been more than 7 percent.

In the Mehra–Prescott (1985) model, the constant relative risk aversion, which would have to be 30 to explain the actual historical excess return of stocks, is also the inverse of the elasticity of intertemporal substitution. A value of 30 is very high and implies very high interest rates. But interest rates since 1925 have not been high enough to justify that risk aversion.

What, then, is the explanation for the high historical excess return on stocks? One possibility is high risk coupled with good luck investing in the U.S. stock market. But bond markets are risky too, and if both stock and bond returns are adjusted for high risk, we are still left with an extraordinary gap in historical returns. Furthermore, most surveys in the 1980s and 1990s of "expert" opinion indicated a high expected equity premium, on the order of 4–6 percent. And current surveys give consistent results. Thaler's observation is that many long-term investors who think that the long-term equity premium is 4–5 percent, or higher, still invest 40 percent in bonds, something that is not easily explained. A firm belief in such a premium should have led to at least a 100 percent allocation to stocks. The size of the historical excess equity return versus the size of the expected equity premium present a puzzle.

Most attempts to explain the puzzle focus on behavioral deviations from the standard assumptions of expected utility maximization. Epstein and Zin (1989) broke the link between the coefficient of relative risk aversion and the elasticity of intertemporal substitution. Constantinides (1990) incorporated "habit formation" to posit rising risk aversion with high returns. Others see further reasons for very high risk aversion; they include Benartzi and Thaler (1995) in their myopic risk aversion model.

Thaler put forward a test for choosing among explanations in the form of two questions: (1) How well does the explanatory theory explain the data? (2) Are the behavioral assumptions consistent with experimental and other evidence about actual behavior?

The answers to both questions, he said, support the myopic loss aversion theory. All the consumption-based models have trouble explaining the behavior of pension funds and endowments. A number of experiments presenting people with choices of different gambles have argued against the high-risk-aversion theory. At the same time, experiments posing a problem of allocating funds between stocks and T-bonds have supported myopic loss aversion. Participants in these experiments were asked to allocate money between stocks and bonds after receiving periodic reports on the investment performance of the two classes. It was found that providing more frequent performance feedback induces greater risk aversion and hence reduces commitment to stocks. Shifting

[1] See Ibbotson Associates (2001).

upward and equally the reported returns for both asset classes such that there were no losses for either led to greater investment in stocks.

A further experiment asking subjects to divide retirement funds between stocks and bonds on the basis of the historical excess return on stocks led to a median 40 percent investment in stocks when the subjects were shown distributions of one-year returns and to a median 90 percent investment in stocks when the distributions shown were of 30-year returns.

When the reported excess return on stocks was cut in half from its historical level and the experiment was repeated, the median allocation to stocks was about 70 percent for the annual and for the 30-year distributions. Thaler referred to this condition as "framing equilibrium." The expected risk premium was now such as to remove the influence of the time period of the performance results studied. The equilibrium was reached at an equity premium of about 3 percent.

His three final conclusions were as follows:

- The historical excess return on equities has been surprisingly high.
- Part of the explanation seems to be that investors are excessively concerned about short-term losses.
- Part may be that willingness to bear risk depends on recent experience, both because past gains provide a psychological cushion against future losses and because high returns can create unrealistic expectations about the future.

Theoretical Foundations II

Clifford S. Asness

AQR Capital Management, LLC
New York City

Historically, high P/Es have led to low returns and low P/Es have led to high returns. So, with today's market at historically high P/Es, there is a real need for rescue. This discussion examines three possible ways in which the market might be saved from decline: high and sustained real earnings growth (which is highly unlikely), low interest rates (which help only in the short term), and investor acceptance of lower future rates of return. The last possibility boils down to a choice between low long-term returns forever and very low (crash-type) returns followed by more historically normal returns. The research presented here found some support for the prescription that investors should accept a 6–7 percent nominal stock return, but evidence indicates that investors do not actually think they are facing such low returns.

My talk does not fit neatly into the category of "theoretical foundations," which makes sense; after all, someone who runs a hedge fund is not going to have much to add to the theoretical foundations that underlie our musings about the equity risk premium, certainly not in this crowd!

My first set of data is intended to be an icebreaker. As a beginning, **Figure 1** plots the S&P 500 Index's P/E from 1881 to 2001. From those data, I created seven P/E buckets, or ranges, covering the 1927–2001 period. For each of the buckets, I calculated the median real annualized stock market return for the following decade and the worst return for any decade. **Table 1** provides the results for each range. We can argue about statistical significance, but these numbers are pretty striking. The infallibility of stocks is typically drawn from a 20-year horizon, so I have cheated by using a 10-year horizon. But the infallibility still exists when stocks are bought at low valuation ratios.

The note "Here Be Dragons" is a caution about what might happen with those P/Es of 32.6 to 45.0. It is a saying (similar to "Terra Incognita") once used on old maps for areas not yet visited. The highest P/E, about 45, was reached in 2000. We don't know what the next 10 years will bring. We still have another eight and a half years to go, but for the one and a half years we have recently visited, the return realization is fitting the chart nicely.

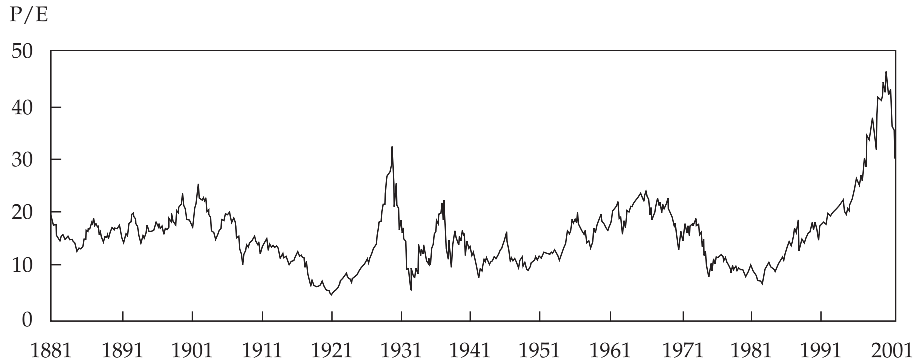
The relationship between starting P/E and subsequent return is potentially exaggerated because much of the strong relationship comes from P/E reversion. What if P/Es did not change?

Figure 2 presents some input into the relationship if P/Es were constant. In the figure, trailing 20-year real S&P earnings growth is plotted for the past 110 years. For this period, annualized real earnings growth averaged 1.5–2.0 percent fairly consistently. Those people who actually still assume 10 percent nominal returns on stocks should recognize that such a return would require 5–6 percent real earnings growth over the next 10–20 years. Such growth has happened only a few times in history, and it has happened only after very depressed market conditions, which we are not really experiencing now, certainly based on the last 10 years. With a 2 percent real earnings growth forecasted, a long-term buy-and-hold investor in the S&P 500 can expect to earn 6–7 percent nominal returns.

What Can Save the Stock Market?

I envision a bad 1920s-type serial in which the villain has tied the stock market to the railroad tracks and a

Figure 1. Historical P/E of the S&P 500, 1881–2001

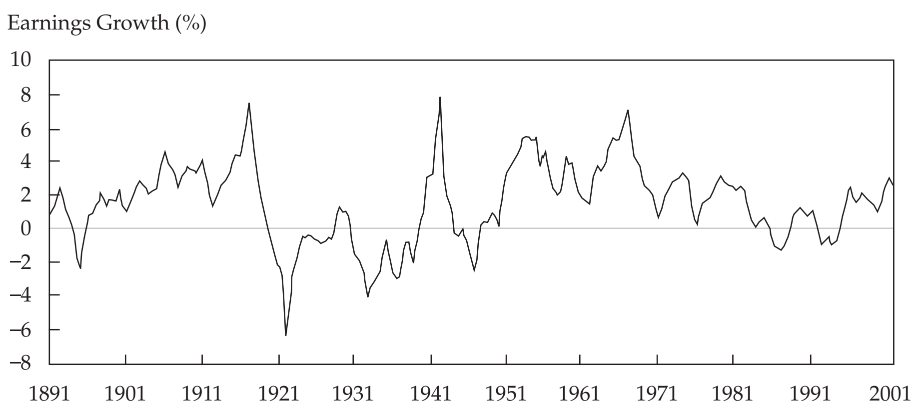


Note: P/E was calculated as the current price divided by the average of earnings for the past 10 years adjusted for inflation.

Table 1. Real Stock Market Return in the Next 10 Years for Historical P/E Ranges of the S&P 500, 1927–2001 Data

P/E Range (low to high)	Median Return (annualized)	Worst Return (total)
5.6 to 10.0	11.0%	46.1%
10.0 to 11.7	10.6	37.3
11.7 to 14.1	10.0	4.1
14.1 to 16.7	9.0	–19.9
16.7 to 19.4	5.4	–23.1
19.4 to 32.6	–0.4	–35.5
32.6 to 45.0	Here Be Dragons!	

Figure 2. S&P 500 Trailing 20-Year Real Earnings Growth, 1891–2001



Note: Earnings growth is annualized.

voice-over is pleading, "What can save stocks?" This question is going to be the organizing principle for my presentation today. I am going to concentrate on three things that could save stocks, although other answers may be possible. One is sustained high real earnings growth—"high" meaning better than the historical average. The second, a Wall Street favorite, is the so-called Fed model, in which the U.S. Federal Reserve lowers interest rates and supports high P/Es. The third is a simple hero—investor acceptance of lower future rates of return in the long term.

HIGH EARNINGS GROWTH. First, something we all probably know: Only if the future brings extra-special, super-high earnings growth are very high starting P/Es justified. For each level of P/E at the start of a 10-year period except very low P/Es (when returns are always on average strong), decades with stronger earnings growth also experienced stronger average stock returns, and even when P/Es were high, if earnings growth came in very high, returns were on average strong. This analysis, however, gives us an *ex post*—not a predictive—measure. If we see extraordinarily high growth in real earnings after 2001, we will probably see high real equity returns. However, the question is: What reason do we now have to be optimistic that such abnormally high earnings growth will occur?

One reason is that higher productivity and technological advancement could create high earnings growth. I think this development is unlikely. Historically, most productivity benefits accrue to workers and consumers, not necessarily to earnings:

> Optimists frequently cite higher growth of real output and enhanced productivity, enabled by the technological and communications revolution, as the source of this higher growth. Yet the long-run relationship between the growth of real output and *per share* earnings growth is quite weak on both theoretical and empirical grounds. (Siegel 1999, pp. 14–15)

So, the first hurdle to believing in high earnings growth is to believe the productivity numbers, and the second is to believe earnings will benefit.

Now, let's look at the empirical data. In **Table 2**, I show the historical relationship between P/E at the beginning of a period and subsequent average 10-year real earnings growth for 1927–2001. The numbers in the 16 quadrants, or 16 buckets, are actual realized real earnings growth over rolling 10-year periods. Each number corresponds to a range of starting P/Es and a range of starting earnings retention rates. Historically, when both the starting P/E and the retention rate are high, the real earnings growth rate is low. On May 30, 2001, the P/E of the S&P 500 was 27.3 and the retention rate was 65.3 percent, which today puts us in the bottom right bucket, so the dragons are off to the right. This position is not promising for saving stocks.

We can interpret Table 2 further. The second way stocks could experience future high earnings growth is through market efficiency. The idea is that in an efficient market, high current P/Es will lead to higher earnings growth because the market must be right. I like this approach. I wish it were the case, but I don't think the data support it well. Table 2 shows no relationship between starting P/E and future earnings growth. In fact, P/E does a lousy job of predicting earnings growth. I will go further. It does no job. In fact, the data show that higher P/Es have not led to higher real earnings growth going forward and lower P/Es have not led to lower growth. The joint hypothesis of constant expected returns and market efficiency should lead to P/Es predicting growth, but the hypothesis doesn't hold, at least in the data.

Finally, Table 2 sheds light on the third reason we might now expect high earnings growth: the idea that high cash retention (low payout ratios) leads to strong growth. Table 2 indicates, however, that the retention rate at the beginning of a period has been *inversely* related to the subsequent 10-year growth in earnings. The impact of the retention rate is incredibly, astronomically backward. Rob Arnott and I have struggled with this phenomenon. We haven't found this impact to be intuitive—it is not a forecasted result—but we do have a few *ex post* theories as to why higher retention rates might lead to lower real growth rates. I'll share three of them quickly.

The first reason relates to company managers. The general idea is that companies retain a lot of cash

Table 2. Average 10-Year Real Earnings Growth, 1927–2001 Data

	Retention Rate (%)				
Starting P/E	Negative to 37.7	37.7 to 44.4	44.4 to 50.3	50.3 to 63.9	63.9 →
5.9 to 10.4	4.1%	2.5%	2.2%	–0.3%	
10.4 to 13.8	4.3	2.5	2.4	0.6	
13.8 to 17.2	3.3	2.5	1.7	–0.4	
17.2 to 26.3	4.3	2.7	0.8	–0.6	
26.3 →					The Dragons Are Here!

to finance projects for behavioral reasons such as empire building. If the cash is for projects, managers are not doing a good job with the cash; they tend to pursue and overinvest in marginal projects, which is reflected in the future lowered growth rates of the company. If this is the explanation, the telecom boom in the late 1990s is going to be the poster child for empire building for all eternity.

Another theory, less plausible in my opinion, is that managers have information that the market doesn't have. It is generally accepted that companies are loath to cut dividends. So, the theory goes that when a company's managers pay high dividends, the market perceives that those managers must have such positive information about the company's prospects that they know they will not have to cut dividends in the future. When managers pay high dividends, they are optimistic because they have information unknown in the market. When managers do not pay high dividends, they must be nervous. So, retention of earnings may reflect a desire by managers to smooth dividends.

The third explanation is that Rob and I are doing something wrong. We have each double-checked our approach and the data repeatedly, but when you get a wacky result, for intellectual honesty, you still have to admit the possibility. That is why I mentioned the dragons, because we are off the charts and into uncharted territory.

If history repeats and higher P/Es and higher retention rates lead to lower real earnings growth and if Rob and I are not making an error, the future does not bode well for real earnings growth.

LOW INTEREST RATES. The second possible way stocks can be saved is low interest rates. **Figure 3** compares the P/E (or the "absolute" value of the S&P 500) with the earnings yield on the S&P 500, E/P, minus the 30-year U.S. T-bond yield, Y (or the "relative" value of the S&P 500); Panel A graphs these indicators for the past 20 years. As you can see, P/E has certainly fallen from its peak in 1999 but is still at the high end of the 20-year range. The equity yield minus the bond yield is one version of the Fed model. In that model, a high value is an indication of good news for the equity market, but for P/E, a high value indicates bad news for the market. Using the Fed model, the situation does not look that bad in 2001; the market is above average on earnings yield minus bond yield.

The same information, but stretching back to 1927, is presented in Panel B of Figure 3. The line for earnings yield minus bond yield is pretty lackluster over the period. When stocks were far cheaper in relation to bonds, stocks used to be bought for their dividend yield; this chart uses earnings yield, but the difference is not really important. As Panel B shows, if Wall Street had a little bit longer perspective, such as looking back to 1927 rather than just 20 years, even the Fed model, or the relative value of the equity market, does not look great.

Forgetting the data, note that the Fed model has little theoretical standing. Nominal earnings growth does correlate nicely with expected inflation over time. A lot of confounding biases, such as depreciation methods, accounting choices, and different inflationary environments, affect the P/E calculation (see Siegel 1998). But by and large, the net of those biases is not clear. What does appear fairly clear, however, is that the market does not seem to understand that if you write down the expected return of a stock (dividend yield plus earnings growth), then if inflation and interest rates fall and earnings growth drops along with them, the P/E does not have to change. I think you understand the concept, but it is an idea I have to explain to most people, and I encourage you to do the same. People believe P/Es have to move with interest rates, and they are probably wrong, or at least overstating the relationship.

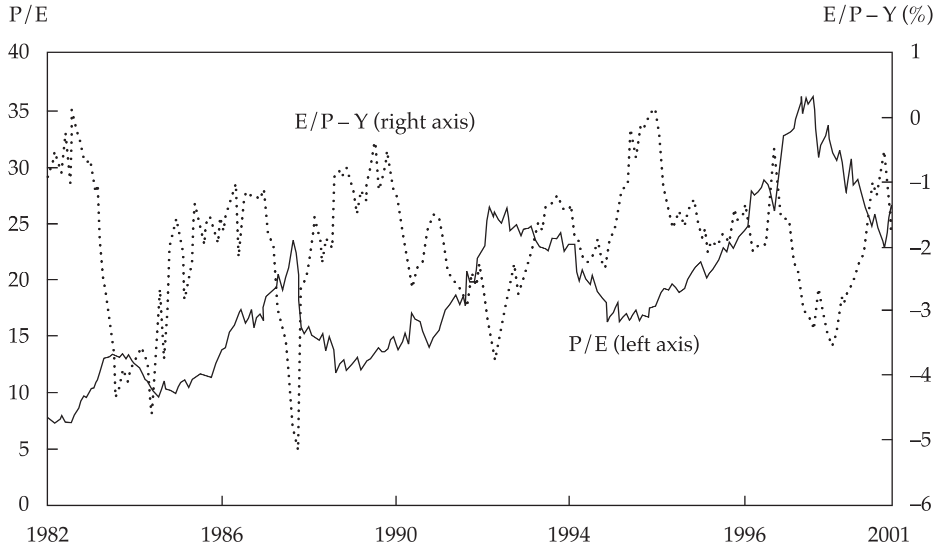
Figure 4 shows a plot of the S&P 500's realized 20-year volatility divided by the bond market's 20-year realized volatility against the relative yield of the stock market for 1950 to 2001.[1] I chose 20 years because I think of 20 years as a generation, so the ratio plotted from the x-axis reflects what a generation thinks in terms of how risky stocks are versus bonds. This ratio is a very robust indicator for each five-year period, up to 30 years. The y-axis is the earnings yield on the S&P 500 minus the 10-year bond yield. Whenever you look at long-term autocorrelated relationships like this, you have to carry out many, many robustness tests. This ratio survived every test we came up with.

Note that the y-axis is not stock yields; it is stock yields *minus* nominal bond yields. The market clearly does trade on interest rates in the short term. Not many models have a high R^2 at forecasting short-term (less than a one-year horizon) market performance. One indicator that is less pathetic than most in this regard is deviation from the fitted [linear (normal)] line in Figure 4. However, for longer horizons, such as forecasting the next 10-year real stock return, neither the bond yield nor the volatility measures matter. P/E alone forecasts the real stock return. So, an investor with a short horizon cares a lot about this line, but an investor with a long horizon doesn't.

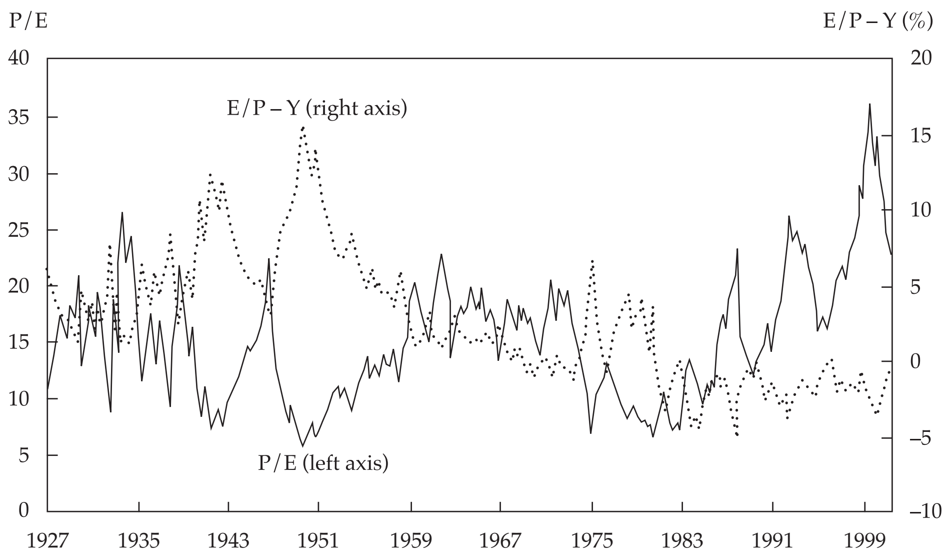
[1] Figure 4 is similar to Figures 7 and 8 in Asness (2000b). In that article, Figure 7 goes back to 1871 and forward to mid-1998 and Figure 8 goes back to 1881 and forward to mid-1998.

Figure 3. S&P 500 "Absolute" and "Relative" Value

A. 1982–2001



B. 1927–2001



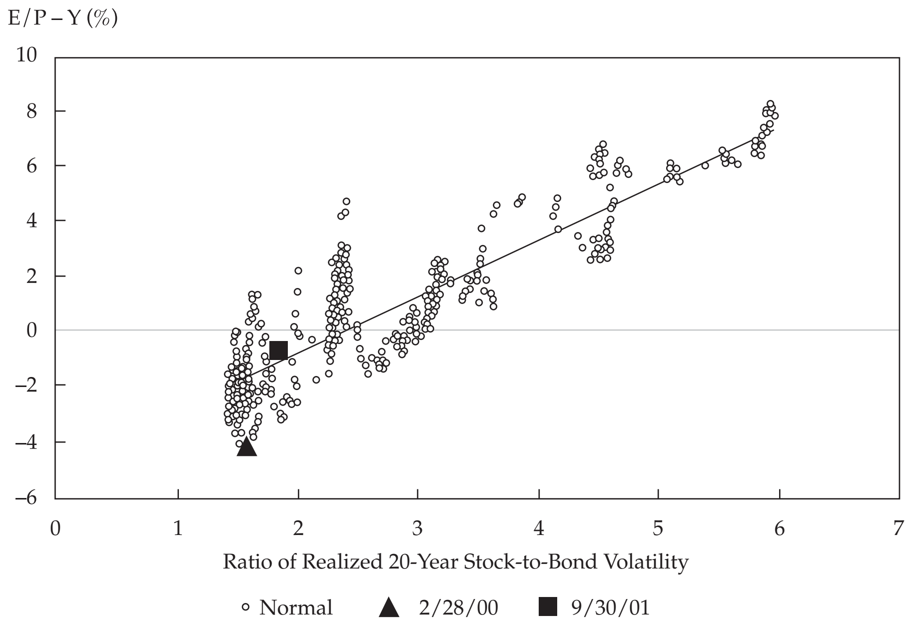
Note: S&P 500 P/E and E/P; 10-year T-bond yield.

I have marked on Figure 4 where we were on February 28, 2000, and on September 30, 2001. On February 28, 2000, short-term traders could not be saved by anything; the solid triangle is well under the line. Stocks were yielding much less than they had historically—even given unusually low volatility and unusually low interest rates relative to the historical average.

The September 2001 mark in Figure 4 indicates that stock performance doesn't look too bad over the very short term. Short-term investors tend to trade on this relationship—that is, trade on the idea that eventually the market moves back to the line for behavioral reasons. Note that this relationship is behavioral because it is based on errors—which does not change what the equity risk premium is in the long term. Over the short term, it is the deviation of E/P from the line that counts; over the long term, it is only the actual E/P that counts.

ACCEPTANCE OF LOW RETURNS. Now for the third possible hero that might save the stock

Figure 4. Stock versus Bond Valuation, 1950–2001



Note: S&P 500 E/P; 10-year T-bond yield.

market: Are investors willing to accept low stock returns? Have they understood the idea that future returns will be low, as so many of us have discussed. A ton of "strategists" will give explanations of why high P/Es are supportable, but then they will follow the explanations with the expectation of 10–12 percent stock returns anyway. That reasoning is questionable to say the least. The first part is believable; no one can say that a 1–2 percentage point return over bonds is bad. But you cannot have your cake and eat it too. Or as I like to say when it comes to Wall Street investors, they cannot have their cake and eat yours too.

What if investors haven't yet realized the conundrum of expectations versus reality? Surveys exist—Campbell Harvey is going to present his survey data [see the "Implications for Asset Allocation, Portfolio Management, and Future Research" session]—that indicate respondents are expecting very high equity returns. Survey data are not always the most reliable, but the data report that the high return expectations are out there. I talk to a lot of pension plans, and not many of them are using assumptions as low as 6–7 percent nominal returns or a 1 percent real equity return over bonds. And investors who plan to retire at 38 because they expect to get a 5 percent equity risk premium and 7 percent real stock returns forever are going to wake up at 62 out of money.

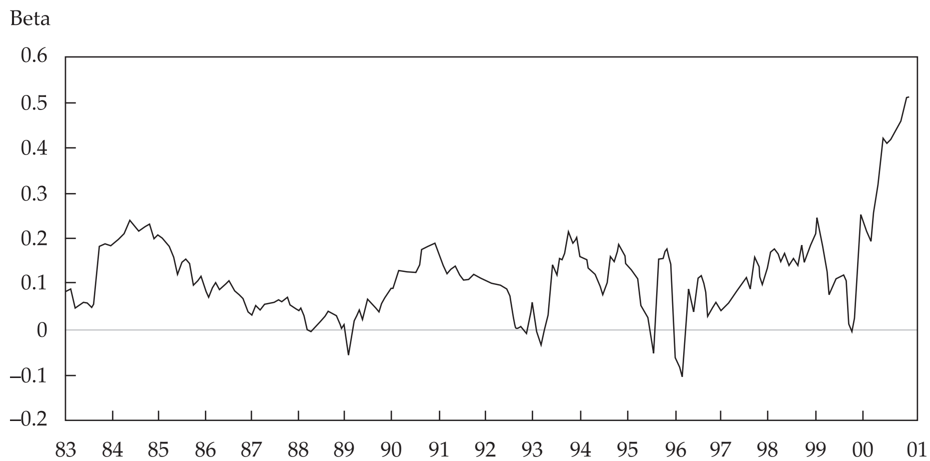
Are investors rationally accepting the low equity risk premium, or are a lot of people still trying to buy lottery tickets?[2] Many have shown that Wall Street's growth expectations are ridiculously optimistic, but investors seem to still believe them. So, Rob and I examined a strategy based on these expectations. We formed a portfolio for a 20-year period that was long high-growth stocks and short low-growth stocks (based on Wall Street's estimates). **Figure 5** shows the rolling 24-month beta of that long–short portfolio from December 1983 to September 2001. For a long time, the beta was mildly positive, but for the past few years, it has been massively positive. It is a dollar long, dollar short 0.5 beta. Figure 5 says that every rally for the past several years has occurred because the high-expected-growth stocks were crushing the low-expected growth stocks. And every market sell-off has been a result of the opposite occurring. Does this pattern indicate rational acceptance of the low equity risk premium or the buying of lottery tickets?

Conclusion

Broad stock market prices are still well above those of most recorded history (and of all history excluding 1999–2000 and just before the crash of 1929). Unless a miracle happens, we must prepare for very low returns as compared with history. In the end, the market offers two choices: low long-term expected

[2] See Statman (2002).

Figure 5. Rolling 24-Month Beta of Long–Short Portfolio, December 1983–September 2001



Note: Except for 2001, dates are as of December.

returns in perpetuity or very bad short-term returns with higher, more normal expected returns in the long run. My personal opinion: Do the events of 1999–2001 strike anyone as a group of rational investors embracing and accepting a permanently low risk premium? If so, I missed it on CNBC.

Theoretical Foundations II

Clifford S. Asness

AQR Capital Management, LLC
New York City

SUMMARY

by Peter Williamson

Amos Tuck School of Business Administration
Dartmouth College, Hanover, New Hampshire

Clifford Asness made the second presentation of the day, beginning with a graph (**Figure 1**) showing the record of the S&P 500 Index's P/E (current price divided by the average of the preceding 10 years' real earnings) for 1881 to 2001. The highest P/E, about 45, was reached in 2000. **Table 1** reports for each of six ranges of P/E the median real stock market return in the next 10 years and the return for the worst decade. In general, high P/Es led to low subsequent returns and to the worst of the worst decades and low P/Es led to high returns and to the best of the worst decades.

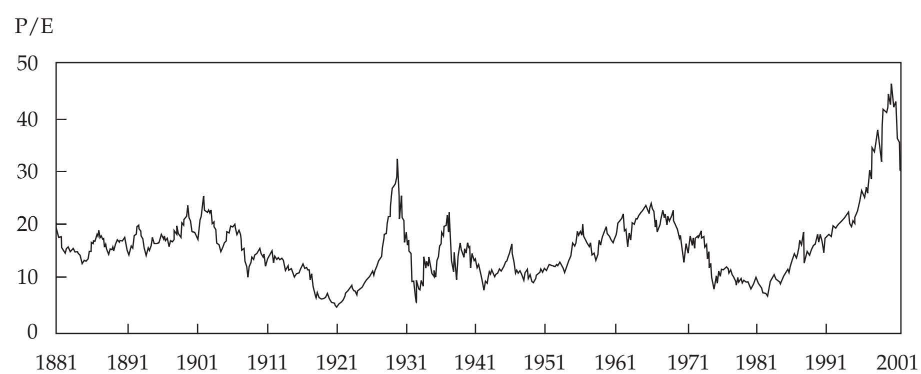
Asness observed that much of what Table 1 shows in terms of consequences of P/E levels comes from P/E reversion. Some would ask: What happens if the ratios do not revert? **Figure 2**, showing S&P 500 trailing 20-year real earnings growth (annualized) helps to answer the question.

Asness next examined three possible ways in which the market might be saved from decline. One is high and sustained real earnings growth. A second (the Wall Street solution) is low interest rates. This is the so-called Fed model. The third way is based on investor acceptance of lower future rates of return. This answer would mean no imminent crash but a less attractive long-term return.

Would high earnings growth work? **Table 2** shows the historical relationship between P/E at the beginning of a period and subsequent average 10-year real earnings growth for 1927–2001. The numbers in the 16 quadrants, or 16 buckets, are actual realized real earnings growth over rolling 10-year periods. Each number corresponds to a range of starting P/Es and a range of starting earnings retention rates. Historically, when both the starting P/E and the retention rate are high, the real earnings growth rate is low.

Why might we expect high earnings growth? Some might say because of increasing productivity and technological advancement. But the relationship between growth of real output and *per share earnings*

Figure 1. Historical P/E of the S&P 500, 1881–2001

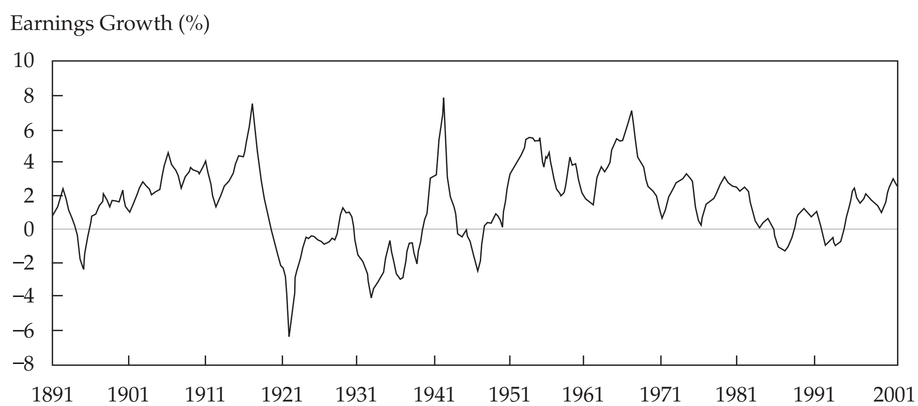


Note: P/E was calculated as the current price divided by the average of earnings for the past 10 years adjusted for inflation.

Table 1. Real Stock Market Return in the Next 10 Years for Historical P/E Ranges of the S&P 500, 1927–2001 Data

P/E Range (low to high)	Median Return (annualized)	Worst Return (total)
5.6 to 10.0	11.0%	46.1%
10.0 to 11.7	10.6	37.3
11.7 to 14.1	10.0	4.1
14.1 to 16.7	9.0	–19.9
16.7 to 19.4	5.4	–23.1
19.4 to 32.6	–0.4	–35.5
32.6 to 45.0	Here Be Dragons!	

Figure 2. S&P 500 Trailing 20-Year Real Earnings Growth, 1891–2001



Note: Earnings growth is annualized.

Table 2. Average 10-Year Real Earnings Growth, 1927–2001 Data

	Retention Rate (%)				
Starting P/E	Negative to 37.7	37.7 to 44.4	44.4 to 50.3	50.3 to 63.9	63.9 →
5.9 to 10.4	4.1%	2.5%	2.2%	–0.3%	
10.4 to 13.8	4.3	2.5	2.4	0.6	
13.8 to 17.2	3.3	2.5	1.7	–0.4	
17.2 to 26.3	4.3	2.7	0.8	–0.6	
26.3 →					The Dragons Are Here!

has been weak. Some would argue that in an efficient market, the current P/E simply *must* be justified by high earnings expectations. Asness thinks the data do not provide much support for this proposition.

A third reason might be that high cash retention leads to above-normal growth. But referring to Table 2, he pointed out that the current retention rate has been significant in relation to real earnings growth and the retention at the beginning of a 10-year period is *inversely* related to the subsequent 10-year growth in earnings! Why should this be? One answer is empire building. Retention of earnings is simply not productive. A second is a desire on the part of managers to smooth dividends. In any case, the current retention rate is about 65 percent, and Table 2 is not encouraging for the future of the stock market.

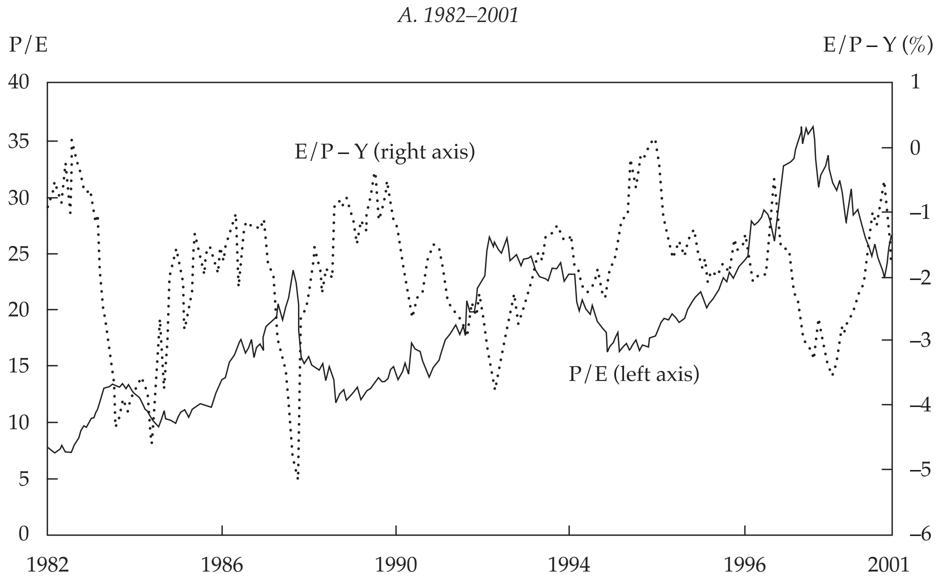
A second way in which the market might be saved is through low interest rates. Can low interest rates save stocks? Panel A of **Figure 3** is encouraging: Interest rates below about 3 percent are very helpful.

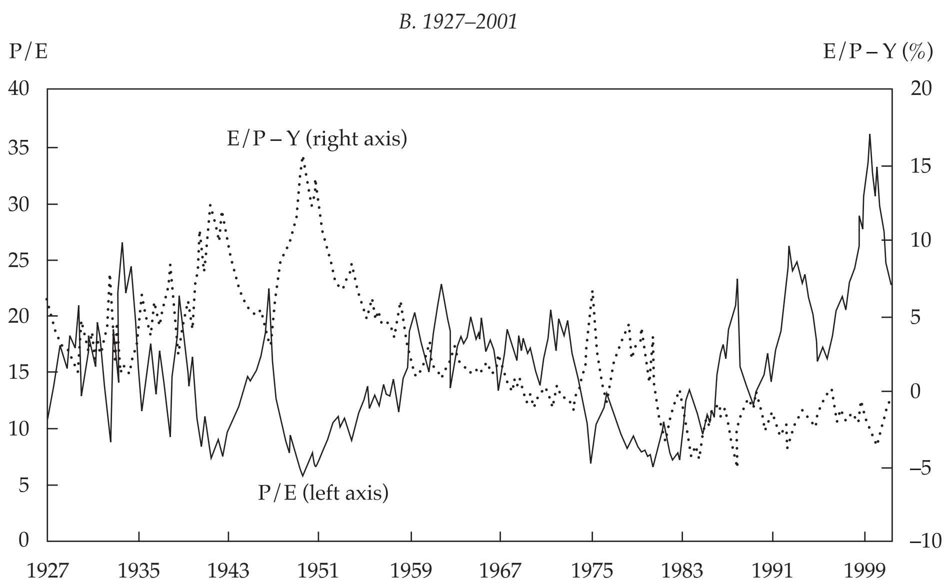
But Panel B shows that over a longer historical period, the news is not so good. The indicator seems to be the earnings yield, E/P, less the bond yield, Y. There is evidence that nominal earnings growth is correlated with inflation. The P/E, however, is mostly a real entity, and comparing it with nominal bond yields cannot be expected to have much long-term forecasting power.

Finally, the willingness of investors to accept low stock returns might save the market. Are investors willing to accept low stock returns? Declining volatility may be justifying high P/Es and low returns. **Figure 4** provides support for this idea, although the vertically plotted E/P minus Y mixes real and nominal data.

Figure 4 seems to work for the short term. The point on the graph for September 30, 2001, represents a high P/E coupled with a low ratio of realized 20-year stock-to-bond volatility. For the longer term, the E/P is a better guide to real stock returns.

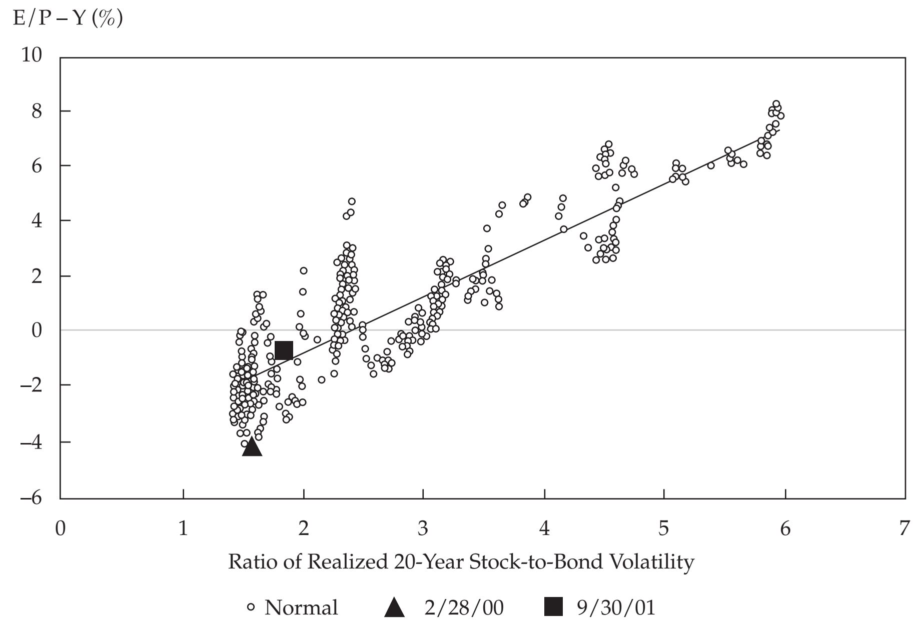
Figure 3. S&P 500 "Absolute" and "Relative" Value





Note: S&P 500 P/E and E/P; 10-year T-bond yield.

Figure 4. Stock versus Bond Valuation, 1950–2001

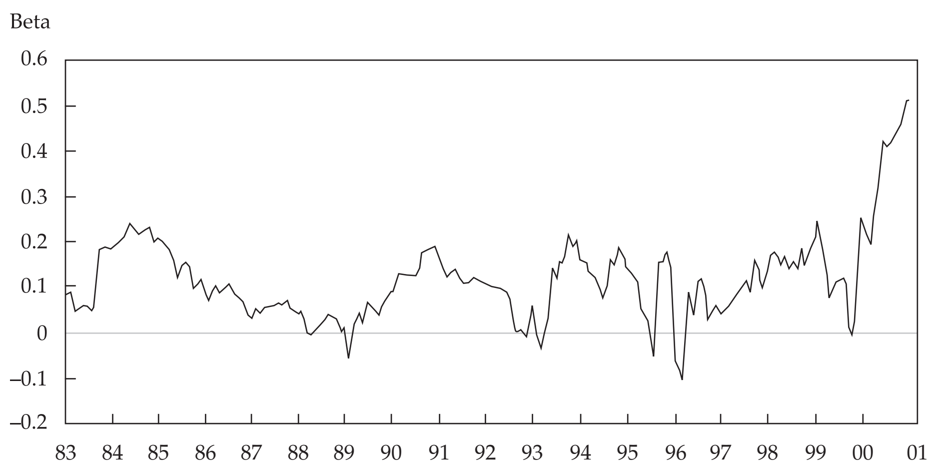


Note: S&P 500 E/P; 10-year T-bond yield.

Acceptance of a 6–7 percent nominal stock return appears not unreasonable. But Asness went on to present evidence that investors do not actually think they are facing such low returns. In this case, when they realize the true prospects, then short- to medium-term returns will be low. To raise the expected return on the S&P 500 by 2 percentage points, the price must fall about 50 percent.

Figure 5 shows the results of forming long–short portfolios (based on Wall Street growth forecasts) in which the portfolios were long the high growers and short the low growers. The rolling 24-month beta of the portfolios has been consistently positive and, in recent years, has been massively positive. Every rally has seen the high-expected-growth stocks crushing the low-expected-growth stocks. Asness thought this

Figure 5. Rolling 24-Month Beta of Long–Short Portfolio, December 1983–September 2001



Note: Except for 2001, dates are as of December.

was not a picture of investors willing to accept lower equity premiums.

In conclusion, he said:

- Broad stock market prices are still well above the levels of most recorded history (and of all history excluding 1999–2000 and just before the crash of 1929). Unless a miracle happens, we must prepare for very low returns as compared with history.
- The choice is between low long-term returns forever and very low (crash type) returns followed by more historically normal returns.

Finally, he offered the following reflection: Do the events of 1999–2001 strike anyone as a picture of rational investors accepting a permanently low risk premium? Answer: No.

Theoretical Foundations: Discussion

Stephen Ross (*Moderator*)
Robert Arnott
Clifford Asness
Ravi Bansal
John Campbell
Bradford Cornell
William Goetzmann
Roger Ibbotson
Martin Leibowitz
Rajnish Mehra
Thomas Philips
Robert Shiller
Richard Thaler

STEPHEN ROSS (Moderator)

I have a few brief comments. They will be brief for two reasons. First, I am confused. Second, even in my confusion, I am in the uncommon position of not having a lot to say. Let me turn first to Cliff Asness's presentation.

What is puzzling to me about Cliff's presentation is that the discussions about P/Es and other broad descriptors of the market seem to me to be discussions that we could have held 100 years ago. The vocabulary would have been a little different, but in fact, not only could we have held the discussion, I suspect these discussions *were* held 100 years ago. So, I don't think we are saying many things differently now than we said back then.

What is troubling to me is that we are supposed to be making progress in the theory. To the contrary, the theory seems to me to be in a wasteland, not just regarding the risk premium but, more generally, in much of finance. We are in a period of time, a phase, in which data and empirical results are just outrunning our ability to explain them from a theoretical perspective. This position is a very tough one for a theorist who used to dine high on the hog when we had derivatives pricing, where theory worked wonderfully. Now, we are interested in theory to explain the problems, which is not working quite so wonderfully.

It seems to me that the issues involving P/Es are issues involving whether or not these processes are mean reverting. Obviously, something like the P/E has to revert to the mean; it is only a yield. Jonathan Ingersoll made a wonderful comment about interest rates and whether interest rates revert or not. He noted that interest rates existed 4,000 years ago in Egypt and if interest rates didn't mean-revert, they would be 11,000 percent today. So, they have to revert.

We know P/Es revert, but they seem to revert very slowly, and we are able to measure the reversion only with great difficulty. Our efforts to measure, for example, stock returns—not actual returns but expected returns—have basically been futile.

I also have some comments about Richard Thaler's presentation. I am often characterized as a defender of the neoclassical faith. I know I am because often I am asked to debate Richard. Sometimes, however, I am characterized as a shill of the neoclassical school. So, it is not clear to me which position I am supposed to represent in the minds of market pundits. But I will say that I feel a bit like one of those physicians with a gravely ill patient to whom I would like to suggest the possible benefits of herbs and acupuncture—alternative medicine. I call for "alternative finance," not behavioral finance as the alternative approach, but an alternative that may offer a little bit of hope.

What I actually think is that our prey, called the equity risk premium, is extremely elusive. We cannot observe the expected return on stocks even with stationarity in time-series data because volatility and the short periods of time we are able to analyze give us little hope of actually pinning down a result. The best hope, from the empirical perspective, seems to lie in cross-sectional analysis, which is not what we are talking about here; we are talking mostly about *time series*, for which we do not have many observations. Cross-sectional analysis says that the excess returns should be the risk premium times the beta. If we could find some way to spread excess returns, maybe through P/Es of individual stocks, then we'd have a better chance of measuring expected return at each point in time—no matter what theory we decide to pin our hopes on.

The theory itself is a myth, and in this case, Richard and I are in complete agreement. Any hope of tickling, or torturing, some reasonable measure of the risk premium out of consumption data is forlorn. It resides in the hope that somehow people are rational.

I love old studies. For example, in one study on consumption data that was done mostly in Holland, the researchers observed shoppers in supermarkets to see what happened when the price of soap was higher than the price of bread. These shoppers did not adjust their marginal rates of substitution to the prices of consumer goods at a single point in time, let alone in the presence of uncertainty and over time. But consumption theory has always said that people would adjust their marginal rates of substitution for prices that evolve over time in a stochastic world.

I am not at all surprised, nor am I troubled, by the fact that we do not find any meaningful correlations between something that we may or may not be able to measure, such as expected return and consumption, and the interplay between them. So, I applaud Richard's view that we ought to consider other reasons to explain why people do what they do.

The real puzzle may be: Why do investors behave the way they do based on what the premiums actually are? And here too, I have to say that even though neoclassical theory is not up to the task of explaining this behavior, and it is not doing a good job, I am not sure that behavioral theory has much more to say to us.

Behavioral anecdotes and observations are intriguing. Behavioral survey work is empirically fortified. But behavioral theory does not seem to have a lot of content yet. In interpreting the study that Richard mentioned about the incompatibility of two gambles, one has to be very careful. Those gambles are incompatible if they are assumed to hold over the entire range of the preference structure. But there is no reason to believe that the gamble holds over the entire range of the preference structure. We do not believe that if the guy wins $20 million he won't take the 110 to 100 gamble. The uniformity requirements in that assumption bend the question. A lot of curious things are going on in those kinds of analyses of behavioral assumptions. And even the richer models, such as those of DeLong and Shleifer (1990), have their own problems.

In summary, I am a theorist and I am confused. I would like theory to make progress, and I would like for us to be able to address some of these issues successfully. I do not really care whether we do so from a neoclassical or another perspective, but I find myself facing an enormous, complicated array of phenomena that come under the heading of "the equity risk premium puzzle" and I'm completely unable to explain any of it.

RAJNISH MEHRA: One thing that Richard Thaler missed was that most of these models do not incorporate labor income. Constantinides, Donaldson, and I (1998) have been doing work in this area for the last couple of years. We have been analyzing the implications of the changes in the characteristics of labor income over the life cycle for asset pricing. The idea is simple: The attractiveness of equity as an asset depends on the correlation between consumption and equity income, and as the correlation of equity income with consumption *changes* over the life cycle of an individual, so does the attractiveness of equity as an asset. Consumption can be decomposed into the sum of wages and equity income. A young person looking forward in his or her life has uncertain future wage *and* equity income; furthermore, the correlation of equity income with consumption will not be particularly high as long as stock income and wage income are not highly correlated. This is empirically the case. Equity will thus be a hedge against fluctuations in wages and a "desirable" asset to hold as far as the young are concerned.

Equity has a very different characteristic for the middle-aged. Their wage uncertainty has largely been resolved. Their future retirement wage income is either zero or fixed, and the fluctuations in their consumption occur from fluctuations in equity income. At this stage of the life cycle, equity income is highly correlated with consumption. Consumption is high when equity income is high, and equity is no longer a hedge against fluctuations in consumption; hence, for this group, equity requires a higher rate of return. The way Constantinides, Donaldson, and I approach this issue is as follows: We model an economy as consisting of three overlapping generations—the young, the middle-aged, and the old—where each cohort, by the members' consumption and investment decisions, affect the demand for, and thus the prices of, assets in the economy. We argue that the young, who should be holding equity, are effectively shut out of this market because of borrowing constraints. In the presence of borrowing constraints, equity is thus exclusively priced by the middle-aged investors, and we observe a high equity premium. We show that if there were no constraints on young people participating in the equity markets, the equity premium would be small.

So, I feel that life-cycle issues are crucial to any discussion of the equity premium.

JOHN CAMPBELL: I want to follow up on the point Rajnish Mehra made because one part of Richard Thaler's talk was normative analysis—the claim that if the equity risk premium is as much as 4–5 percent, long-term investors should obviously hold their money in stocks or even leverage a position to hold their money in stocks. I think that, as a normative statement, that prescription is simply wrong.

I am going to take as a benchmark a model with constant relative risk aversion at some reasonable, traditional low number. The simple formula for the share you should put into stocks if you are living off your financial wealth alone and if returns are distributed identically every period is as follows: the risk premium divided by risk aversion times variance. Suppose the risk premium is 4 percent and the standard deviation of stocks is 20 percent; square that and you get 4 percent. Now, you have 4 percent divided by risk aversion times 4 percent. So, if your risk aversion is anything above 1—say, 3 or 4—you should be putting a third of your money in stocks or a quarter of your money in stocks. It is just not true that with low risk aversion and a risk premium of 4–5 percent you should put all your money in stocks.

So, what's happened to the puzzle? Why don't I get an equity risk premium puzzle when I look at it from this point of view? Well, the key assumption I made is that *you are living off your financial wealth entirely.* It follows then that your consumption is going to be volatile because it will be driven by the returns on your financial wealth. The only way to get an equity risk premium puzzle is that when you look at the smoothness of consumption, you see that it is much smoother than the returns on the wealth portfolio. Why is that?

Rajnish's point is that other components of wealth, such as human capital, are smoother, which is keeping down the total risk of one's position. If you have these other, much smoother human assets, then of course, stocks look very attractive. But I think it's important not to assert that a risk premium of 4 percent should induce aggressive equity investment.

I am reminded of Paul Samuelson's crusade over many years to get people to use utility theory seriously, as a normative concept. He was always trying to combat the view that you should just maximize the expected growth rate of wealth. He got so frustrated by his inability to convince people of this that he finally wrote an article called, "Why We Should Not Make Mean Log of Wealth Big Though Years to Act Are Long" (1979). It is a wonderful article, and the last paragraph says, "No need to say more, I've made my point and but for the last word, I've done so in words of but one syllable." And every word in the article is a one-syllable word except for the last word. It is almost impossible to read, of course, but the point is important: We may not want to use standard utility theory as a positive theory, but we should try to use it as a normative theory, in my view.

ROSS: If you are going to use it as a normative theory, though, you do not have to place your attention entirely on the constant relative-risk-aversion utility function. The broader class of linear risk-tolerance models has exactly the same function (with the addition of deterministic parts to the income stream), except they work in the opposite direction. So, if someone has a linear risk tolerance with a high threshold for that risk tolerance, then the equity risk premium puzzle reappears because the desire to invest is huge even when the risk premium is relatively low.

RICHARD THALER: Let me respond briefly. You have all these models that are based on consumption, and it is true (and I appreciate John Campbell's clarification) that to really understand this puzzle, you need to emphasize consumption smoothing. Otherwise, you get precisely the result that John suggested.

But the puzzle I was informally identifying before refers to other investors that I think have been neglected in much of this theoretical research. Those simulations that Marty Leibowitz was doing were mostly for defined-benefit pension funds, and I did some similar simulations for a foundation that I've been associated with over the years. Foundations have 5 percent mandatory spending rules. Now, if you crunch the numbers and you are investing in bonds, basically you are certain to be out of business in the near future unless you can find some bonds providing a 5 percent real rate of return. With TIPS we were getting close for a while.[1] But if the real interest rate is 2 percent and you have to spend 5 percent, you are soon going to be out of business. One question I have for the theorists, of which I am not one, is: What's the normative model we want to apply for those investors and what does it tell us about the kind of risk premium we should expect?

BRADFORD CORNELL: I have one question: Most of you are involved in one way or another with investment firms, and it is almost a mystery to me that you read academic papers where you see things like "consumption process," "labor income," "risk aversion," and so on, and then you attend an actual investment meeting—where none of these concepts are even remotely talked about. So, how do you bridge the gap between the supposed driving factors of the models and equilibrium returns and the way people who are actually making decisions make them? Is there a way to tie all of it together?

ROSS: There does seem to be a disconnect between the two areas and the two literatures. It is, actually, a fundamental theoretical disconnect. In these markets, with their many institutional players, the institutions are typically run by managers under some type of agency structure. So, there must be some sort of agency model for the people who run the pension funds and other institutions. They are the ones who

[1] Treasury Inflation-Protected Securities; these securities are now called Treasury Inflation-Indexed Securities.

make investment decisions. In the theoretical structures we build that include consumption, we seem to have the view, or maybe just the wishful thinking, that whatever the underlying forces in the economy are, these institutions will simply be transparent intermediaries of those forces, so the agents who are representing these institutions will simply be players in people's desire to allocate consumption across time or will be dealing with the life-cycle problems of people. Some take a Modigliani view that the *people* will adjust their actions around whatever the agents do. The net result is that the actions of the agents and the people coincide, which seems to me overly hopeful. I don't believe it is the case.

CLIFFORD ASNESS: Is it more complicated than saying the description Richard Thaler presented works better for what actually happens in a boardroom than any of the theory? Behavior like myopic loss aversion is true. Many of us have behaved that way. The fact that people make choices in the ways that they do does not have to be proven by a survey. As a manager who has gotten way too much money after a good year and too many redemptions after a bad year, I can tell you people focus on the short term.

I have one comment about Steve Ross's initial response. I don't think anyone would argue about the fact that P/Es are mean reverting. But that is not the exciting part of the puzzle. The exciting part, which is incredibly challenging, is that if we all accept that P/Es are mean reverting to an unconditional mean, what we are disagreeing about is what that unconditional mean either should be, in theory, or is. Mean reversion is a pull toward something, and the open issue is not mean reversion but whether the "right" (meaning unconditional mean) P/E is 15. If it is and we are in the high 20s, then mean reversion is not going to work as a good model for the next year. But the pull was downward for a long time, so I do not think my comments were trying to be insightful about P/Es being mean reverting. They have to be, or else they are unbounded in some direction.

MARTIN LEIBOWITZ: This is just strictly an observational comment, not a theoretical one, and it has to do with the comment about myopic loss aversion or myopic return attraction, which is the other side of the coin. As Cliff Asness said, there's clearly some pain in the short term and also some joy in the short term, depending on your outcomes. But I think what actually happens is that people incorporate a kind of Bayesian revision, that the prospects for the future are based on what have been the most immediate short-term returns.[2] We see it in terms of the flow of funds into, for example, TIPS—a wonderful instrument with a great yield, a +4 percent real rate. We couldn't get anyone to invest in them until, suddenly, we had a 12.76 percent return year in the equity market, at which point, of course, the real return on equities was a lot lower than it had been and money started flowing into TIPS big time. Short-term return is a very powerful force.

THALER: Aren't you too Bayesian, then, to be sarcastic?

LEIBOWITZ: Yes, Bayes would recoil because in the fixed-income area, this short-term focus is clearly, you know, a kind of nuttiness, although there's something to it. It does show that real rates can decline. I think some people were thinking: Why were we stuck with real rates in the area of +4 percent? So, myopic loss aversion is not totally irrational, even in the fixed-income area. In the equity area, where the risk premium is so elusive and unmeasurable, I think that investors do place a lot of weight on these myopic results, and not just in the short term; they are interested in what the data say about the long term.

ASNESS: Can we call it Bayesian without priors?

LEIBOWITZ: I think there are priors. I think there really is a Bayesian division going on.

THALER: I want to explain that in the study by Marty Leibowitz, which I so meanly presented, one of the conclusions he reached is that those 20-year numbers look really, really good but that the plan sponsors, the target audience of Marty's study, were going to have to answer some difficult questions over the next two or three years. This problem is an agency problem. The investment committee or whoever is making the investment decisions will get a lot of heat if lots of losses occur on their watch. Typically, the manager running the pension plan is going to be in that job for only two or three years and will then rotate into another job.

ROSS: That agency problem exacerbates this issue even further. With the distinction between the real economy (represented by Rajnish Mehra and John Campbell) and the financial markets, the transmission

[2] Bayes' Law determines a conditional probability (for example, the probability that a person is in a certain occupation conditional on some information about that person's personality) in terms of other probabilities, including the base-rate (prior) probabilities (for example, the unconditional probability that a person is in an occupation and the unconditional probability that the person has a certain personality).

mechanism through institutions becomes even more difficult to explain. Are those who run institutions subject to a variety of psychological vagaries of this sort? Why, if this is an agency problem, has it been so poorly solved to date? It seems to throw up even more theoretical puzzles for us.

LEIBOWITZ: Just a real quick response. That research of mine that Dick Thaler mentioned actually spurred a whole series of papers in which we looked at all kinds of reasons why people would not be 100 percent in stocks. We looked at it from all kinds of different angles—both theoretical and empirical—and we always kept getting this kind of lognormal type of distribution with nice, beautiful tails; it was pretty weird never to see underperformance over long periods of time.

The only conclusion we could finally come to was that, basically, as people peer into the future, they see risk. They are not talking about something with volatility characteristics. They are not talking about return that behaves in a linear fashion. But they see something out there that, basically, fundamentally, scares them. They can't articulate it, but it keeps them from being 100 percent in stocks.

CAMPBELL: I want to defend the relevance of consumption, even in a world with both behavioral biases and agency problems. It would be ludicrous to deny the importance of those phenomena, but even in a world with those phenomena playing a major role, consumption should have a central role in our thinking about risk in financial markets. In the long run, consumption drives the standard of living, which matters to people. So, consumption is a very influential force in investors' decisions.

Can consumption models be applied to endowments, to long-term institutions? I argue that they can, and I have some knowledge of this issue from talking to the managers of the Harvard endowment. Harvard's new president, Lawrence Summers, is trying to make sense of Harvard's spending decisions, which have always been made on an *ad hoc* basis. The endowment maintains very stable spending for a number of years, and then spending rises periodically. Now, in many universities, endowments generally have a smoothed spending rule, so spending levels are linked to past spending levels and the recent performance of the endowment. This rule makes perfect sense if you think that universities get utility from spending but also have some sort of habit formation. It is internal as related to their own history: They hate to cut the budget because it is really painful, the faculty are up in arms, and the students are screaming. And it is related to external situations: They hate to fall behind their competitors. I know that the Harvard endowment managers look very carefully at the management of the Yale endowment, because there's nothing worse than having Yale outperform Harvard. So, habit formation and consumption spending are extremely relevant to endowments. The relationship may be a little more complicated than just saying, "Oh, they have power utility," but you can make sense of the way they think by reference to spending, not only at the micro level but also in terms of the aggregate consumption in the economy.

In the long term, the correlation between consumption growth and the stock market has been quite strong—in the United States and in other countries. And it makes sense. We know that when the economy does well, the stock market does well, and vice versa. There is a link, a correlation, and it represents a form of risk over the longer run.

Aggregate consumption is also an amazingly accurate measure of the sustainable long-term position of the economy. We know that consumption, financial wealth, and labor income are all held together by budget constraints. You can't let your consumption grow indefinitely without some reference to the resources that are available to support it. So, no matter what the behavioral influence is, there is still a budget constraint that is bound to hold consumption, wealth, and income together. You can ask the empirical question when you look at the data: What adjusts to what? If you have a behaviorist's view, you might think that consumption would adjust to the harsh realities of the budget constraint over time. Instead, what seems to happen is that consumption follows a random walk—as if it is set to the level that is sustainable at each point in time. When wealth gets out of line or income gets out of line, they adjust to consumption. So, there's short-term volatility in the financial markets, but when financial wealth is very high relative to consumption, what tends to happen is financial wealth falls. That is just a fact, it does not suggest a particular model, but I think it does suggest the relevance of consumption—together with agency problems and very interesting and important behavioral phenomena—in thinking about the markets.

CORNELL: If consumption is relevant, what type of information would you expect to see flowing through the pipeline of an organization such as TIAA-CREF? How would you expect to see information flowing from the ultimate clients, who are the consumers, into the organization so that the organization can act as the agent on their behalf?

CAMPBELL: Well, TIAA-CREF is running a defined-contribution pension plan. So that, in a sense, information does not have to flow into it. But it seems to me the way to think about defined-benefit pension plans is that they have evolved over a long period of time to reflect the conservatism of the ultimate clients. For example, labor unions negotiate pension arrangements to give their members very stable income in retirement. And even if we accept that agency problems introduce imperfections, it seems to me that the liabilities defined-benefit pension plans have are very stable because of an expressed preference for stable consumption streams.

THALER: The residual claimant to those plans is the company, and the company is supposed to be virtually risk neutral. So, I think the model John Campbell described, which is sort of a habit-formation model, has some plausibility to it as applied to endowments. What is more difficult is to try to use that model in explaining the behavior of the typical plan sponsor of a defined-benefit pension plan.

ROBERT SHILLER: The general public of investors does not, of course, have an economic model like those produced by economists. They do, however, know the definition of stocks and bonds. They know that bondholders get paid first and stockholders are the residual claimants after the bondholders are paid. They know that. The original idea for a stock market was that stockholders are the people who can bear risk and that buying stocks is designed to be a risky contract—which, I think, is very much on investors' minds. So, if we tell them, "Well, in this last century, we were really lucky. Nothing really went wrong. We had five consecutive 20-year periods in which stockholders did really well," I believe that investors then think, rationally, that what we are telling them about low risk for stocks is pretty unconvincing. Investing in stocks is still investing in an asset that was designed for people who can take a lot of risk. There are no promises, and the government isn't going to bail you out if the stock market collapses. The government is perfectly free to throw on a big corporate profits tax; they've moved it up and down. And the shareholder gets no sympathy when the government does so. So, people are rational to be wary, to require a high expected return to take that risk.

ROBERT ARNOTT: I think in this whole discussion of risk premiums we have to be very careful of definitions. In terms of expected returns on stock, there is the huge gap between rational expectation based on a rational evaluation of the sources of return, current market levels, and so forth, versus hope. The investors out there are not investing because they expect to earn TIPS plus 1 percentage point.

And we have a semantic or definitional problem in terms of past *observed* risk premiums, exemplified by the Ibbotson data, between a *normal or unconditional* risk premium, which a lot of the discussion so far seems to have centered on, and the *conditional* risk premium based on current prospects. So, one of the things that we have to be very careful of is that we clarify what we're talking about—past observed risk premiums, normal (unconditional) risk premiums, or conditional premiums based on current prospects.

ROGER IBBOTSON: We have talked mostly about either the behavioral perspective or the classical (or neoclassical) perspective. The classical approach can be interpreted or reinterpreted in many ways as we get more and more sophisticated in our understanding of what the risk aversion might be for the predominant people in the market. And we can put behavioral overlays on classical theory. Ultimately, I think this topic is a rich land for research, and I encourage it, but we are not very close now to getting a fix on an estimate for the risk premium. At first, it appeared that theory suggested low risk premiums, as per Mehra and Prescott (1985), but I think at this stage of the game, using classical theory with behavioral overlays, we can't pinpoint the answer.

THOMAS PHILIPS: An idea that ties together many of the discussions associated with the risk premium is the notion of how to estimate something if you don't have a model or if you're not sure what you are doing. The typical answer is to take the historical average or the sample mean. If we stop to consider why investors buy TIPS at certain times and pull out of hedge funds at other times, we find, more often than not, that the answer is grounded in their use (and abuse!) of the sample mean of the historical returns of that asset class. The trouble is that the sample mean is a terrible estimator. It is easy to show that the sample mean can have huge biases; you just have to vary the risk premium a little bit, for example, or have slightly different economic assumptions, and the estimate and reality diverge sharply. But the sample mean does seem to be the driving force behind most people's behavior. What you observe at cocktail parties or working with clients is this enormous drive toward investing in the asset class with the highest historical return. And I believe it is a fundamentally bad way to think about the problem.

MEHRA: I want to say a couple of things in defense of neoclassical economics. First, for psychological vagaries and other behavioral phenomena to affect prices, the effect has to be systematic. Unless these phenomena occur in a systematic way, the behavior will not show up in prices. So, one has to be very careful about saying, "This is how I behave so I should model market behavior that way." Many of our idiosyncrasies may well cancel out in the aggregate.

Second, most of our economic intuition is actually based on neoclassical models. Ideally, new paradigms must meet the criteria of cross-model verification. Not only must the model be more useful for organizing and interpreting observations under consideration, but it must not be grossly inconsistent with other observations in growth theory, business cycle theory, labor market behavior, and so on. So, I think we should guard against this tendency of model proliferation in which one postulates a new model to explain each phenomenon without regard to cross-model verification. A model that is going to explain one part of reality but then is completely inconsistent with everything else does not make much progress. That is my biggest concern.

ROSS: It seems to me also that there is a vocabulary issue at work here. We have heard the phrase "habit formation" used by many people to mean many different things. On the one hand, the term is used by the behavioralists as though it is some kind of psychological phenomenon. On the other hand, John Campbell uses it as a description of the way universities behave. In either case, it is difficult to tell the difference between whether some fundamental underlying costs that universities face produce a behavioral pattern that looks like habit formation on the preference side but might have nothing to do with it or whether the universities' preferences are perfectly independent across time, are intertemporally independent, but the basic cost structure induces a net behavior that looks like they're concerned about what they did in the past or they are concerned about preserving what they did in the past.

The same is true on the behavioral side. It could well be that there is some fundamental psychological underpinning that we can argue for in terms of habit formation. All you are really saying is that, on the preference side, people don't have adequately separable preferences all the time, that there is some induced link between preferences at one point in time and consumption at one point in time and consumption at another time. There may be some substitutability that we are not capturing in the additive case. So, I think that all of these phenomena have the funny and interesting property that both the neoclassical economist and a purely psychological economist, or behavioral economist (I don't know what the proper phrase is anymore), could wind up saying that the reduced form could be the same for both of them. They just have different ways of getting there.

SHILLER: I think the difference between behavioral economics and classical economics is totally a difference of emphasis. The behaviorists are more willing to look at experimental evidence, a broad array of evidence. Indeed, expected utility is a behavioral model; psychologists also talk about expected utility. So, I think the difference is somewhat methodological; it is not a subject matter difference. It is a question of how willing you are to experiment with different variations.

THALER: Well, habit formation is obviously to some extent a description of preferences. Nothing says it's irrational. The simple additive (and separable) model is the easiest to use, so we naturally started with that model. But you could add completely hypo-rational agents who have preferences that change from one period to another, and you could, of course, have agents who are making the so-called Bayesian forecasts that Marty Leibowitz referred to with those same preferences.

ROSS: There are some exceptions, though, like framing or path dependence. Those tend to be time inconsistent, and time consistency is required in what we typically think of as rational models.

WILLIAM GOETZMANN: A lot of interesting theoretical work is going on, but I want to put in a plug for empirics. Theorists have looked at the price behavior of markets and of individual securities, but a lot of the models have this behavioral component, rational or otherwise, at their heart—whether in identifying the marginal investor or what have you. Yet, we have almost no information about how actual investors behave. Organizations have a lot of that information, but it may never see the light of day for our research purposes. We're beginning to see a little bit of this information cropping up here and there (and sometimes companies that allow us to have it are sorry they did). But imagine the ability to take hundreds of thousands of accounts, time series of accounts, identify the people who seem to exhibit myopic loss aversion, and then test to see whether their behavior has any influence on prices. That work would provide a way to identify whether pathologically behaved people have a short-term or a long-term influence on price behavior. In the long run, empirical study is how we are going to be able to answer some of these questions.

RAVI BANSAL: There is a lot of discussion about preferences, and many of the implementations of this theory lead to the result that asset price fluctuations are a result of cost-of-capital fluctuations. The models do not have much room for expected growth rates. The models build on a long-held belief in economics that consumption growth rates and dividend growth rates are very close to being identically and independently distributed (i.i.d.). It is the notion that most people have. I think we need to rethink that idea. A lot of hidden persistent components are in these growth processes; the realized growth process looks like an i.i.d. process, but if these growth rates have a small persistent component, the ramifications are huge. Small persistent components of any of these growth rates would have dramatic implications for how we think about what is causing asset prices to fluctuate. Statistically, there is actually some evidence to support the view that there are some persistent components in both consumption and growth rates. If such components are put into a model, the unforeseen components can explain equity premiums because consumption goes up at the same time dividends go up. News about consumption and dividend growth rates continuously affects perceptions about long-run expected growth rates, which leads to a lot of asset volatility. This channel is important for interpreting what goes on in asset markets.

Behavior is important, clearly, but understanding the dynamics of cash flows, of consumption, is equally, if not more, important. So, in a paper that Amir Yaron and I wrote (Bansal and Yaron 2000), we allowed for that possibility. And we actually show that when you rely on the Epstein–Zin (1989) preference structure and allow for intertemporal elasticity of substitution to be more than 1.0 (which makes intuitive sense to me), then you can actually get the result that during periods of high anticipated consumption growth rates, the wealth-to-consumption ratio rises. So, in terms of the asset markets, asset valuations will rise simply because of higher expected growth rates. When you require the intertemporal elasticity of substitution to be more than 1.0, then when people expect good times, they want to buy assets. I find this quite intuitive. When you allow for this possibility, you can explain through these neoclassical paradigms a lot of the equity premium and volatility in the market. So, focusing on aggregate output growth is a pretty important dimension.

Historical Results I

Jeremy J. Siegel
Wharton School of Business
Philadelphia

Analysis of the very long term in U.S. markets indicates that average real stock market returns have been about 7 percent and average real T-bond and T-bill returns have been about half that figure. Downward bias in the more recent bond returns and upward bias in recent valuations may be skewing the analysis. Valuations have been rising for three possible reasons: declining transaction costs, declining economic risks, and investors learning that stocks have been undervalued on average throughout history. An analysis of the historical relationships among real stock returns, P/Es, earnings growth, and dividend yields and an awareness of the biases justify a future P/E of 20 to 25, an economic growth rate of 3 percent, expected real returns for equities of 4.5–5.5 percent, and an equity risk premium of 2 percent (200 bps).

Table 1 shows historical returns and the equity risk premium (on a compounded and an arithmetic basis) for the U.S. markets from 1802 through September 30, 2001. The last columns display the equity risk premium based on a comparison with U.S. T-bonds and T-bills, which is just the difference between the real return for stocks and the real return for bonds and bills. I broke out these returns and premiums into the three major subperiods since 1802 and also into 20-year post-World War II periods.

When I wrote the book *Stocks for the Long Run* (Siegel 1998), I was struck by the fact that for all the very long periods (and the definition of "long" is more than 50 years), the average real annual stock market return is just about 7 percent a year, maybe a tad under. This return also holds true for the three subperiods 1802–1870, 1871–1925, and 1926–2001 and for the whole 1946–2001 post-WWII period. (By the way, almost all of the inflation the United States has suffered over the past 200 years has come since World War II, and as we economists should not find surprising, stocks—being real assets—were not at all adversely affected by post-WWII inflation). So, 7 percent appears to be a robust measure of the long-term annual real stock return.

For periods of several decades, however, the real return on stocks can deviate quite a bit from that 7 percent average. Some of those extreme periods since WWII include the bull market of 1946–1965, the bear market of 1966–1981, and the great bull market that lasted from 1982 to the end of 1999. From 1982 through 1999, the average real return on stocks was 13.6 percent, which is double the 200-year average.

That recent experience may color investors' estimates of the equity risk premium today. In the roundtable Discussion for the opening session ["Theoretical Foundations"], there was talk about Bayesian updating, and I do believe that investors place greater weight on the more recent past than we economists think they should. Perhaps investors believe that the underlying parameters of the system have shifted or the model or paradigm has changed or whatever, but I think some of the high expectations investors have for future returns have certainly come from the recent bull market. For many investors, their bull market experience is the only experience they have ever had with the markets, which could certainly pose a problem in the future if excess-return expectations are widespread and those expectations are frustrated.

Table 1. Historical Returns and Equity Premiums, 1802–September 2001

	Real Return						Stock Excess Return over			
	Stocks		Bonds		Bills		Bonds		Bills	
Period	Comp.	Arith.	Comp.	Arith.	Comp.	Arith.	Comp.	Arith.	Comp.	Arith.
1802–2001	6.8%	8.4%	3.5%	3.9%	2.9%	3.1%	3.4%	4.5%	3.9%	5.3%
1871–2001	6.8	8.5	2.8	3.2	1.7	1.8	3.9	5.3	5.0	6.6
Major subperiods										
1802–1870	7.0%	8.3%	4.8%	5.1%	5.1%	5.4%	2.2%	3.2%	1.9%	2.9%
1871–1925	6.6	7.9	3.7	3.9	3.2	3.3	2.9	4.0	3.5	4.7
1926–2001	6.9	8.9	2.2	2.7	0.7	0.8	4.7	6.2	6.1	8.0
Post World War II										
1946–2001	7.0%	8.5%	1.3%	1.9%	0.6%	0.7%	5.7%	6.6%	6.4%	7.8%
1946–1965	10.0	11.4	–1.2	–1.0	–0.8	–0.7	11.2	12.3	10.9	12.1
1966–1981	–0.4	1.4	–4.2	–3.9	–0.2	–0.1	3.8	5.2	–0.2	1.5
1982–1999	13.6	14.3	8.4	9.3	2.9	2.9	5.2	5.0	10.7	11.4
1982–2001	10.2	11.2	8.5	9.4	2.8	2.8	1.7	1.9	7.4	8.4

Note: Comp. = compound; Arith. = arithmetic.

Sources: Data for 1802–1871 are from Schwert (1990); data for 1871–1925 are from Cowles (1938); data for 1926–2001 are from the CRSP capitalization-weighted indexes of all NYSE, Amex, and Nasdaq stocks. Data through 2001 can be found in Siegel (2002).

The annual real bond returns provided in Table 1 show an interesting trend. From 1802 through September 30, 2001, the average annual real T-bond return was 3.5 percent, about half the equity return. In the major subperiods, this return has been trending decidedly downward. Beginning in the 19th century, it was nearly 5 percent; it then fell to 3.7 percent in the 1871–1925 period; it was 2.2 percent for the 1926–2001 period; and since the end of WWII, it has been only 1.3 percent. From 1982 onward, as interest rates and inflation have fallen, bonds have produced a much greater real return than average. When I was studying finance in the 1970s, we learned that both T-bill and T-bond real returns were close to zero. Yet, over the past 20 years, those real returns have definitely risen.

When TIPS were first issued, they were priced to yield a real return of 3.5 percent, which is close to the average 200-year long-term real return of bonds.[1] Investors rightfully ignored the low real returns on bonds of the past 75 years (the period made popular by Ibbotson and the standard benchmark for the profession) in determining the TIPS yield. In fact, in 2000, during the stock market boom, TIPS were priced to yield a real return of almost 4.5 percent. Currently, the long-term TIPS yields have fallen back to a 3.0–3.2 percent range, depending on the maturity.

The real returns on T-bills tell the same story as for bonds, although for bills, the return is generally a bit lower. Of course, bills do not generate the capital gains and losses that bonds do, so in the post-WWII period, bill returns have not fluctuated as much as bonds. Note that from 1982 forward, the annual real return for bills is 2.8 percent, far higher than the nearly zero average real return realized in the previous 55 years. In other words, periods as long as a half century can be quite misleading in terms of predicting future returns.

The problem is that while real stock returns were maintaining their long-term historical average real return of about 7 percent, real bond and bill returns were very low over the past 75 years, particularly up to 1980. Recognition of this phenomenon might help us understand why the equity premium has been so high in data from 1926 to the present.

The equity premium calculated for the past 75 years is biased downward for two reasons—bias in bond returns and bias in equity valuations.

Bias in Bond Returns

First, real historical government bond returns were biased downward over the 1926–2001 period. I say so because all the evidence points to the fact that bondholders simply did not anticipate the inflation of the late 1960s and 1970s. Investors would not have been buying corporate and government bonds of 30-year duration with 3.5 percent coupons (as they did in the 1960s) had they had any inkling of the inflation risk. I attribute part of that ill-fated confidence to the fact that few had a complete understanding of the inflationary implications of the shift from a gold-based to a paper monetary standard.

[1] TIPS are Treasury Inflation-Protected Securities; these securities are now called Treasury Inflation-Indexed Securities.

The gold standard was prevalent during the 19th century and much of the early 20th century when prices were stable over the long term. The United States (and most of the rest of the world) went off the gold standard in the early 1930s, but the effect was not immediately apparent. Although we had a pop of inflation following World War II, inflation was quite low up to the mid-1960s. So, in the 1960s, bond buyers were pricing 30-year bonds as if 30 years later their purchasing power would be nearly the same.

As inflation accelerated, bond buyers began to catch on. Bond yields rose, bond prices fell, and real bond returns were severely depressed. Table 1 shows that during the 15-year period from 1966 through 1981, the real return on bonds was a *negative* 4 percent. That period was long, and its effect is to bias downward the real return of bonds over the longer 1926–2001 period. I thus believe we should use higher real returns on fixed-income assets in our forecasting models, returns that are consistent with the real return on TIPS of 3–4 percent.

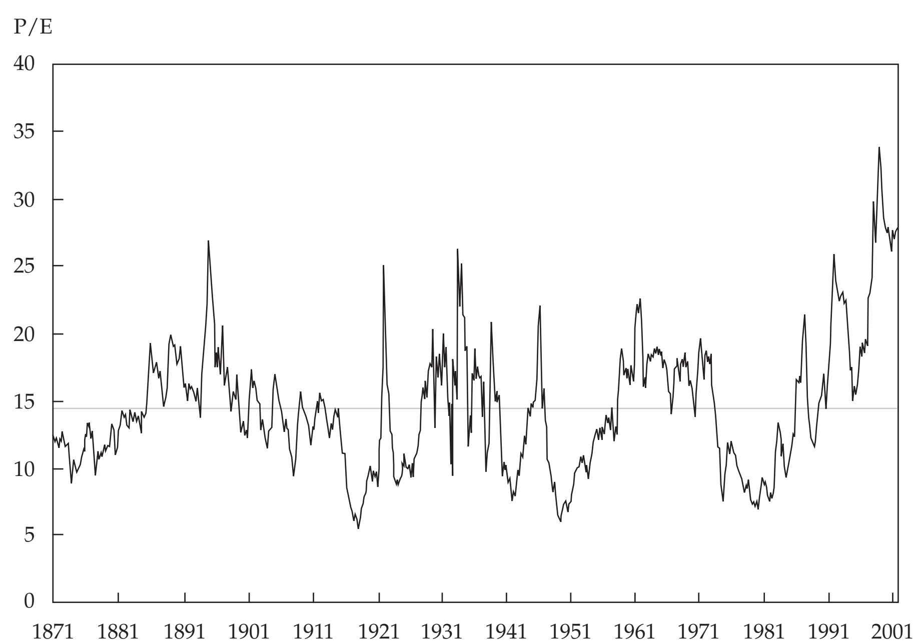
Bias in Equity Valuations

The second reason the equity risk premium is too high is that historical real stock returns are biased upward to some extent. **Figure 1** plots historical P/Es (defined here as current price of the S&P 500 Index divided by the last 12 months of *reported* earnings) from 1871 through September 2001. The straight line is the 130-year mean for the P/E, 14.5. The latest P/E is about 37, surpassing the high that was reached in late 1999 and early 2000. So, the collapse of earnings that we have experienced this year has now sent the P/E to an all-time high.

Let me add a warning here: Part of the incredibly high P/E that we have now is a result of the huge losses in a few technology companies. For instance, JDS Uniphase Corporation wrote down its investments $36 billion in the second quarter of 2001. The write-down was in reported earnings, not in operating earnings, and translates into a 5-point drop in the S&P 500 Index's valuation. So, approach these recent data on reported earnings with caution; $36 billion from just one company's write-down has a huge impact on the market. Some of the technology issues are now essentially out-of-the-money options. When we compute numbers like the P/E of the market, we are adding together all the earnings of all the companies and dividing that into the market value. Because one company has big losses, it sells at option value, but another company with positive earnings can sell at a more normal valuation level. Adding these together might lead to upward biases in P/Es.

Nevertheless, there is no question that P/Es have risen in the past 10 years. If the market's P/E were to return to the historical (since 1871) average of 14.5 tomorrow, the annual real return on equities would fall 50 bps. And if the P/E had always remained at its

Figure 1. Historical Market P/E, 1871–2001



Note: Ending month for 2001 is September.

historical average level but the dividends paid had been reinvested, the annual real return on equities would be 115 bps lower than where it is today. The reason is that much of the real return on equities comes from the times when stock prices are very depressed and the reinvested dividends are able to buy many more shares, boosting stock returns. Much of the historically high returns on stocks has come when the market was extremely undervalued and cash flows were reinvested at favorable prices.

I believe there are several reasons for rising valuation ratios.

- *Declining transaction costs.* One reason for rising valuations is the extensive decline in equity transaction costs. One-way transaction costs were more than 1 percent of the value of the transaction as late as 1975; costs are less than 0.2 percent today.[2] In the 19th and early 20th centuries, the (two-way) costs of maintaining a diversified portfolio could have been as high as 2 percent a year, whereas today indexed funds enable even small investors to be completely diversified at less than 0.2 percent a year.
- *Declining risk.* Another reason for rising valuations may be declining levels of real economic risk as the U.S. economy has become more stable. The increased stability of labor income has enabled workers to accept a higher level of risk in their savings.
- *Investor learning.* We cannot dismiss the fact that investors may have learned about the long-term risk and return characteristics of stocks. If investors have learned that stocks have been chronically undervalued on average, and in particular during recessions and crises, they will be less likely to let prices become undervalued, which leads to higher average valuations.
- *Taxes.* Tax law has become increasingly favorable to equities. And low inflation, because the capital gains tax is not indexed, causes after-tax returns to rise. There has also been a proliferation of tax-deferred savings accounts, although it is not clear whether the taxable or tax-deferred investor sets stock prices at the margin.

[2] Charles Jones of Columbia University discussed declining transaction costs in "A Century of Stock Market Liquidity and Trading Costs" (2001).

Historical Growth Rates

As **Table 2** shows, the real return on stocks has been 7 percent for the 1871–2001 period and is almost exactly the inverse of the P/E. If you divide this period into two subperiods—before World War II and after World War II—the real return for stocks remains roughly 7 percent but the dividend yield drops significantly from the first subperiod to the second, as does the payout ratio, and earnings growth rises.

In his presentation, Cliff Asness mentioned that he could not find in the data an increase in earnings growth when the payout ratio decreased [see "Theoretical Foundations" session]. But his findings are inconclusive because of the confusion between cyclical and long-term trends. In a recession, because dividends remain relatively constant as earnings plummet, payout ratios rise and earnings fall. In the subsequent economic recovery, earnings growth is higher and appears to follow a high dividend payout ratio. But this phenomenon is purely cyclical. Over long periods, a drop in the payout ratio and a drop in the dividend yield are matched almost one-to-one with an increased growth rate of real earnings. I find this relationship comforting because it is what finance theory tells us should happen over long periods of time.

Projecting Real Equity Returns

The link between the P/E and real returns is given by the following equation:

$$\text{Expected future real returns} = \frac{E}{P} + g\left(1 - \frac{RC}{MV}\right),$$

where

E/P = earnings yield, the inverse of the P/E
g = real growth
RC = replacement cost of capital
MV = market value of capital
RC/MV = book-to-market value, or 1/Tobin's q

I will call it the "Tom Philips equation" for projecting the real return of equity (Philips 1999). (I modified the formula somewhat.) According to this equation, if replacement cost does not equal market value, then the link between the P/E and future real returns must be modified. If Tobin's q is not 1, you have to correct

Table 2. Historical Growth Rates, 1871–September 2001

Period	Real Stock Return	Average P/E	Inverse of Average P/E	Real Earnings Growth	Dividend Yield	Real Dividend Growth	Real Capital Gains	Average Payout Ratio
1871–2001	7.06%	14.45	6.92%	1.27%	4.66%	1.09%	2.17%	62.24%
1871–1945	6.81	13.83	7.23	0.66	5.31	0.74	1.32	70.81
1946–2001	7.38	15.30	6.54	2.08	3.78	1.57	3.32	50.75

the earnings yield for the growth rate in the real economy to find expected future real returns. According to the equation, when the market value of equity exceeds the replacement cost of capital, as is the case today, the earnings yield *underestimates* future returns. The reason is that higher equity prices allow companies to fund capital expenditures by floating less equity, thereby reducing the dilution that this investment entails.

How much downward is the earnings yield biased? The Tobin's q on the latest data that I have is about 1.2. It was about 1.5, or even higher, in 2000. With long-run real growth at 3 percent, the last term, $g[1 - (RC/MV)]$, adds about 50 bps to the forecast of real return going forward. It added more in 2000 because Tobin's q was higher. So, if the P/E settles down to 20 (and I believe that a future P/E should not be back at 14 or 15 but that a higher P/E is justified for the reasons I listed previously) and we emerge from the recession, then in terms of a long-term trend, E/P will be about 5 percent. Add the half a percentage point for the cheaper investment to maintain capital and you get a 5.5 percent expected real rate of return for equities. If the P/E is 25 in the future, with 1/25 = 4 percent, adding the growth correction produces an expected real return for equities of 4.5 percent.

Keep in mind that TIPS are now priced to yield a real return of about 3 percent. So, because I believe that the long-run P/E in the market will settle between 20 and 25, the real future equity return is about 5 percent and the equity risk premium will be 2 percent (200 bps).

Historical Results I

Jeremy J. Siegel

Wharton School of Business
Philadelphia

SUMMARY

by Peter Williamson

Amos Tuck School of Business Administration
Dartmouth College, Hanover, New Hampshire

Jeremy Siegel began his presentation with a table of U.S. market historical returns and excess equity returns for five time periods. **Table 1** provides returns for two very long periods, from the 1800s to September 30, 2001, for three subperiods making up the long periods, and for five post-World War II periods. What is most noteworthy in Table 1 is the geometric (compounded) average real return on stocks of close to 7 percent for the long periods, for both of the major subperiods, and for the 1946–2001 period. Equally significant are the wide deviations above and below 7 percent over quite long periods after World War II, especially since 1982. The geometric average for 1982–1999 was 13.6 percent, and Siegel concluded that this high average return has influenced the high expectations of today's investors, many of whom have little experience of the pre-1982 period.

Table 1 indicates that average real U.S. T-bond returns fell over the years until the post-1982 period, when very high returns resulted from a decline in interest rates. The 1926–2001 period produced a 2.2 percent average real bond return, biased downward by unexpected inflation in the 1960s and 1970s. Siegel observed that TIPS were priced originally in 1997 at about 3.375 percent, with the yield later rising to about 4 percent, and are now down to about 3 percent.[1] This pricing is close to the 200-year average real return on bonds.

[1] TIPS are Treasury Inflation-Protected Securities; these securities are now called Treasury Inflation-Indexed Securities.

Table 1. Historical Returns and Equity Premiums, 1802–September 2001

	Real Return						Stock Excess Return over			
	Stocks		Bonds		Bills		Bonds		Bills	
Period	Comp.	Arith.	Comp.	Arith.	Comp.	Arith.	Comp.	Arith.	Comp.	Arith.
1802–2001	6.8%	8.4%	3.5%	3.9%	2.9%	3.1%	3.4%	4.5%	3.9%	5.3%
1871–2001	6.8	8.5	2.8	3.2	1.7	1.8	3.9	5.3	5.0	6.6
Major subperiods										
1802–1870	7.0%	8.3%	4.8%	5.1%	5.1%	5.4%	2.2%	3.2%	1.9%	2.9%
1871–1925	6.6	7.9	3.7	3.9	3.2	3.3	2.9	4.0	3.5	4.7
1926–2001	6.9	8.9	2.2	2.7	0.7	0.8	4.7	6.2	6.1	8.0
Post World War II										
1946–2001	7.0%	8.5%	1.3%	1.9%	0.6%	0.7%	5.7%	6.6%	6.4%	7.8%
1946–1965	10.0	11.4	–1.2	–1.0	–0.8	–0.7	11.2	12.3	10.9	12.1
1966–1981	–0.4	1.4	–4.2	–3.9	–0.2	–0.1	3.8	5.2	–0.2	1.5
1982–1999	13.6	14.3	8.4	9.3	2.9	2.9	5.2	5.0	10.7	11.4
1982–2001	10.2	11.2	8.5	9.4	2.8	2.8	1.7	1.9	7.4	8.4

Note: Comp. = compound; Arith. = arithmetic.

Sources: Data for 1802–1871 are from Schwert (1990); data for 1871–1925 are from Cowles (1938); data for 1926–2001 are from the CRSP capitalization-weighted indexes of all NYSE, Amex, and Nasdaq stocks. Data through 2001 can be found in Siegel (2002).

Real returns on T-bills averaged 2.8 percent from 1982 to September 30, 2001—a surprisingly high return for those who were accustomed to the popular position a few years ago that bills offered a zero real rate.

The equity excess return, over both bonds and bills, from 1982 to 1999 and from 1926 to 2001 was much higher than it had been for the long periods, and Siegel commented that the 3–4 percent range that characterized the longer periods was probably reasonable for the long term.

Figure 1 shows the historical P/E of the equity market (calculated from the current price and the last 12 months of reported earnings) for 1871 through September 2001. The collapse of earnings recently pushed the ratio up to 37, past the high of 1999. The average P/E over 130 years was only 14.5. Siegel noted that huge losses in only a few technology companies accounted for a lot of this valuation change. Real stock returns have been biased upward with the rise in P/Es. If the market's P/E were to return to the historical (since 1871) average overnight, the real return on equities would fall 50 bps. And if the P/E had always remained at its average level, without reinvestment of the dividends that actually were paid, real returns would be 115 bps lower than where they are today.

Siegel offered three reasons for rising P/E multiples. First is declining transaction costs, which could have accounted for 2 percent a year in the 19th and early 20th centuries and are presently perhaps as low as 0.2 percent for a one-way trade. Second is declining real economic risk. And third is investors learning more about the long-term risk characteristics of common stocks, especially investors realizing that there are periods of significant undervaluation.

Table 2 shows the relationships among real stock returns, P/Es, earnings growth, and dividend yields. For 130 years, the real stock return, averaging 7 percent, has been almost exactly the earnings yield (reciprocal of the P/E). The periods before and after World War II show close to the same 7 percent. Faster post-WWII earnings growth matches the decline in the dividend yield and the rise in retained earnings. Siegel noted that this long-term relationship between payout and growth is in accord with theory, but over short periods, the change in earnings growth does not always accompany a change in dividend yield.

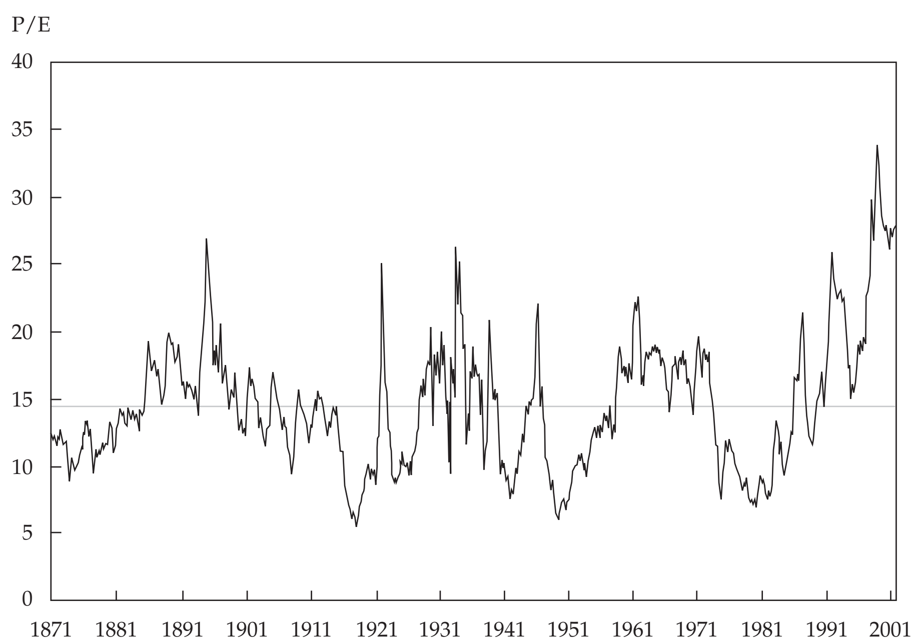
The link between P/E and real returns is given by

$$\text{Expected future real returns} = \frac{E}{P} + g\left(1 - \frac{RC}{MV}\right),$$

where

E/P = earnings yield, the inverse of the P/E
g = real growth
RC = replacement cost of capital
MV = market value of capital
RC/MV = book-to-market value, or 1/Tobin's q

Figure 1. Historical Market P/E, 1871–2001



Note: Ending month for 2001 is September.

Table 2. Historical Growth Rates, 1871–September 2001

Period	Real Stock Return	Average P/E	Inverse of Average P/E	Real Earnings Growth	Dividend Yield	Real Dividend Growth	Real Capital Gains	Average Payout Ratio
1871–2001	7.06%	14.45	6.92%	1.27%	4.66%	1.09%	2.17%	62.24%
1871–1945	6.81	13.83	7.23	0.66	5.31	0.74	1.32	70.81
1946–2001	7.38	15.30	6.54	2.08	3.78	1.57	3.32	50.75

Tobin's q is currently about 1.2, and the long-run growth rate, g, is about 3 percent, so the term $g[1 - (RC/MV)]$ adds about 0.5 percentage point to the E/P term. At a P/E of 20, appropriate for today, the expected real return is about 5.5 percent. At a P/E of 25, it is 4.5 percent. With the TIPS return at about 3 percent and a P/E of 20 to 25, Siegel's equity risk premium is about 2 percent (200 bps).

Historical Results II

Bradford Cornell

University of California
Los Angeles

The basic investment and constant-growth models, used with some justifiable simplifying assumptions about the U.S. market, indicate that the earnings growth rate cannot be greater than the GNP growth rate because of political forces and that the expected return, or cost of capital, in the long run should unconditionally be about 1.5 times the dividend-to-price ratio plus GNP growth. Adding reasonable assumptions about inflation produces a finding that equity risk premiums cannot be more than 3 percent (300 bps) because earnings growth is constrained by the real growth rate of the economy, which has been in the 1.5–3.0 percent range. In a consideration of today's market valuation, three reasons for the high market valuations seem possible: (1) stocks are simply seen as less risky, (2) valuation of equities is fundamentally determined by taxation, or (3) equity prices today are simply a mistake. A research question that remains and is of primary interest is the relationship between aggregate stock market earnings and GNP.

The very basic investment and constant-growth models from introductory finance courses can be used to interpret the long-run unconditional historical data on returns. So, let's begin with the basic model:

$$\frac{E_{t+1}}{E_t} = 1 + [(b)(ROE)],$$

where

E = earnings
b = the retention rate
ROE = return on equity

So that, with investment at time t denoted by I_t,

$$ROE = \frac{E_{t+1} - E_t}{I_t}$$

and

$$b = \frac{I_t}{E_t};$$

therefore, the growth rate of earnings is

$$(b)(ROE) = \frac{E_{t+1} - E_t}{E_t}.$$

This model implies that the growth rate in earnings is the retention rate times the return on equity, $(b)(ROE)$. In discussing the models, I would like to stress an important point: If you are interpreting the growth in earnings as being the retention rate times the return on equity, you have to be very careful when you are working with historical data. For example, does the retention rate apply only to dividends or to dividends and other payouts, such as share repurchases? The distinction is important because those proportions change in the more recent period. And if you make that distinction, you have to make a distinction between aggregate dividends and per share dividends because the per share numbers and the aggregate numbers will diverge. In working with the historical data, I have attempted to correct for that aspect.

What simplifying assumptions can be made to work with the unconditional data? I have made some relatively innocuous simplifying assumptions. First, that b should adjust until the cost of capital equals the ROE at the margin. To be very conservative, therefore, I will assume that the ROE equals the cost of capital, or expected returns, in the aggregate. The problem that arises is: What if the retention rate times the cost of capital (that is, the minimal expected return on equity), bk, is greater than GNP growth? The second assumption deals with this possibility: I assume bk cannot be greater than GNP growth because political forces will come into play that will limit the ROE if earnings start to rise as a fraction of GNP.

The relationship between aggregate earnings and GNP is one of the research questions that I have been unable to find interesting papers on—perhaps because I have not searched well enough—but I want to bring up the subject to this group. It seems to me that if aggregate earnings start to rise, and Robert Shiller mentioned several reasons why it can happen [see the "Current Estimates and Prospects for Change" session], then tax rates can change, antitrust regulation can change (one of Microsoft's problems probably was that it was making a great deal of money, which is an indication that some type of regulation may be necessary), labor regulation can change, and so forth. And these variables can change *ex post* as well as *ex ante*. So, once a company starts making superior returns using a particular technology, the government may step in *ex post* and limit those returns. The critical research question is how earnings relate to GNP.

The constant-growth model is

$$P = \frac{D}{k-g}$$

or

$$k = \frac{D}{p} + g,$$

where

P = price
D = dividends
k = cost of capital
g = growth rate

What I am going to do is just an approximation because I am going to work with aggregate, not per share, data. I am going to assume that total payouts are 1.5 times dividends.[1] Payouts will probably be lower in the future, but if I work with aggregate payouts, then g should be the growth rate in aggregate potential payouts, which I will characterize as earnings.

One of the implications of the simplifying assumptions I have made, and it relates to the data that Jeremy Siegel just produced ["Historical Results I"], is that the expected returns on stocks should be equal to the earnings-to-price ratio. (In the more complicated equations, you have situations in which the ROE is not exactly equal to expected returns, but for my long-run data, the simplifying assumption that earnings yield equals the expected ROE is fine.) So, with these assumptions,

$$\begin{aligned} P &= \frac{D}{k-g} \\ &= \frac{D}{k-bk} \\ &= 1-(b)\left(\frac{E}{1-b}\right)(k) \\ &= \frac{E}{k} \end{aligned}$$

or

$$k = \frac{E}{P}.$$

A further implication is that if g is constrained to be close to the growth of GNP, then it is reasonable to substitute GNP growth for g in the constant-growth model. The implication of this conclusion is that the expected return, or cost of capital, in the long run should unconditionally be about 1.5 times the dividend-to-price ratio plus GNP growth:

$$k = 1.5\frac{D}{P} + \text{GNP growth}.$$

With this background, we can now look at some of the data.

Earnings and GNP

Figure 1 allows a comparison of dividends/GNP and (after-tax) earnings/GNP for 1950 through July 2001.[2] The data begin in 1950 because Fama believed that the data before then were unreliable. Figure 1 shows that, historically, earnings have declined as a fraction of GNP in this period. My assumption that earnings keep up with GNP works from about 1970 on, but I am looking at the picture in Figure 1 in order to make that conclusion. The ratio of earnings to GNP depends on a lot of things: the productivity of labor, capital, the labor-to-capital ratio, taxes, and (as I said earlier) a host of political forces. Figure 1 shows that earnings have, at best, kept up with GNP.

[1] This choice is based on recent findings by Jagannathan, Stephens, and Weisbach (2000) that we are seeing significant payouts today.

[2] These data were provided by Eugene Fama, who attributed them to Robert Shiller.

Figure 1. S&P 500 Earnings and Dividends to GNP, 1950–July 2001

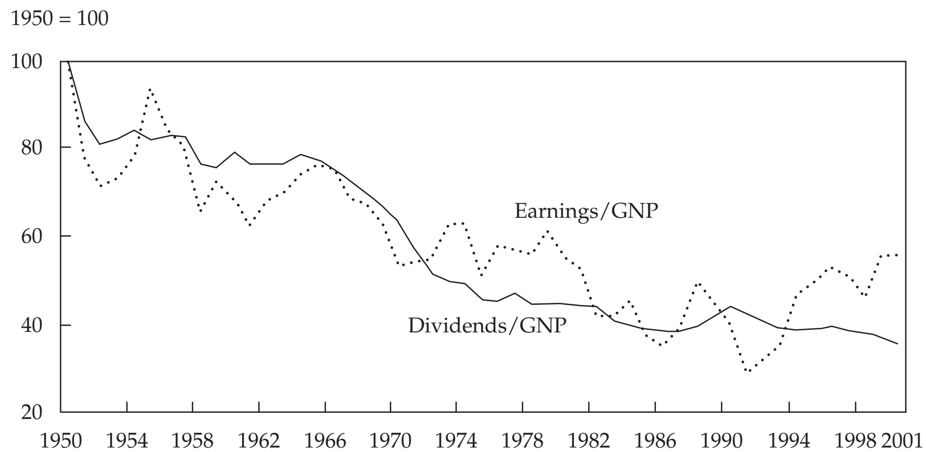


Table 1 gives the arithmetic average data for growth rates in GNP, earnings, and dividends for two periods: 1951–2000 and 1972–2000. (I used the 1972–2000 period because it mirrors the same period shown in Figure 1.) The earnings growth rates are so much more volatile than the dividend growth rates. And because of the volatility effect on arithmetic averages, GNP and earnings exhibit very similar growth rates from the early 1970s to the present. Dividends (and Table 1 shows the growth rate of actual dividends, not payouts) have grown much less than earnings for two reasons: First, dividends are less volatile, and second, dividend substitution is occurring. Corporations are not providing shareholders the same constant fraction of earnings (in the form of dividends) that they were in the past.

Despite the 1972–2000 data, it seems to me that earnings are not going to grow as fast as or faster than GNP in the future. This notion seems to be consistent with long-term historical data, and it fits my view of how politics works on the economy. If you accept that notion, it has immediate implications for the future.

Table 1. Historical Growth Rates of GNP, Earnings, and Dividends: Two Modern Periods

Period/Measure	GNP	Earnings	Dividends
1951–2000			
Mean	3.21%	2.85%	1.07%
Standard deviation	2.89	14.29	4.13
1972–2000			
Mean	2.62%	3.79%	0.96%
Standard deviation	2.94	15.72	3.58

Note: Growth rates for earnings and dividends are based on aggregate data.

First, under any reasonable underlying assumptions about inflation, equity risk premiums cannot be much more than 3 percent (300 bps) because the earnings growth rate is constrained unconditionally in the long run by the real growth rate of the economy, which has been in the range of 1.5–3.0 percent. Second, as **Table 2** shows, for an S&P level of about 1,000, you simply cannot have an equity risk premium any higher than 2 percent, 2.5 percent, or (at most) 3 percent.

Table 2. Value of the S&P 500 Index Given Various Real (Earnings or GNP) Growth Rates and Equity Risk Premiums

Real Growth Rate	Equity Risk Premium						
	2.0%	2.5%	3.0%	4.0%	5.0%	6.0%	7.0%
1.5%	845	724	634	507	423	362	317
2.0%	1,014	845	724	563	461	390	338
2.5%	1,268	1,014	845	634	507	423	362
3.0%	1,690	1,268	1,014	724	563	461	390

Assumptions: Inflation = 3 percent; long-term risk-free rate = 5.5 percent; payout = 1.5(S&P 500 dividend). The S&P 500 dividend used in the calculation was $16.90, so P = 1.5($16.90)/($k - g$), where k = 5.5 percent (the risk-free rate minus 3 percent inflation plus the risk premium) and g = real growth rate.

Valuation

Why is the market so high? As an aside, and this concern is not directed toward our topic today of the equity risk premium, but I think it is an interesting question: Why is the market where it is today relative to where it was on September 10 or September 9 or just before the events of September 11, 2001? The market then and now is at about the same level. Almost every economist and analyst has said that the September 11 attacks accelerated a recession, that they changed perceptions of risk, and so forth. It is curious to me that such a situation does not seem to be reflected in market prices.

But in general, why is the market so high? I believe three possible explanations exist. One idea, and I consider it a "rational" theory, is that stocks are simply seen as less risky than in the past. I do not know whether the behavioral theories are rational or not, in the sense that prices are high because of behavioral phenomena that are real and are going to persist. If so, then those phenomena—as identified by Jeremy Siegel and Richard Thaler [see the "Theoretical Foundations" session]—are also rational. In that case, the market is not "too high"; it is not, in a sense, a mistake. It is simply reflecting characteristics of human beings that are not fully explained by economic theories.

Another rational explanation has been given less attention but is the subject of a recent paper by McGrattan and Prescott (2001). It is that the valuation of equities is fundamentally determined by taxation. McGrattan and Prescott argue that the move toward holding equities in nontaxable accounts has led to a drop in the relative tax rate on dividends. Therefore, stock prices should rise relative to the valuation of the underlying capital and expected returns should fall. This effect is a rational tax effect.

Both this theory and the theory that stocks are now seen as less risky say that the market is high because it should be high and that, looking ahead, equities are going to have low expected returns, or low risk premiums—about 2 percent—but that investors have nothing to worry about.

The final explanation, which I attribute to John Campbell and Robert Shiller, focuses on the view that equity prices today are simply a mistake. (I suppose mistakes are a behavioral phenomenon, but presumably, they are not as persistent as an underlying psychological condition.) Now, when people realize they have made a mistake, they attempt to correct the behavior. And those corrections imply a period of *negative* returns from the U.S. equity market before the risk premium can return to a more normal level.

Closing

To close, I want to repeat that, to me, the fundamental historical piece of data that needs more explanation is the relationship between the aggregate behavior of earnings and GNP—what it has been in the past and what it can reasonably be going forward. This relationship is interesting, and I look forward to hearing what all of you have to say about it. In my view, it is the key to unlocking the mystery of the equity risk premium's behavior.

Historical Results II

Bradford Cornell

University of California
Los Angeles

SUMMARY

by Peter Williamson

Amos Tuck School of Business Administration
Dartmouth College, Hanover, New Hampshire

To interpret long-run unconditional features of historical returns, Bradford Cornell began with the following basic model:

$$\text{Earnings growth} = (b)(ROE),$$

where b is the rate at which earnings are retained and ROE is return earned on equity. He noted that we have to be careful when working with historical data in this model. For example, does payout apply only to dividends or to dividends and other payouts, such as share repurchases? And we need to distinguish between aggregate dividends and per share dividends. The two have been diverging.

Now, b should adjust until ROE at the margin equals k, the cost of capital. Cornell assumed that $k = ROE$ in the aggregate, but a critical question is how earnings relate to GNP (see **Figure 1**). What if bk is greater than GNP growth? Cornell assumed that political forces—such as taxation, antitrust laws, and labor regulations—would affect *ex ante* and *ex post* returns in such a way as to bring about

$$(b)(ROE) = bk \le \text{GNP growth}.$$

The constant-growth model is

$$P = \frac{D}{k-g}$$

or

$$k = \frac{D}{P} + g$$

where

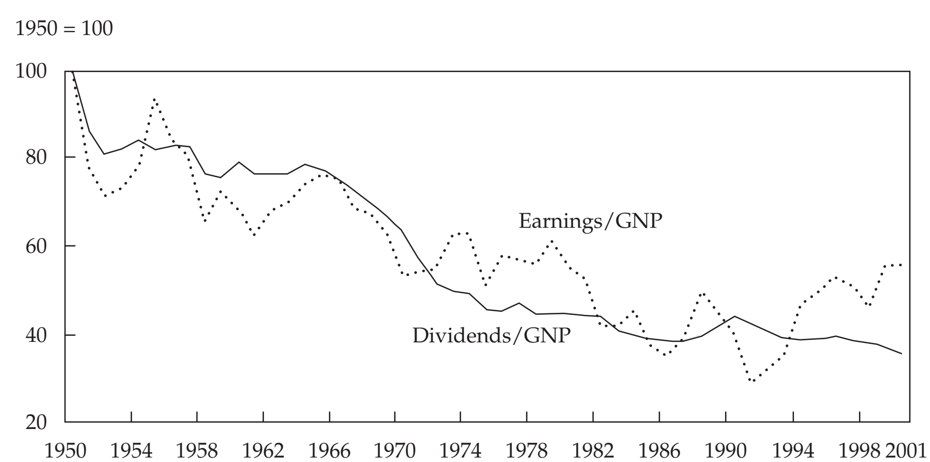
P = price
D = dividends
k = cost of capital
g = growth rate

Because D is equal to $E(1 - b)$ and g is equal to bk, the constant-growth model becomes, in real terms,

$$P = \frac{E}{k}$$

or

Figure 1. S&P 500 Earnings and Dividends to GNP, 1950–July 2001



$$k = \frac{E}{P}.$$

Cornell had so far been working with aggregates, but share repurchases and other nondividend cash flows between companies and their shareholders should be considered. So, he assumed that the total of cash distributions is approximately 1.5D.

Finally, if g is constrained to be close to GNP growth, then $k = 1.5(D/P) + \text{GNP growth}$.

Table 1 shows that since 1950, aggregate S&P 500 Index earnings and dividends have both grown less than GNP, although from 1972 to 2000, earnings actually grew faster. (Earnings may appear to have kept up with or even exceeded GNP because of the high volatility of the earnings, which leads to high arithmetic average rates of growth for the same geometric averages.) The dividend growth rates have been lower because of falling payout ratios. The picture conveyed to Cornell is that earnings growth will not exceed GNP growth in the future. (The relationship of earnings to GNP is an interesting measure having to do with, among other things, the productivity of labor and capital.)

Table 1. Historical Growth Rates of GNP, Earnings, and Dividends: Two Modern Periods

Period/Measure	GNP	Earnings	Dividends
1951–2000			
Mean	3.21%	2.85%	1.07%
Standard deviation	2.89	14.29	4.13
1972–2000			
Mean	2.62%	3.79%	0.96%
Standard deviation	2.94	15.72	3.58

Note: Growth rates for earnings and dividends are based on aggregate data.

Finally, putting together an inflation assumption of 3 percent, a long-term nominal risk-free rate of 5.5 percent, and the relationships developed previously produces **Table 2**. An example of the calculations for Table 2 under the assumptions given in the table is as follows: At real growth of 3 percent and with a risk premium of 2.5 percent, $P = [1.5(\$16.90)]/(0.055 - 0.03 + 0.025 - 0.03) = \$1{,}268$. What Table 2 indicates is that as long as g is limited by GNP growth of 1.5–3.0 percent, the equity risk premium must be no more than about 3 percent to be consistent with an S&P 500 of about 1,000.

Cornell asked why, in general, is the market so high? (In particular, he questioned why the market is currently at the level of pre-September 11, 2001, if, as so many say, the events of that date accelerated a recession and changed perceptions of risk.) One explanation is that investors see the market generally as less risky than in the past. Cornell found that explanation rational. Another rational explanation is that the value of equities is fundamentally determined by taxation. Perhaps the market's level is explained by human behavior that is rational but for which we have no explanation. Both propositions imply that there is nothing wrong with current prices. Still, another explanation is that equity prices are a mistake and that a downward correction will produce negative returns before a normal risk premium prevails.

A key subject on which we might focus is the relationships among aggregate earnings, GNP, and other economic variables.

Table 2. Value of the S&P 500 Index Given Various Real (Earnings or GNP) Growth Rates and Equity Risk Premiums

Real Growth Rate	Equity Risk Premium						
	2.0%	2.5%	3.0%	4.0%	5.0%	6.0%	7.0%
1.5%	845	724	634	507	423	362	317
2.0%	1,014	845	724	563	461	390	338
2.5%	1,268	1,014	845	634	507	423	362
3.0%	1,690	1,268	1,014	724	563	461	390

Assumptions: Inflation = 3 percent; long-term risk-free rate = 5.5 percent; payout = 1.5(S&P 500 dividend). The S&P 500 dividend used in the calculation was \$16.90, so $P = 1.5(\$16.90)/(k - g)$, where k = 5.5 percent (the risk-free rate minus 3 percent inflation plus the risk premium) and g = real growth rate.

Historical Results: Discussion

Ravi Bansal (*Moderator*)
Robert Arnott
Clifford Asness
John Campbell
Peng Chen, CFA
Bradford Cornell
William Goetzmann
Campbell Harvey
Roger Ibbotson
Martin Leibowitz
Rajnish Mehra
Thomas Philips
William Reichenstein, CFA
Stephen Ross
Robert Shiller
Jeremy Siegel
Kevin Terhaar, CFA
Richard Thaler

RAVI BANSAL (Moderator)

I would like to make a couple of observations. One aspect that we could consider is the time-series evidence on aggregate consumption volatility. I am thinking of consumption as a way to measure economic uncertainty in the data, but it can be done by other means as well. The time-series evidence suggests that a decline in conditional volatility has without doubt occurred over the past 40 years or so. This reduced volatility suggests that there should be some decline in risk premiums. Another aspect that could be considered, which Steve Ross mentioned earlier, is that much of the risk premium discussion draws on the cross-sectional evidence. It is where a lot of the bodies are buried in terms of understanding where risks are coming from.

We heard some debate in the first session [“Theoretical Foundations”] about whether consumption models are plausible or not, and my view is that consumption data are not in a usable form for explaining the cross-sectional differences, although there may be new evidence in this regard. The consumption models can actually go a long way, however, in explaining the difference in the risk premiums on different assets. In fact, in “Consumption, Dividends, and the Cross-Section of Equity Returns” (Bansal, Dittmar, and Lundblad 2001), we show that if you take the earnings growth or the dividend growth of different portfolios and regress actual growth on historical (say, the past 25–30 years) consumption growth smoothed for 12 or 14 quarters, and if you consider (what has almost become the industry benchmark) 10 portfolios composed on the basis of size, 10 on momentum, and 10 on the book-to-market ratio, you will see that the regression coefficient almost entirely lines up with the *ex post* excess returns on these different assets. So, for example, the regression coefficient of extreme “loser” momentum portfolios is negative and that of “winner” portfolios is strongly positive. The value stocks have a very high exposure to the consumption growth rate, and what I call the loser value stocks—that is, the growth stocks—have a low exposure, which maps the differences in equity premiums also. So, there is a link between consumption and risk premiums, which creates a *prima facie* case for aggregate economic uncertainty, defined as consumption, being a very useful measure.

The cross-sectional evidence also highlights that what determines the risk premium on an asset is “low-frequency” movements (long-run growth prospects) and the exposure of different portfolios to them. Long-run growth prospects are the key source of risk in the economy.

Still, a puzzle remains because the equity market risk premiums have decreased—to 2 percent, 2.5 percent, or so on—and of course, people disagree about what the risk premium is. It seems to me that the right way to approach the equity risk premium puzzle is through the Sharpe ratio on the market. If we argue that the risk premium has fallen, then the Sharpe ratio is quite likely to have fallen also.

CLIFFORD ASNESS: If I understood correctly, Jeremy Siegel was saying that Rob Arnott and I were picking up a short-term mean-reversion effect that is not relevant over the long term. I would like to make two points: First, we were forecasting over several decades and found a pretty strong negative relationship between the retention rate and real earnings growth. So, Jeremy, if this relationship reverses itself in the longer term, we should find a very, very strong positive relationship later. Yes? Second, in the draft of our paper (Arnott and Asness 2002), which has only been seen by Rob, me, and a few people we trusted not to laugh at us, we tested the relationship against other proxies for pure, univariate mean reversion in earnings growth—prior growth, growth versus a 20-year average—added to the equation. We still found over a 10-year horizon (we would like to have used a longer horizon but were trying to avoid having too few periods) that the relationship is very negative. Therefore, I have a hard time believing that over longer periods the relationship is going to be very positive. We did find that simple measures of mean reversion and earnings do not knock out the relationship. I am curious about the data you were using and what you are citing in the longer term. Maybe we can reconcile the apparent differences.

JEREMY SIEGEL: Well, I did not run the tests that you did. I just know that there is very strong evidence from cycles. In recessions, the payout ratio goes very high because companies choose to maintain the same level of dividends they were paying before the recession, and earnings drop. Then, subsequent growth in real earnings is very high because it is happening relative to the slow or negative growth experienced during the recession. The same phenomenon, but in the opposite direction, occurs during and after an economic boom. For these reasons, I found in the two long periods, 1871–1945 and then 1946–2000, that the decrease in the dividend yield during each period was matched by an increase in real earnings growth [see Siegel's Table 2]. The result is the same approximate 7 percent real return in the later period as in the earlier period, which is comforting from a theoretical point of view. Otherwise, we would have to turn to such theories as that "companies that retain more earnings must be totally wasting them because the companies do worse after the earnings retention." That theory is very much a concern.

JOHN CAMPBELL: I want to focus attention on an issue that is in Jeremy Siegel's tables but which he didn't talk about in his presentation—the geometric versus the arithmetic average. This issue is one that causes people's eyes to glaze over. It seems a pedantic thing, like worrying about split infinitives—the sort of thing that pedantic professors do but other people shouldn't bother about. But it is actually an important issue for risky assets because the difference between the arithmetic and the geometric average is on the order of about half the variance, which for stocks, is about 1.5–2.0 percent. That's a big difference, and it shows up in Jeremy's tables very clearly. So, when we're bandying about estimates of the equity premium and we say, "Maybe it's 2 percent; maybe it's 3 percent," clearly the difference between these two averages is large relative to those estimates.

Which is the right concept, arithmetic or geometric? Well, if you believe that the world is identically and independently distributed and that returns are drawn from the same distribution every period, the theoretically correct answer is that you should use the arithmetic average. Even if you're interested in a long-term forecast, take the arithmetic average and compound it over the appropriate horizon. However, if you think the world isn't i.i.d., the arithmetic average may not be the right answer.

As an illustration, think about a two-lane highway to an airport. Suppose that to increase traffic capacity, you repaint the highway so that it has three, narrower lanes. Traffic capacity is thus increased by 50 percent. But suppose the lanes are now too narrow, causing many accidents, so you repaint the highway with only two lanes. Arithmetically, the end result appears to be a great success because the net effect is an increase in capacity. A 50 percent increase in capacity has been followed by only a 33.3 percent decrease. The arithmetic average of the changes is +8.5 percent. So, even though you're back to your starting point, you delivered, on average, an 8.5 percent increase in traffic capacity. Obviously, that's absurd. In this case, the geometric average is the right measure. The geometric average calculates a change in capacity to be zero, which is the correct answer; nothing has been accomplished with the lane rearrangement and reversal.

The difference between the i.i.d. case and the highway story is that in the highway story, you have extreme negative serial correlation. You could get to –33.3 percent in the end only by having had the +50 percent and –33.3 percent occur on a higher base than +50 percent. So, the geometric average is the correct measure to use in an extreme situation like the highway illustration.

I think the world has some mean reversion. It isn't as extreme as in the highway example, but whenever any mean reversion is observed, using the arithmetic average makes you too optimistic. Thus, a measure somewhere between the geometric and the arithmetic averages would be the appropriate measure.

BRADFORD CORNELL: You see that difference in the GNP and earnings data. Although the ratio of earnings to GNP is falling from 1972 on [see Cornell's Table 1], the growth rate of earnings is higher as an *arithmetic* mean precisely for the reason you suggest.

CAMPBELL: Right, right. Mean reversion has the effect of lowering the variance over long horizons, which is, of course, a major theme of Jeremy Siegel's work. And you could imagine taking the geometric average and then adding half of long-term variance to get an appropriate long-term average.

SIEGEL: That's a good point. You discussed in your new book with Lewis Viceira (Campbell and Viceira 2002) whether we should use the arithmetic or the geometric average and that when mean reversion occurs, we perhaps have more reason to use the geometric average. I've found in my data that at 30-year horizons, the standard deviation is about half the number that i.i.d., random walk theory would predict. So, you can actually add half the variance to the geometric average and use that number as the appropriate arithmetic risk premium on long horizons.

CAMPBELL: It was striking that you did focus your presentation on the geometric average. A lot of the other calculations that have been presented here today evolve out of these deterministic models in which no distinction is made between geometric and arithmetic calculations. But I think that when you face randomness, as we do in the world, you have to think about this issue.

ROBERT ARNOTT: I had just a quick follow-up to Cliff Asness's question about the link between payout ratios and earnings growth. I think one possible source of the difference that we're seeing is not the time horizon but that, in Jeremy Siegel's work, if I understand correctly, he is looking at the *concurrent* payout ratio versus earnings growth. Cliff Asness and I are looking at *leading* payout ratio versus *subsequent* earnings growth; in effect, we're using the payout ratio as a predictor of earnings growth.

ASNESS: I'll add one thing to that: What Jeremy Siegel is saying is that a high and falling dividend yield is replaced by increased earnings growth over that period. What Rob Arnott and I are saying is that perhaps there is mean reversion but if you look at the start of that period, the high dividend yield was leading to a high payout ratio, which tended to forecast the declining actual earnings growth. So, I think we're actually saying the same thing. That's a limb I'm going to go out on.

CAMPBELL HARVEY: One thing that completely baffles me is the TIPS yield right now. The breakeven inflation rate for 10 years is about 1.2 percent. Brad Cornell showed that valuation table [Cornell's Table 2] with a reasonable assumption of inflation at 3 percent. And Jeremy Siegel's Table 1 showed the historical data in terms of real bond return, which was significantly higher on average than 1.5 percent. It just seems there's something going on with TIPS that I don't understand. For me, an inflation rate of 1.2 percent over 10 years doesn't seem reasonable.

PENG CHEN: It depends on how you define the equity risk premium. Some define the equity risk premium in relation to the real return earned on TIPS. It's a good observation, but TIPS is a new asset class, started just several years ago. The TIPS market is still immature; the market size is relatively small. So, I'm not sure how much inference you should draw by just looking at the current yield. A current yield of 3 percent doesn't mean that the real interest rate is 3 percent. If you had followed the TIPS market for a while, you probably would have heard rumors that the U.S. Treasury Department is going to suspend issuing TIPS—which would have a huge impact on how TIPS behave in the marketplace. So, we need to be careful when using TIPS as part of the benchmark in trying to calculate the actual risk premium.

SIEGEL: On that issue, I think there is a liquidity issue with TIPS, but it's not that great. I think there's \$70, \$80, \$90 billion worth of TIPS in the market. You can do a trade of fairly decent size at narrow bid–ask spreads. My opinion of what's going on right now is that nominal bonds are seen as a hedge. I think there is fear of deflation in the market. And as in 1929, 1930, and 1931, investors were thinking that if the world markets, such as Japan, were going to be in a bad state, in a deflationary sense, holding nominal assets was the thing to do. So, as a result, the demand for nominal bonds is rising as a hedge against deflation, which will be bad for the economy and for real assets. The difference between TIPS and nominal bonds doesn't measure unbiased expected inflation; there's a negative risk premium in the picture. It is not what we think of as "there's inflation risk so nominal bonds should sell at a higher-than-expected return." I think right now the premium is a negative risk premium as investors use nominal bonds as a hedge against deflationary circumstances in the economy.

STEPHEN ROSS: In all of these computations of the equity risk premium on the stock market, does anyone take into account the leverage inherent in the stock market and the volatility premium that you would get from it? I don't have a clue about the empirical size of that premium. Can someone help me?

MARTIN LEIBOWITZ: I can. If you take the formulas that have been discussed today and translate them to assume a particular risk premium on unlevered assets, you can see how that premium translates into the typical level of leverage in the equity markets. You find that it is exactly what you'd expect. The risk premium that you actually see in the market reflects the leverage that is endemic in the equity market, and if you back out that premium to find the risk premium on unlevered assets, you find that the premium on unlevered assets is less.

RAJNISH MEHRA: The Sharpe ratio won't change. It's invariant to leverage.

LEIBOWITZ: It's exactly linear.

ROBERT SHILLER: Let's remember correctly the McGrattan and Prescott article (2001) that Brad Cornell mentioned. They use a representative agent model, and they compare the late 1950s and early 1960s with a recent year. And they say that because of 401(k)s and similar vehicles, the tax rate on dividends for a representative agent has fallen—from 50 percent in 1950–1962 to 9 percent in 1987–1999. That fall seems to me like an awfully big drop, and I question whether there could have been such a big drop for the representative investor. I wonder if anyone here has looked carefully at their model? Are they right?

SIEGEL: They use the average investor; they don't use the marginal investor. They say that X percent of assets are in a 401(k), and they equate that amount with the marginal rate. My major criticism of the McGrattan–Prescott paper is that we don't know whether the marginal investor is a taxable investor, which would change their results dramatically.

CORNELL: That criticism doesn't mean their results are wrong. We simply don't know.

SIEGEL: We don't know. But I have a feeling that the marginal investor has a much higher tax rate than the marginal investor used to have.

ROSS: Yes, James Poterba told me that his calculations indicate that 401(k)s have far less tax advantage at the margin than one might think. Because of the tax rate "upon withdrawals," those vehicles can be dramatically attacked from a tax perspective. If you make a simple presumption that 401(k)s are simply a way of avoiding taxes, you're missing the point.

THOMAS PHILIPS: I'd like to go back to the equation for expected future real returns that Jeremy Siegel attributes to me: Expected future real returns = Earnings yield + $g \times$ [1 – (Book value/Market value)]. It really is an expression for the expected future *nominal* return. When I derived that equation, I derived it in *nominal* terms. In particular, the growth term, g, is nominal, not real, growth (Philips 1999). When you subtract inflation, you have Expected future real returns = Earnings yield + Nominal growth $\times$ [(1 – Book value/Market value) – Inflation]; the last two terms go to approximately zero. You're left with the earnings yield being approximately the real expected return.

In the special case that Brad Cornell talked about, in which the cost of capital and the return on capital are the same, the second term disappears because the book-to-market ratio becomes 1. In that case, the earnings yield is actually the *nominal* expected return. The truth, in practice, lies somewhere in between the two results because some of these quantities will vary with inflation, real interest rates, and the economywide degree of leverage.

The approximation that Brad used is biased up or down depending on where inflation, growth, and the cost of capital relative to the return on capital lie. It's a great first-order approximation, a great historical approximation, but you can be talking about the nominal rate of return instead of the real rate of return when the cost of capital starts coming very close to the return on capital.

SIEGEL: Well, I disagree with you. In your slides, the earnings yield—if you're in equilibrium and book value equals market value equals replacement cost—is an estimate of the real return, not the nominal return. Your equation is extraordinarily useful, but I think we do have to interpret it as the real return.

ROGER IBBOTSON: I'd like to say something about Brad Cornell using aggregate calculations to get an estimate of the equity risk premium. I did some work on aggregate calculations in a paper I wrote with Jeffrey Diermeier and Laurence Siegel in 1984. Relating to merger and acquisition activities, we looked at how best to use cash: For example, do you use cash for dividends or share repurchases? (You could take the same approach for investing in projects.) When you look at which data to use in the context of cash mergers or acquisitions, you can see that the per share estimates are going to be very different from the aggregate estimates because you're buying other companies on a per share basis. Thus, EPS can grow much faster than aggregate corporate earnings.

CORNELL: That's why I like looking at aggregate earnings; it's the whole pot, and you're not as concerned about how things are moving around within the pot or being paid out to shareholders. But even looking at aggregate earnings, and this is based on Bob Shiller's data series going back to 1872, the earnings don't keep up with GNP, despite the greater volatility of earnings; even the arithmetic averages are less. Can you explain that phenomenon? What does it imply for the future?

SHILLER: The national income and product account (NIPA) earnings keep up a lot better. So, it's probably because earnings in the market indexes are not representing the new companies that come into the economy and existing companies' earnings are growing at a slower rate.

SIEGEL: I looked at it very closely. The trend in the ratio of NIPA profits to GDP is virtually zero, the mean being 6.7 percent. You can do a linear regression—any regression—and you get a trend of absolutely zero: The ratio of NIPA profits to GDP has remained constant. Aggregate S&P 500 Index profits have slipped because the S&P 500 back in the 1950s and 1960s represented a much higher percentage of the market's value than it does today. You can look at both aggregate S&P 500 profits and aggregate NIPA profits and see the trends.

MEHRA: I found the same thing in my 1998 paper. The ratio of aggregate cash flows to national income (NI) is essentially trendless. In the afternoon, I'll be talking about the difference when you look at stock market valuation relative to national income [see the "Current Estimates and Prospects for Change" session]. That ratio fluctuates from about 2 × NI to about 0.5 × NI, whereas cash flows, which are the input for all these valuation models, are trendless relative to NI.

KEVIN TERHAAR: I want to go back to the representative investor or the marginal investor and Brad Cornell's first "rational" reason that the market might be high—that stocks are seen as less risky. One thing that hasn't been brought up is that all the discussions so far have focused primarily on the U.S. equity market. To the extent that the marginal investor looks at U.S. equities in the context of a broader portfolio (as opposed to looking at them only in a segmented market), the price of risk (or the aggregate Sharpe ratio) can stay the same while the equity premium for *U.S.* equities can fall. As the behavior of investors becomes less segmented—as they become less apt to view assets in a narrow or isolated manner—the riskiness of the assets can decline. Risk becomes systematic rather than total, and as a result, the compensation for risk falls commensurately.

WILLIAM GOETZMANN: I have a related comment in reference to Brad Cornell's presentation. An interesting aspect was his reference to changes in diversification of individual investors. There's not much empirical evidence on this issue, but it's interesting because we did have a boom in mutual funds through the 1980s and 1990s, with investors becoming more diversified. And the result was that the volatility of their equity portfolios dropped. We saw a similar trend in the 1920s, at least in the United States, through much growth in the investment trusts.[1] We think of trusts as these terrible entities that we clamped down on in the 1930s, but nevertheless, they did provide diversification for individual investors. So, maybe there is some relationship between the average investor's level of diversification and valuation measures of the equity premium.

It's hard to squeeze much more information out of the time-series data because we don't have many booms like I just described. But we might get something from cross-sectional studies—looking internationally—because we have such differences in the potential for investors in each country to diversify—different costs associated with diversification and so forth. So, maybe we could find out something from international cross-sectional data.[2]

CAMPBELL: On the diversification issue, I have a couple of cautionary notes. First, I think that diversification on the part of individual investors probably is part of this story, but what matters for pricing ought not to be the diversification of investors with investors equally weighted but with investors *value weighted*. Presumably, the wealthy have always been far more diversified than the small investor. So, if small investors succeed in diversifying a bit more, it may not have much effect on the equity premium.

Second, you mentioned the trend toward increased diversification in recent years. There has also been a trend toward increased idiosyncratic risk in recent years. So, although marketwide volatility has not trended up, there has been a very powerful upward trend since the 1960s in the volatility of a typical, randomly selected stock. So, you *need* to be more diversified now in order to have the same level of idiosyncratic risk exposure as before 1960. It's not clear to me whether the increase in diversification of portfolios has outstripped that other trend or merely kept pace with it.

ROSS: It's not at all obvious to me that the wealthy are more diversified. The old results from estate tax data I found are really quite striking. Keep in mind that the data contain survivorship bias and that the rich got wealthy by owning a company that did well, but as I remember, the mean holding of the wealthy is about four stocks, which is really quite small. Conversely, if you look at the less wealthy investor, many of their assets are tied up in pension plans,

[1] Investment trusts existed solely to hold stock in other companies, which frequently held stock in yet other companies.
[2] For a discussion of long-term equity risk premiums in 16 countries, see Dimson, Marsh, and Staunton (2001).

where the diversification—even in defined-benefit plans—is subtle and not easy to detect. The same can be said for Social Security.

SIEGEL: I think we should also keep in mind the absolutely dramatic reduction in the cost of buying and selling stocks. Bid–ask spreads are sometimes pennies for substantial amounts of stocks, and transaction costs have decreased virtually to zero. I would think that, even with the increase in idiosyncratic risk, if individual investors *want* to diversify (leaving aside the question of whether they want to diversify or pick stocks), they can do so at a much lower cost today than they could, say, 20 or 30 years ago.

BANSAL: So, your argument for the falling equity premium would be that the costs have gone down more for equities than for bonds?

SIEGEL: Yes.

ASNESS: We still see many investors with tremendously undiversified portfolios. There are psychological biases and errors that can lead to a lack of diversification; we haven't had a rush to the Wilshire 5000 Total Market Index.

RICHARD THALER: To follow up, I want to point out that research on the prevalence of ownership of company stock in 401(k) plans indicates that it's quite high—in some companies, shockingly high. At Coca-Cola, for example, at one time, more than 90 percent of the pension assets were in Coca-Cola stock. The same pattern was common in the technology companies. Talk about investments being undiversified *and* positively correlated with human capital! These situations are very risky.

ASNESS: Have you ever tried to convince an endowment started by one family that what they should really do is diversify?

THALER: Right, right.

ASNESS: You never succeed.

THALER: Research on the founders of companies indicates that they hold portfolios with very low returns and very high idiosyncratic risk.

ASNESS: But they had *had* very high returns at some point.

THALER: Right.

PHILIPS: I'd like to re-explore the earnings versus GDP question. Rob Arnott and Peter Bernstein (2002) find that per share earnings grow more slowly than the economy for a very simple reason: A large chunk of the growth of the economy is derived from new enterprises, and therefore, the growth in earnings per dollar of capital will be inherently lower than the growth of earnings in the entire economy. Their empirical result is that per share earnings grow at roughly the same rate as per capita GDP. Let's call that the rate of growth of productivity. I, on the other hand, am much more comfortable with the notion of EPS growing at roughly the same rate as the economy as a whole. Why? Because the old economy spins off dividends that it cannot reinvest internally. Those dividends, in turn, can be invested in the new economy, which allows you to capture the growth in the new economy. In effect, you have a higher growth rate and a lower dividend yield, and your per share earnings keep growing at roughly the same rate as the economy as a whole. Do you have a take on that, Jeremy? Do you have an instinctive feel for whether we're missing something here or not?

SIEGEL: If companies paid out all their earnings as dividends (with no reinvestment or buying back of shares) and because (based on the long-run-growth literature) the capital output ratio is constant, then EPS would not grow at all. You would have new shares as the economy grew, through technology or population growth, because companies would have to float more shares over time to absorb new capital. But EPS wouldn't really grow at all. What happens, of course, is that the companies withhold some of their earnings for reinvestment or buyback of shares, which pushes EPS upward. If the earnings growth also happens to be the rate of productivity growth or GDP growth, I think it's coincidental, not intrinsic.

IBBOTSON: I have done work on the same subject, and I agree.

WILLIAM REICHENSTEIN: I have a concern. If you're buying back shares, EPS grow (corporate earnings don't necessarily grow, but earnings per share do). The argument that when companies reinvest their earnings rather than paying out their earnings to shareholders they must be wasting some of that money just doesn't jibe with the reality that the price-to-book ratio on the market today is about 4 to 1. If the market is willing to pay $4.00 for the $1.00 equity that is being reinvested, companies cannot be wasting the reinvested money.

SIEGEL: The confusing thing is that the price-to-book ratio for the S&P 500 or the DJIA is about 4 or 5 to 1 but the Tobin's *q*-ratio—which uses book value adjusted for inflation and replacement costs—is nowhere near that amount. I think it could be very misleading to use historical market-to-book ratios.

LEIBOWITZ: Still, whether you use the market-to-book ratio or not, the idea of having high P/Es in an environment where monies are reinvested at less than the cost of capital produces the same inconsistency. Something doesn't compute.

IBBOTSON: The burden is on the people who are challenging the Miller–Modigliani theorem. M&M said that dividends and retention of earnings have the same effect so which number is used doesn't matter; you're saying it does matter.

ARNOTT: I believe the Miller–Modigliani theorem is an elegant formula that should work. But it doesn't match 130 years' worth of historical data.

IBBOTSON: We'll investigate that!

PHILIPS: In part, the difference may be something already mentioned: NIPA (which covers all businesses) versus the set of publicly traded securities (which is a subset of NIPA). Examining both groups separately might provide us some answers to the reinvestment question. Another angle on reinvestment is: Suppose we idealize the world so that businesses reinvest only what they need for their growth (so, it's a rational reinvestment, not empire building). What is our view now of how EPS should be growing? Is there a consensus? Rob Arnott has some very strong numbers showing that per share earnings grow more slowly than the economy. Will you be putting up that graph this afternoon, Rob?

ARNOTT: Yes, that's why I'm not saying anything.

SIEGEL: What's interesting is that growth has occurred over time in the marketable value of securities versus what would be implied by the NIPA profits. Many more companies are now public than used to be. A lot of partnerships have gone public in the past 10–20 years. A lot of small companies, private companies, have gone public recently. Part of the reason could be the good stock market, and part could be a long-term trend. At any rate, in NIPA, a very big decline has occurred in "proprietors' income," which is derived from partnerships and individual owners, and an increase has occurred in corporate income as these private companies and partnerships went public. You have to be aware of this trend if you are using long-term data. It is one reason I think there is an upward trend in market value versus GDP. I'm not saying the ownership change alone explains the market value trend, or that it explains the whole amount, but changes between corporate income and noncorporate income are important.

IBBOTSON: So, as I've just said, either go to per share data to do this type of analysis or make sure you make all these adjustments to the aggregate data. See Diermeier, Ibbotson, and Siegel (1984) if you want to see how to make the adjustments.

TERHAAR: For the per share data, however, most people use the S&P 500, and the S&P 500 isn't really passive. It's a fairly actively managed index, particularly in recent years; the managers at Standard & Poor's have a habit of adding "hot" stocks, such as their July 2000 inclusion of JDS Uniphase. These substitutions have effects on the per share earnings and the growth rate that would not be present in a broader index or in the NIPA index.

SIEGEL: That's a very important point. Whenever the S&P 500 adds a company that has a higher P/E than the average company in the index, which has been very much the case in the past three years, the result is a dollar bias in the growth rate of earnings as the index is recomputed to make it continuous. My calculations show that the bias could be 1–2 percent a year in recent years as companies with extraordinarily high P/Es were added.

Current Estimates and Prospects for Change I

Robert J. Shiller
Yale University
New Haven, Connecticut

The equity premium puzzle and the foundations of behavioral finance are inseparable. The equity premium puzzle is a puzzle only if we assume that people's expectations are consistent with past historical averages, that expectations are rational. But behavioral finance has shown repeatedly the weakness of the assumption that rational expectations consistently drive financial markets. This presentation explores, in the context of recent stock market behavior, a number of reasons to doubt that rational expectations always find their way appropriately into stock prices. The reasons stressed have to do with psychological factors: (1) the difficulty that committees, groups, and bureaucracies have in changing direction, (2) the inordinate influence of the recent past on decisions, (3) the tendency (perhaps the need) to rely on "conventional wisdom," and (4) group pressure that keeps individuals from expressing dissent.

I will discuss here some issues in behavioral finance related to the so-called equity premium puzzle. The academic literature on the puzzle is based on the assumption that people are perfectly rational and consistent in their financial decision making and that their expectations for future returns are at all times in line with facts about past historical returns. The term "equity premium puzzle" refers to the fact that the performance of the stock market in the United States has just been too strong relative to other assets to make sense from the standpoint of such rationality. But behavioral finance research has provided strong evidence against the very assumptions of rationality, at least against the idea that the rationality is consistent and responsive to relevant information and only relevant information. The equity premium puzzle and the foundations of behavioral finance are inseparable.

People's expectations cannot be equated with mathematical expectations, as the equity premium literature assumes. Expectations for future economic variables, to the extent that people even have expectations, are determined in a psychological nexus. I want to describe, in the context of recent experience in the stock market, some of the psychology that plays a role in forming these expectations. Considering recent experience will help provide concreteness to our treatment of expectations. The U.S. equity market became increasingly overpriced through the 1990s, reaching a phenomenal degree of overpricing by early 2000.[1] This event is a good case study for examining expectations in general.

I will be following here some arguments I presented in my 2000 book *Irrational Exuberance*, and I will also develop some themes that I covered in my 2002 paper, "Bubbles, Human Judgment, and Expert Opinion," which concentrated attention on the behavior of institutional investors—particularly, college endowment funds and nonprofit organizations (see Shiller 2002).

The theme of "Bubbles, Human Judgment, and Expert Opinion" is that even committees of experts can be grossly biased when it comes to actions like those that are taken in financial markets.

A lot of behavioral finance depicts rather stupid things going on in the market, but (presumably) trustees and endowment managers are pretty intelligent people. Yet, they, as a group, have not been

[1] See the testimony by John Y. Campbell and Robert J. Shiller before the Federal Reserve Board on December 3, 1996. Summarized in Campbell and Shiller (1998).

betting against the market during this recent bubble. They seem to be going right along with it. One of the biggest arguments for market efficiency has been that if the market is inefficient, why are the smart people still investing in the market. So, the question of how expert opinion can be biased will be one of the focal points of this talk.

The Recent Market Bubble

Figure 1 is the Nasdaq Composite Index in real terms from October 1984 to October 2001. Anyone who is thinking about the equity premium puzzle ought to reflect on what an event like the recent bubble we have had implies about the models of human rationality that underlie the equity premium puzzle. There has never been a more beautiful picture of a speculative bubble and its burst than in the Figure 1 chart of the Nasdaq; the price increase appears to continue at an ever increasing rate until March 2000; then, there is a sudden and catastrophic break, and the index loses a great deal of its value. We will have to reflect on what could have driven such an event before we can be comfortable with the economic models that imply a high degree of investor consistency and rationality.

Figure 2 shows the same speculative bubble from 1999 to late 2000 in the monthly real price and earnings of the S&P Composite Index since 1871. This bubble is almost unique; the only other one like it for the S&P Composite occurred in the 1920s; we could perhaps add the period just before the mid-1970s as a similar event. So, because we have a record of only two (possibly three) such episodes in history, a lot of short-run historical analysis may be misleading. We are in very unusual times, and this circumstance is obvious when we look at Figure 2.

The bubble that was seen in the late 1990s was not entirely confined to the stock market. Real estate prices also went up rapidly then. Karl Case[2] and I have devised what we call the "Case–Shiller Home Price Indexes" for many cities in the United States. **Figure 3** is our Los Angeles index on a quarterly basis from the fourth quarter of 1975 to the second quarter of 2001. (The smoothness in price change is not an artifact; real estate price movements tend to be smooth through time. The real estate market is different from the stock market.) Figure 3 tells an interesting and amazingly simple story. The two recessions over the period—1981–1982 and 1990–1991—are easy to see. Los Angeles single-family home prices were trending up when the 1981–82 recession hit. Then, although nominal home prices did not go down, prices did drop in real terms. After that recession, prices moved up again, only to fall again in the 1990–91 recession. Following that recession, prices soared back up. In the fall of 2001, we are again entering a recession. So, our prediction is that home

[2] Of Wellesley College, Massachusetts, and the real estate research firm of Case Shiller Weiss, Inc.

Figure 1. Real Nasdaq Composite, October 1984–October 2001

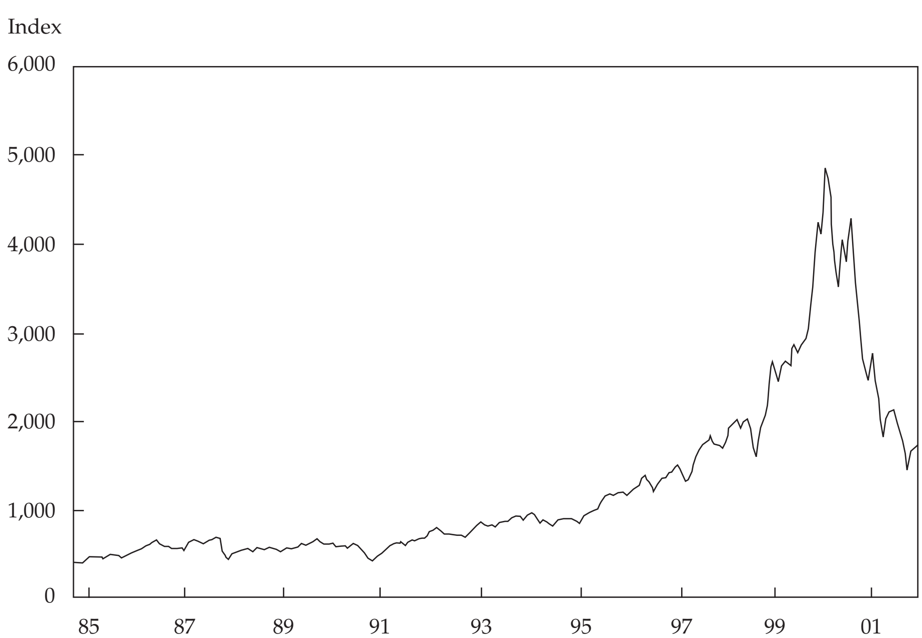
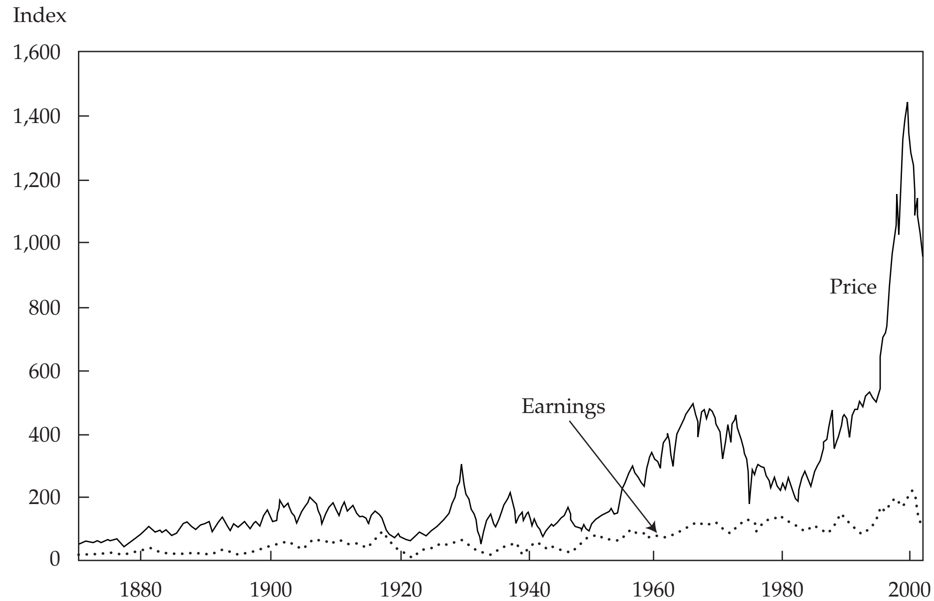
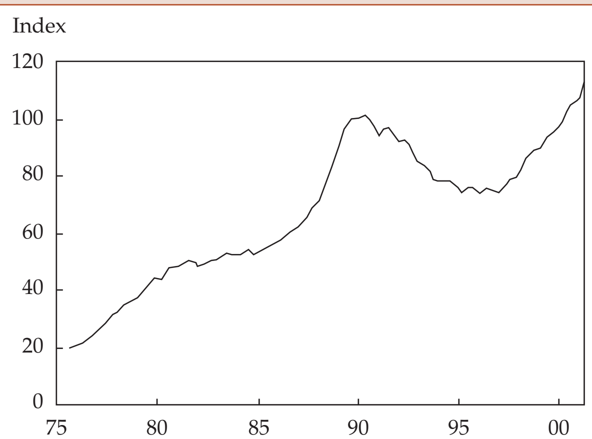


Figure 2. S&P Composite: Real Price and Earnings, January 1871–2001



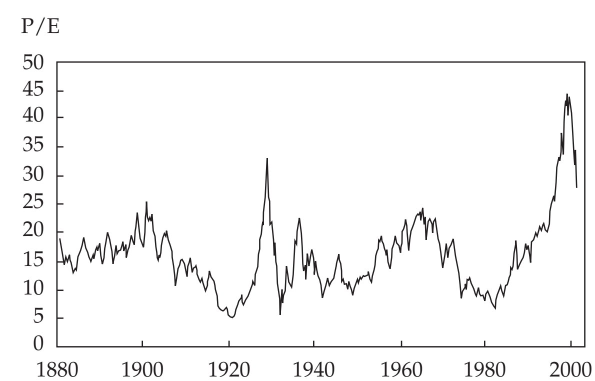
Note: Measured monthly.

Figure 3. Case–Shiller Home Price Index: Los Angeles Single-Family Home Prices, Fourth Quarter 1975 to Second Quarter 2001



Note: Measured quarterly.

Figure 4. P/E for the S&P Composite, January 1881–October 2001



Note: P/E calculated as price over 10-year lagging earnings (a calculation recommended by Graham and Dodd in 1934).

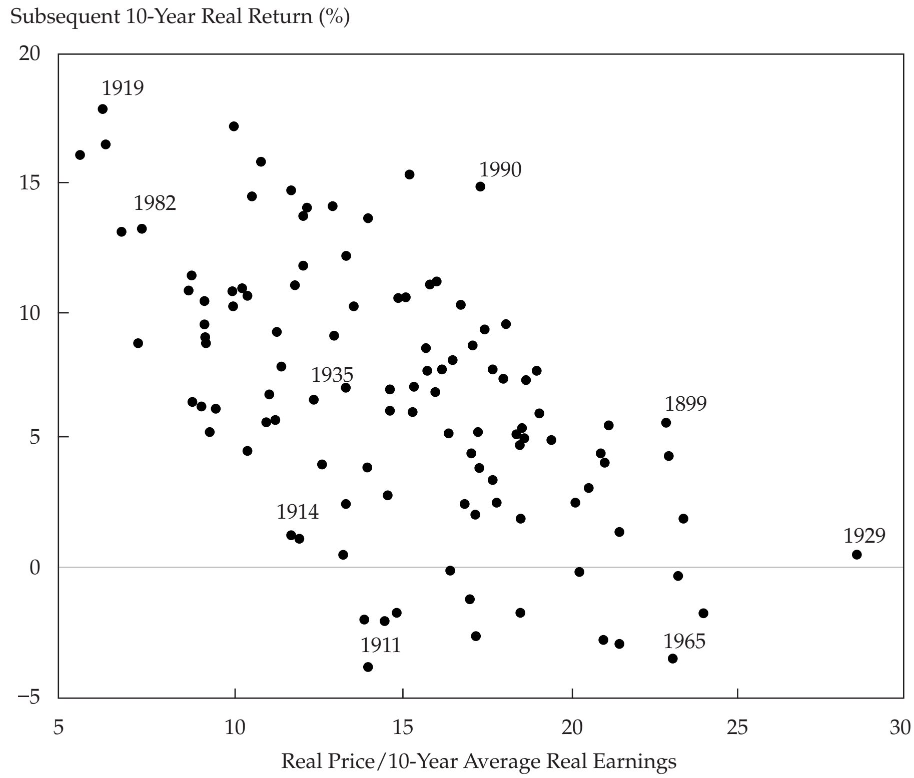
prices may trend lower as a result. We do not expect to see in the market for homes a sharp bubble and burst pattern such as we saw in the Nasdaq, but we might well see some substantial price declines.

Figure 4, the S&P Composite P/E for 1881 to 2001, shows once again the dramatic behavior in the stock market recently, behavior matched only by the market of the late 1920s and (to a lesser extent) around 1900 and the 1960s.

Figure 5 is a scatter diagram, which John Campbell and I devised, depicting the historical negative correlation between P/Es and subsequent 10-year returns. Figure 5 shows how the S&P Composite P/E predicts future S&P Composite returns. The P/E is now around the 1929 level, which suggests that high valuation is the dominant issue in judging the equity premium at this time.

It seems there is sufficient evidence in these markets, not only in their outward patterns but also in their correlation with each other and with other events, to feel pretty safe in concluding that we have seen a speculative bubble here. I know that there are

Figure 5. P/E for the S&P Composite in Relation to Subsequent 10-Year Real Composite Returns



Notes: P/E for 1881–1990; average real returns for 1891–2000. A similar scattergram was used in the Campbell–Shiller presentation to Congress in 1996 (see Campbell and Shiller 1998) .

some academics who still apparently believe that there are no such things as speculative bubbles.[3] But these academics are increasingly in the minority in the profession.

Why Speculative Bubbles?

In *Irrational Exuberance*, I begin by showing the historical data that I just reviewed with you. The question that I addressed in the book is why we have speculative bubbles. I take three behavioral approaches to answering the question. In the first part, I consider structural factors—precipitating factors and amplification mechanisms—that encourage people to buy more stocks. The second part deals with cultural factors, such as the news media and "new era" theories. The third part deals with psychological factors, which include overconfidence, anchoring, and attention anomalies.

I have not heard many of these factors mentioned at our meeting today. It is puzzling to me that economists rarely seem to express an appreciation of the news media as important transmitters of speculative bubbles and of the idea that we are in a new era. Every time a speculative bubble occurs, many people who work in the media churn out stories that we are in a new era. I documented this phenomenon in my book by looking at a number of different cases in which the stock markets in various countries rose over a brief period, and I was able to find in each of them a new era theory in the newspaper.

Expert Theories

"Bubbles, Human Judgment, and Expert Opinion" was written to be of interest to practitioners. The objective was to observe how investors react to a market bubble and then try to interpret that phenomenon.

During the book tour for *Irrational Exuberance* in 2000 and 2001, I was often speaking to investment professionals, and although I had the sense that many times I was engaging their interest, I often did not have the sense that I was really connecting with them.

[3] For example, Peter Garber, in his recent (2000) book *Famous First Bubbles: The Fundamentals of Early Manias*, argues that even the tulipmania in Holland in the 1600s was essentially rational. He concludes, "The wonderful tales from the tulipmania are catnip irresistible to those with a taste for crying bubble, even when the stories are obviously untrue" (p. 83).

In many cases, they were not a really receptive audience. There was a sense of momentum or inertia among many of these people. They appeared to be of two minds—the one of an interested book reader and the other of a more rigid committee member or bureaucrat. I wanted to talk about that type of behavior in the "Bubbles" paper.

Why would that behavior be happening? What evidence would help us understand it? The reason I set forth in the paper is that the market is like a supertanker that cannot make sudden changes in course: Even if people like me present a case that the market is overpriced and is going to fall and even if people like me convince investment professionals that the market outlook is not so good, the professionals will not really make substantive changes in their portfolios. They may well continue to hold the 55 percent of their portfolios in U.S. equities and 11 percent in non-U.S. equities. University portfolio managers and other institutional investors were not withdrawing from the market in 1999.

In the paper, I discuss the *feedback* theory of bubbles that Andrei Shleifer and Nicholas Barberis (2000), I (1990), and others have talked about. In the feedback theory, demand for shares is modeled as a distributed lag of past returns plus the effect of precipitating factors. When returns have been high for a while, investors become more optimistic and bid up share prices, which amplifies the effects of precipitating factors. I consider this behavior to be an inconstancy in judgment, not naive extrapolation; for portfolio managers to respond naively to past returns seems implausible. Inconstancy in judgments arises because committees and their members find it difficult to respond accurately and incrementally to evidence, especially when the evidence is ambiguous, qualitative rather than quantitative, and ill defined. Ultimately, recent past returns have an impact on the decisions committee members make, even if they never change their conscious calculations. This feedback behavior thus amplifies the effect on the market of any precipitating factors that might initiate a speculative bubble.

The critical point is that the problem faced by institutional investors in deciding how much to put in the stock market is extremely complex; it has an infinite number of aspects that cannot possibly be completely analyzed. In such situations, people may fall into a pattern of behavior given by the "representative heuristic"—a psychological principle described by Kahneman and Tversky (1974, 1979) in which people tend to make decisions or judge information based on familiar patterns, preconceived categories or stereotypes of a situation. We tend to not take an objective outlook but to observe the similarity of a current pattern to a familiar, salient image in our minds and assume that the future will be like that familiar pattern.

Part of the problem that institutional investors face is the impossibility of processing all the available information. Ultimately, the decision whether to invest heavily in the stock market is a question of historical judgment. There are so many pieces of information that no one person can process all of them.

Therefore, institutional investment managers must rely on "conventional wisdom." They make decisions based on what they perceive is the generally accepted expert opinion. A problem with that approach is that one cannot know how much information others had in reaching the judgments laid out in conventional wisdom. In addition, investors do not know whether others were even relying on information or were, for their part, just using their judgment.

These kinds of errors that professionals make are analogous to the errors we sometimes make when, for example, we walk out of a conference and cross the street as a group. We may be talking about something interesting, so each person in the group assumes that someone else is looking at oncoming traffic. Sometimes, nobody is.

The tendency to follow conventional wisdom is increased by the strange standard we have called "the prudent person rule," part of fiduciary responsibility that is even written into ERISA. It is a strange standard because what it's really saying is not clear. As set forth in the ERISA regulations adopted in 1974, the prudent person rule states that investments must be made with

> the care, skill, and diligence, under the circumstances then prevailing, that a prudent man acting in a like capacity and familiar with such matters would use in the conduct of an enterprise with like character and like aims.

I interpret the statement to mean that an investment manager or plan sponsor must make judgments based on what is considered conventional at the time, not independent judgments.

The prudent person rule is a delicate attempt to legislate against stupidity, but the way the problem is addressed basically instructs the trustee or sponsor to be conventional. "Conventional" is exactly how I would describe what I think has happened to institutional investors and the way they approach the market. In 2000, many institutional investors believed they should not be so exposed to the market, but they could not justify to their organizations, within the confines of the prudent person rule, cutting back equity exposure. This dilemma is a serious problem.

Another problem that managers of institutional investments have can be described as "groupthink," a term coined in a wonderful book of the same name by the psychologist Irving Janis (1982). In the book, Janis gives case studies of committees or groups of highly intelligent people making big mistakes. In particular, he discusses the mistakes that arise because of group pressures individuals feel to conform. Janis points out that people who participate in erroneous decisions often find themselves censoring their statements because they believe, "If I express my dissenting view too often, I will be marginalized in the group and I will not be important." He uses the term "effectiveness trap" to describe this thinking. Dissenters, although they may be correct in their opinions, fear that they are likely to see their influence reduced if they express their opinions. Janis describes, for example, responses in the Lyndon Johnson administration to a Vietnam bombing fiasco. When Johnson wrote about this episode in his memoirs, he did not mention any substantial dissent. Yet, those involved remember having dissenting views. Evidently, they did not express their views in such a way that Johnson remembered the dissent after the fact.

As economists, we talk a great deal about models, which concretize the factors in decisions, but when you are making a judgment about how to manage a portfolio, you face real-world situations. The real world is fundamentally uncertain. And fundamental uncertainty is what Knight talks about in *Risk, Uncertainty and Profit* (1964): How do we react in committees or as groups or as individuals within groups?

An argument Shafir, Simonson, and Tversky (2000) recently made that they applied to individual decisions is, I think, even more applicable to group decisions. The authors stated that when we are making what seems like a portentous decision, our minds seek a *personalized* way to justify the decision; we do not simply consider what to do. They asked people to make hypothetical custody decisions about divorcing couples. They described the two parents and then asked each participant to choose which parent would get custody of the child. They framed the question in two different ways. One question was, "Which parent would you give the child to?" And the other was, "Which parent would you deny custody to?" Of course, the question is the same either way it is framed. Nevertheless, the authors found systematic differences in the responses. When the parents were described, one person was described in bland terms and the other person in very vivid terms—both good extremes and bad extremes. Participants tended to point their decisions to the more salient person (the more vividly described person) in the couple. For example, when the question was framed for awarding custody, participants tended to award custody to the person who was vividly described—even though the description included bad things. And when the question was framed for denying custody, participants tended to deny custody to the person who was vividly described—even though the description included good things.

This research points to a fundamental reason for inertia in organizations: Institutions have to have a very good reason to change any long-standing policy, but the kinds of arguments that would provide that good reason are too complicated (not salient enough) to be persuasive.

Conclusion

My talk has taken us a little bit away from the abstract issue of the long-run equity premium that has been talked about so much at this forum. I have described a shorter-run phenomenon, the recent stock market bubble, and I have described some particular psychological principles that must be borne in mind if we are to understand this recent behavior. But we cannot see the weaknesses of faulty abstract principles unless we focus on particular applications of the principles. I hope that my discussion today has raised issues relevant to understanding whether we ought to consider the markets efficient, whether we ought to be "puzzled" by the past equity premium, and whether we should expect this historical premium to continue in the future.

Current Estimates and Prospects for Change I

Robert J. Shiller

Yale University
New Haven, Connecticut

SUMMARY

by Peter Williamson

Amos Tuck School of Business Administration
Dartmouth College, Hanover, New Hampshire

Robert Shiller described the equity premium puzzle as inseparable from the foundations of behavioral finance. The three bases of his presentation were

- Campbell and Shiller, testimony before the Federal Reserve Board on December 3, 1996,[1]
- *Irrational Exuberance* (published in April 2000; see Shiller 2000), and
- "Bubbles, Human Judgment, and Expert Opinion" (Shiller 2002).

[1] Summarized in Campbell and Shiller (1998).

The third publication was aimed at (nonprofit) practitioners (particularly, those at U.S. educational endowments). Much behavioral finance describes apparently foolish behavior in the market, but trustees are, presumably, intelligent people. Yet, even they have not been betting against the market during the recent bubble. Despite warnings, intelligent people have not lost faith in the stock market. Why is expert opinion so biased?

Shiller's **Figure 1** showed the real Nasdaq Composite Index from October 1984 to October 2001. It provided clear evidence of a perfect bubble from 1999 to late 2000. The same could be seen in his **Figure 2** of the S&P Composite Index from 1871 to 2001. Two other, lesser bubbles appeared—in the late 1920s and the late 1960s. Similarly, the **Figure 3** graph of real estate prices in Los Angeles, California, showed a clear bubble (although it was smoother than the market bubble) around 1990. **Figure 4**, of the S&P

Figure 1. Real Nasdaq Composite, October 1984–October 2001

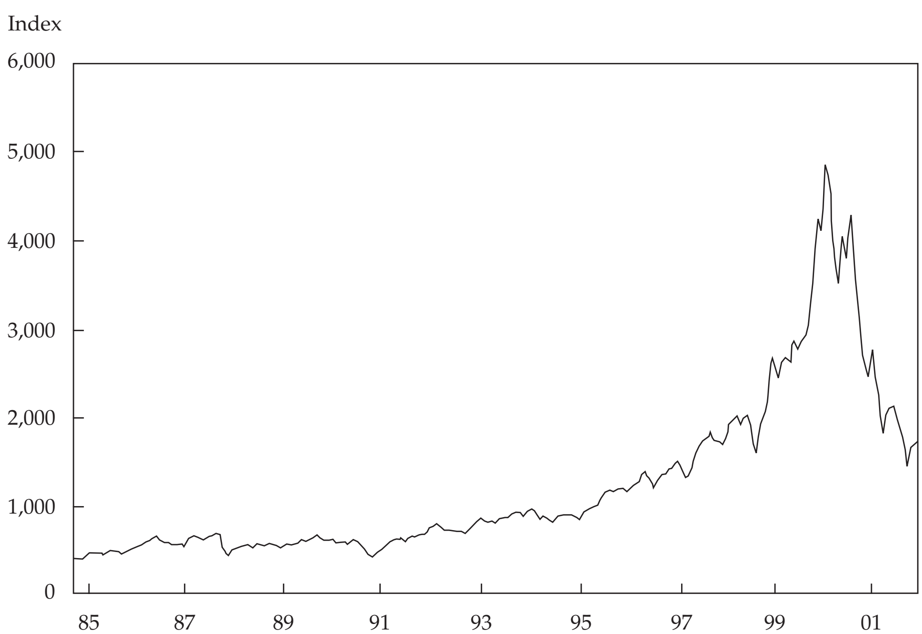
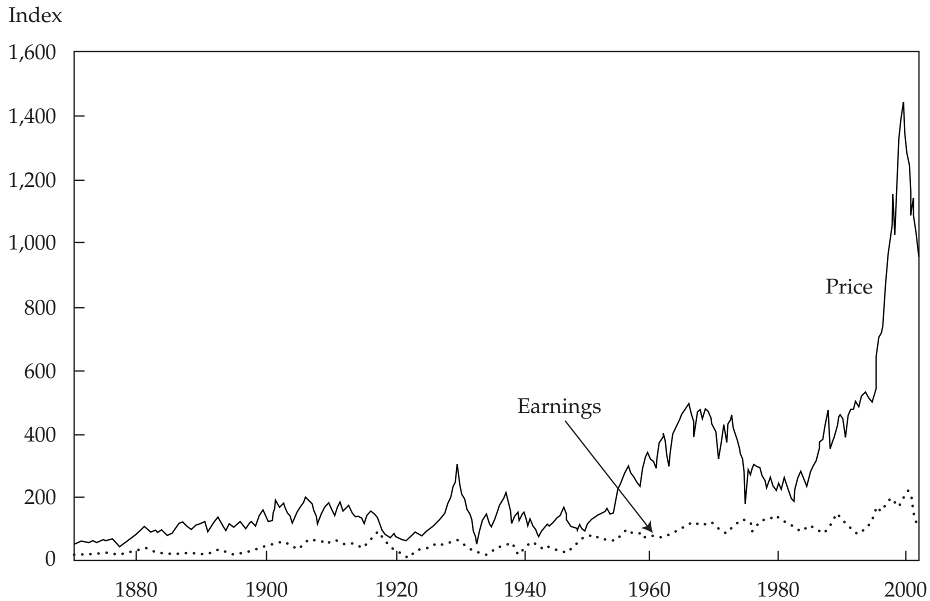
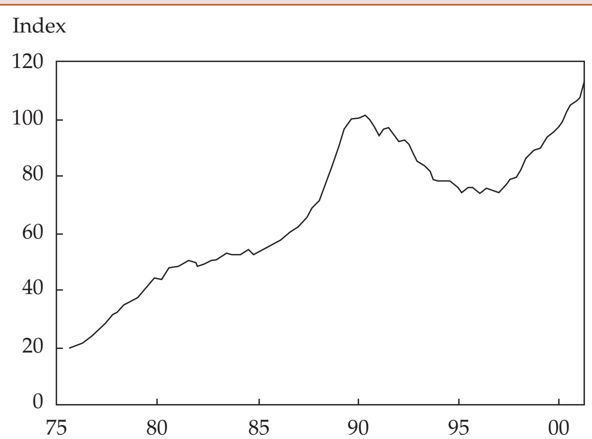


Figure 2. S&P Composite: Real Price and Earnings, January 1871–2001



Note: Measured monthly.

Figure 3. Case–Shiller Home Price Index: Los Angeles Single-Family Home Prices, Fourth Quarter 1975 to Second Quarter 2001

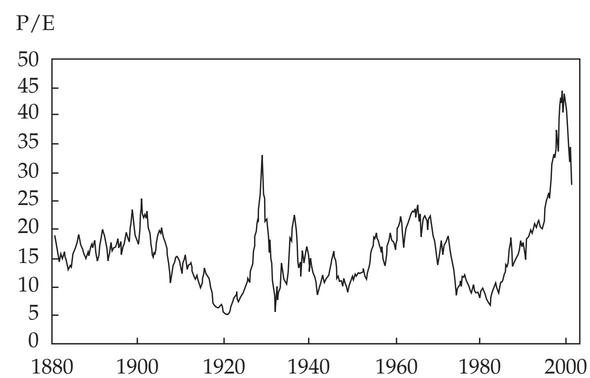


Note: Measured quarterly.

Composite P/E (real price divided by average real earnings over the preceding 10 years) from 1881 to 2001, showed bubbles recently, in the late 1920s, around 1900 (to a lesser extent), in the late 1930s, and in the 1960s.

Figure 5 is a scattergram showing how the S&P Composite P/E predicts future S&P Composite returns. The P/E is now around the 1929 level, which suggests that valuation is the dominant issue in terms of the equity premium at this time.

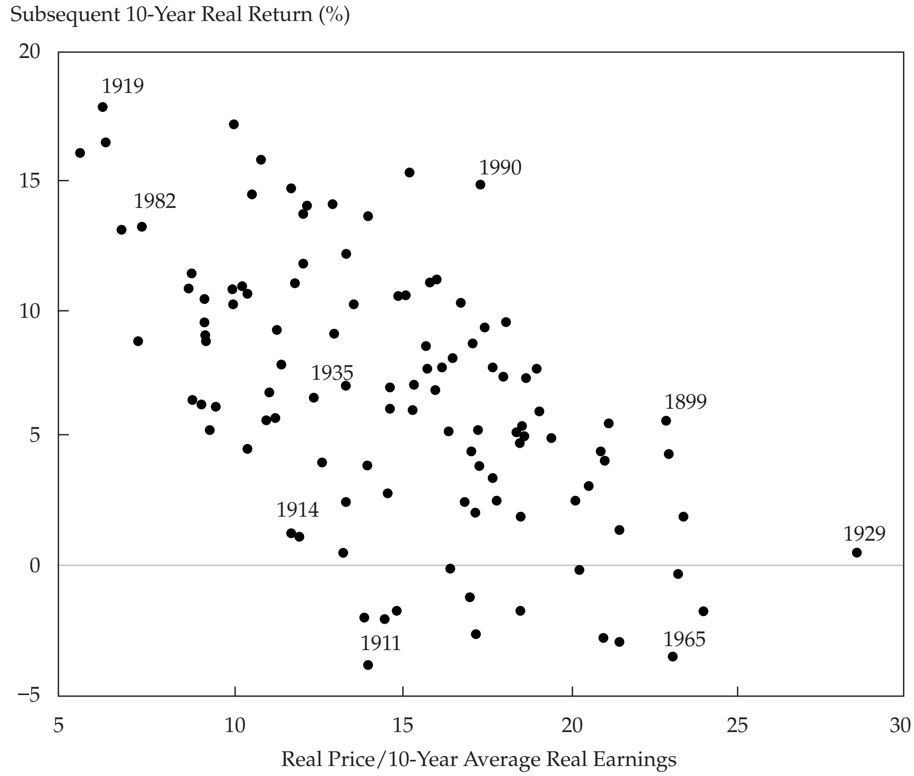
Figure 4. P/E for the S&P Composite, January 1881–October 2001



Note: P/E calculated as price over 10-year lagging earnings (a calculation recommended by Graham and Dodd in 1934).

In his book *Irrational Exuberance*, Shiller dealt with three types of factors leading to excessive valuations: structural, cultural, and psychological. Cultural factors included the news media and "new era" theories. The news media are important transmitters of speculative bubbles, and every bubble is accompanied by a new era theory to explain the rise in prices. Among psychological factors are overconfidence, anchoring, and attention anomalies.

Figure 5. P/E for the S&P Composite in Relation to Subsequent 10-Year Real Composite Returns



Notes: P/E for 1881–1990; average real returns for 1891–2000. A similar scattergram was used in the Campbell–Shiller presentation to Congress in 1996 (see Campbell and Shiller 1998) .

Turning to the subject of his "Bubbles" paper, Shiller discussed a number of aspects of behavioral finance behind the behavior of investment professionals that drove equity prices up. The most important factor is the inertia of a bureaucratic process. No matter how convincing the evidence that stock prices are too high, institutional committees do not change their asset allocations, which were generally about 60 percent in U.S. and non-U.S. equities in 1999.

The influence of recent past returns is powerful. Reliance on recent returns might be thought of as naive extrapolation, but Shiller prefers to think of it as inconstancy in judgment. It is difficult for committees to maintain the same judgment at all times when the evidence is ambiguous and complicated. The tendency is to assume that the future will be like the past.

The impossibility of processing all available information leads to reliance on conventional wisdom. Institutional investors have a tendency to trust the opinions of others without knowing what information those others are making use of. Moreover, the "prudent person rule" is, unfortunately, to "do what is conventional."

Shiller also cited examples of the "effectiveness trap"—the group pressure to conform—described in *Groupthink* (Janis 1982). Dissenters, although they may be correct in their opinions, fear that they are likely to see their influence reduced if they express their opinions. Other references Shiller made dealt with the difficulty of getting organizations to change long-standing policy. Committees need a *very* good reason to change a policy.

Shiller's conclusions included the following:

- Bubble behavior and the equity risk premium are tied up with many issues of human cognition and judgment.
- Institutional investors have generally been too slow to react to the negative equity premium today.

Current Estimates and Prospects for Change II

Rajnish Mehra

Professor of Finance
University of California, Santa Barbara
National Bureau of Economic Research and Vega Asset Management

Analysts have more than 100 years of good, clean economic data on asset returns that support the persistence of a historical long-term U.S. equity risk premium over U.S. T-bills of 5–7 percent (500–700 bps)—but the expected equity risk premium an analyst might have forecasted at the beginning of this long period was about 2 percent. The puzzle is that stocks are not so much riskier than T-bills that a 5–7 percent difference in rates of return is justified. Analyses of the long series of data indicate that the relationship between *ex ante* and *ex post* premiums is inverse. The relationship between the market and the risk premium is also inverse: When the value of the market has been high, the mean equity risk premium has been low, and vice versa. Finally, investors and advisors need to realize that all conclusions about the equity risk premium are based on and apply only to the very long term. To predict next year's premium is as impossible as predicting next year's stock returns.

I took the topic of the equity risk premium literally and considered, given current valuation levels, what is the expected equity risk premium. I would argue that this question is an exercise in forecasting and has little to do with the academic debate on whether the historically observed equity risk premium has been a puzzle. Let me illustrate.

Table 1 shows the data available to us from various sources and research papers on U.S. equity returns (generally proxied by a broad-based stock index), returns to a relatively riskless security (typically a U.S. Treasury instrument), and the equity risk premium for various time periods since 1802. The equity premium can be different over the same time period, primarily because some researchers measure the premium relative to U.S. T-bonds and some measure it relative to T-bills. The original Mehra–Prescott paper (1985) measured the premium relative to T-bills. Capital comes in a continuum of risk types, but aggregate capital stock in the United States will give you a return of about 4 percent. If you combine the least risky part and the riskier part, such as stocks, their returns will be different but will average about 4 percent. I can, at any time, pry off a very risky slice of the capital risk continuum and compare its rate of return with another slice of the capital risk continuum that is not at all risky.

Table 1 provides results from a fairly long series of data—almost 200 years—and the premium exists even when the bull market between 1982 and 2000 is

Table 1. Real U.S. Equity Market and Riskless Security Returns and Equity Risk Premium, 1802–2000

Period	Mean Real Return on Market Index	Mean Real Return on Relatively Riskless Asset	Risk Premium
1802–1998	7.0%	2.9%	4.1%
1889–2000	7.9	1.0	6.9
1889–1978	7.0[a]	0.8	6.2[b]
1926–2000	8.7	0.7	8.0
1947–2000	8.4	0.6	7.8

[a]Not rounded, 6.98 percent.
[b]Not rounded, 6.18 percent.

Sources: Data for 1802–1998 are from Siegel (1998); for 1889–2000, from Mehra and Prescott (1985).

excluded. That bull market certainly contributed to the premium, but the premium is pretty much the same in all the periods. One comment on early-19th-century data: The reason Edward Prescott and I began at 1889 in our original study is that the earlier data are fairly unreliable. The distinction between debt and equity prior to 1889 is fuzzy. What was in a basket of stocks at that time? Would bonds actually be called risk free? Because the distinction between these types of capital was unclear, the equity premium for the 1802–1998 period appears to be lower in Table 1 than I believe it really was. As **Table 2** shows, the existence of an equity premium is consistent across developed countries—at least for the post-World War II period.

The puzzle is that, adjusted for inflation, the average annual return in the U.S. stock market over 110 years (1889–2000) has been a healthy 7.9 percent, compared with the 1 percent return on a relatively riskless security. Thus, the equity premium over that time period was a substantial 6.2 percent (620 basis points). One could dismiss this result as a statistical artifact, but those data are as good an economic time series as we have. And if we assume some stationarity in the world, we should take seriously numbers that show consistency for 110 years. If such results occurred only for a couple of years, that would be a different story.

Is the Premium for Bearing Risk?

This puzzle defies easy explanation in standard asset-pricing models. Why have stocks been such an attractive investment relative to bonds? Why has the rate of return on stocks been higher than on relatively risk-free assets? One intuitive answer is that because stocks are "riskier" than bonds, investors require a larger premium for bearing this additional risk; and indeed, the standard deviation of the returns to stocks (about 20 percent a year historically) is larger than that of the returns to T-bills (about 4 percent a year). So, obviously, stocks are considerably more risky than bills!

But are they?

Why do different assets yield different rates of return? Why would you expect stocks to give you a higher return? The deus ex machina of this theory is that assets are priced such that, *ex ante*, the loss in marginal utility incurred by sacrificing current consumption and buying an asset at a certain price is equal to the expected gain in marginal utility contingent on the anticipated increase in consumption when the asset pays off in the future.

The operative emphasis here is the *incremental loss or gain* of well-being resulting from consumption, which should be differentiated from incremental consumption because the same amount of consumption may result in different degrees of well-being at different times. (A five-course dinner after a heavy lunch yields considerably less satisfaction than a similar dinner when one is hungry!)

As a consequence, assets that pay off when times are good and consumption levels are high—that is, when the incremental value of additional consumption is low—are less desirable than those that pay off an equivalent amount when times are bad and additional consumption is both desirable and more highly valued.

Let me illustrate this principle in the context of a popular standard paradigm, the capital asset pricing model (CAPM). This model postulates a linear relationship between an asset's "beta" (a measure of systematic risk) and expected return. Thus, high-beta stocks yield a high expected rate of return. The reason is that in the CAPM, good times and bad times are captured by the return on the market. The performance of the market as captured by a broad-based index acts as a surrogate indicator for the relevant state of the economy. A high-beta security tends to pay off more when the market return is high, that is, when times are good and consumption is plentiful; as

Table 2. Real Equity and Riskless Security Returns and Equity Risk Premium: Selected Developed Markets, 1947–98

Country	Period	Mean Real Return on Market Index	Mean Real Return on Relatively Riskless Asset	Risk Premium
United Kingdom	1947–1999	5.7%	1.1%	4.6%
Japan	1970–1999	4.7	1.4	3.3
Germany	1978–1997	9.8	3.2	6.6
France	1973–1998	9.0	2.7	6.3

Sources: Data for the United Kingdom are from Siegel (1998); the remaining data are from Campbell (2002).

discussed earlier, such a security provides less incremental utility than a security that pays off when consumption is low, is less valuable to investors, and consequently, sells for less. Thus, assets that pay off in states of low marginal utility will sell for a lower price than similar assets that pay off in states of high marginal utility. Because rates of return are inversely proportional to asset prices, the latter class of assets will, on average, give a lower rate of return than the former.

Another perspective on asset pricing emphasizes that economic agents prefer to smooth patterns of consumption over time. Assets that pay off a relatively larger amount at times when consumption is already high "destabilize" these patterns of consumption, whereas assets that pay off when consumption levels are low "smooth" out consumption. Naturally, the latter are more valuable and thus require a lower rate of return to induce investors to hold them. (Insurance policies are a classic example of assets that smooth consumption. Individuals willingly purchase and hold them in spite of their very low rates of return.)

To return to the original question: Are stocks that much riskier than bills so as to justify a 7 percent differential in their rates of return?

What came as a surprise to many economists and researchers in finance was the conclusion of a research paper that Prescott and I wrote in 1979. Stocks and bonds pay off in approximately the same states of nature or economic scenarios; hence, as argued earlier, they should command approximately the same rate of return. In fact, using standard theory to estimate risk-adjusted returns, we found that stocks on average should command, at most, a 1 percent return premium over bills. Because for as long as we had reliable data (about 100 years), the mean premium on stocks over bills was considerably and consistently higher, we realized that we had a puzzle on our hands. It took us six more years to convince a skeptical profession and for our paper (the Mehra and Prescott 1985 paper) to be published.

Ex Post versus *Ex Ante*

Some academicians and professionals hold the view that at present, there is no equity premium and, by implication, no equity premium puzzle. To address these claims, we need to differentiate between two interpretations of the term "equity premium." One interpretation is the *ex post* or realized equity premium over long periods of time. It is the actual, historically observed difference between the return on the market, as captured by a stock index, and the risk-free rate, as proxied by the return on T-bills.

The other definition of the equity premium is the *ex ante* equity premium—a forward-looking measure. It is the equity premium that is *expected* to prevail in the future or the conditional equity premium given the current state of the economy. I would argue that it *must* be positive because all stocks must be held.

The relationship between *ex ante* and *ex post* premiums is inverse. After a bull market, when stock valuations are exceedingly high, the *ex ante* premium is likely to be low, and this is precisely the time when the *ex post* premium is likely to be high. After a major downward correction, the *ex ante* (expected) premium is likely to be high and the realized premium will be low. This relationship should not come as a surprise because returns to stock have been documented to be mean reverting. Over the long term, the high and low premiums will average out.

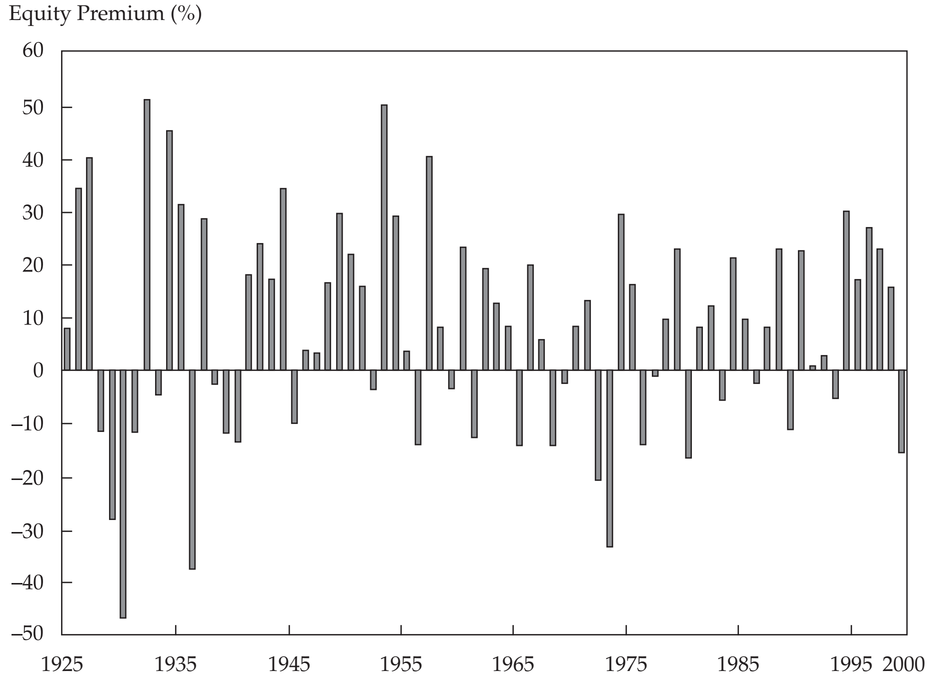
Which of these interpretations of the equity risk premium is relevant for an investment advisor? Clearly, the answer depends on the planning horizon.

The historical equity premium that Prescott and I addressed in 1985 is the premium for very long investment horizons, 50–100 years. And it has little—in fact, nothing—to do with what the premium is going to be over the next couple of years. Nobody can tell you that you are going to get a 7 percent or 3 percent or 0 percent premium next year.

The *ex post* equity premium is the realization of a stochastic process over a certain period, and as **Figure 1** shows, it has varied considerably over time. Furthermore, the variation depends on the time horizon over which it is measured. Over this 1926–2000 period, the realized equity risk premium has been positive and it has been negative; in fact, it has bounced all over the place. What else would you expect from a stochastic process in which the mean is 6 percent and the standard deviation is 20 percent? Now, note the pattern for 20-year holding periods in **Figure 2**. This pattern is more in tune with what Jeremy Siegel was talking about [see the "Historical Results" session]. You can see that over 20-year holding periods, there is a nice, decent premium.

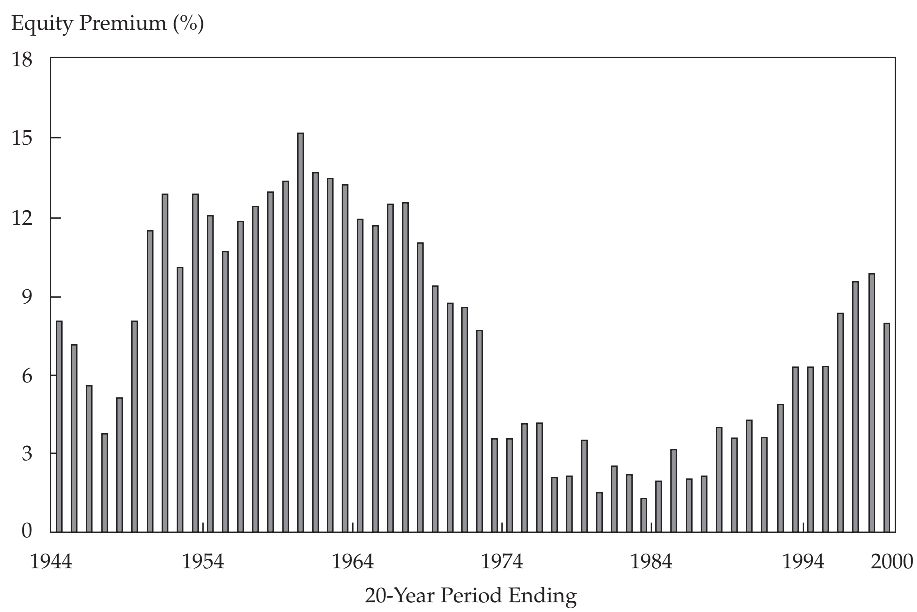
Figure 3 carries out exactly the exercise that Brad Cornell recommended [see the "Historical Results" session]: It looks at stock market value (MV)—that is, the value of all the equity in the United States—as a share of National Income (NI). These series are cointegrated, so when you divide one by the other, you get a stationary process. The ratio has been as high as approximately 2 times NI and as low as approximately 0.5 NI. The graph in Figure 3 represents risk. If you are looking for stock market risk, you are staring at it right here in Figure 3. This risk is low-frequency, persistent risk, not the year-to-year volatility in the market. This persistence defies easy

Figure 1. Realized Equity Risk Premium per Year, January 1926–January 2000



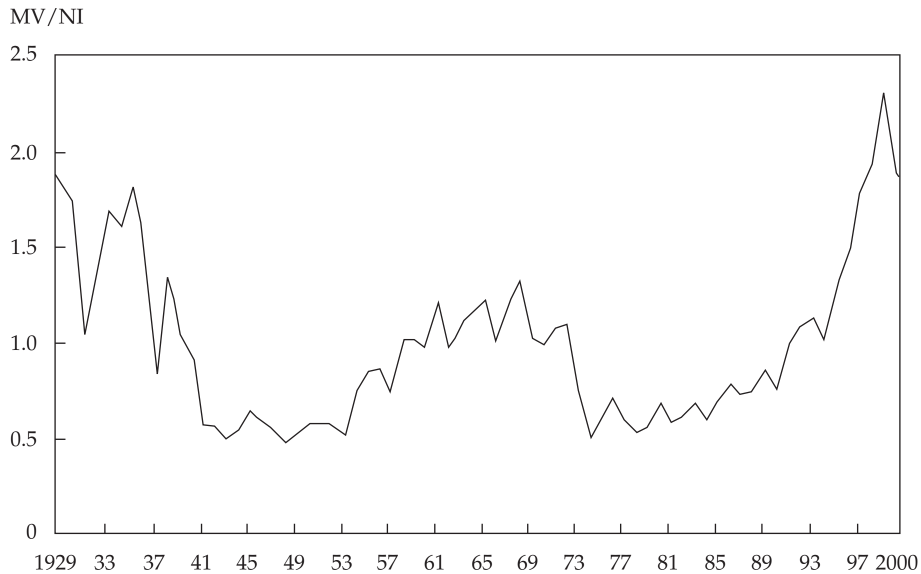
Source: Ibbotson Associates (2001).

Figure 2. Mean Equity Risk Premium by 20-Year Holding Periods, January 1926–January 2000



Source: Ibbotson Associates (2001).

Figure 3. U.S. Stock Market Value/National Income, January 1929–January 2000

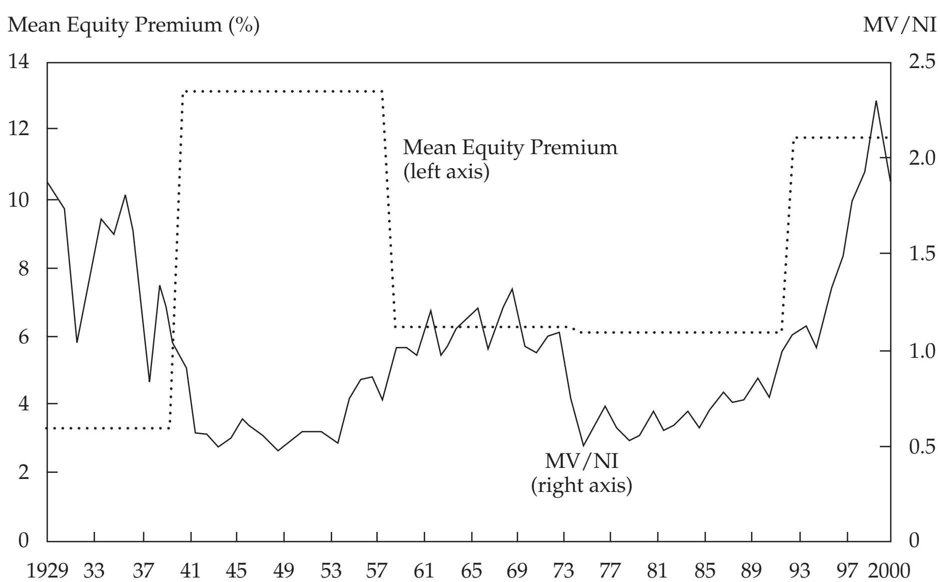


Source: Data updated from Mehra (1998).

explanation for the simple reason that if you look at cash flows over the same period of time relative to GDP, they are almost trendless. There are periods of relative overvaluation and periods of undervaluation, and they seem to persist over time.

When I plotted the contemporaneous equity risk premium over the same period, the graph I got was not very informative, so I arbitrarily broke up the data into periods when the market was more than 1 NI and when the market was below 1 NI. I averaged out all the wiggles in the equity premium graph, and **Figure 4** shows the smoothed line overlaid on the graph from Figure 3 of MV/NI. As you can see, when the market was high, the mean equity risk premium was low, and when the market was low, the premium was high.

Figure 4. Mean Equity Risk Premium and Market Value/National Income, January 1929–January 2000



The mean equity risk premium three years ahead is overlaid on the graph of market value to net income in **Figure 5**. (The premium corresponding to 1929 on the dotted line represents the mean equity risk premium averaged from 1929 to 1932. So, the premium line ends three years before 2001). You can clearly see that the mean equity risk premium is much higher when valuation levels are low.

I might add that the MV/NI graph is the basis of most of the work in finance on predicting returns based on price-to-dividends ratios and price-to-earnings ratios. Essentially, we have historical data for only about two cycles. Yet, a huge amount of research and literature is based on regressions run with only these data.

A scatter diagram of MV/NI versus the mean three-year-ahead equity risk premium is shown in **Figure 6**. Not much predictability exists, but the relationship is negative. (The graphs and scatter diagrams for a similar approach but with the equity risk premium five years ahead are similar).

Finally, **Figure 7** plots mean MV/NI versus the mean equity risk premium three years ahead, but I arbitrarily divided the time into periods when MV/NI was greater than 1 and periods when it was less than 1, and I averaged the premium over the periods. This approach shows, on average, some predictability: Returns are higher when markets are low relative to GDP. But if I try to predict the equity premium over a year, for example, the noise dominates the drift.

Operationally, because the volatility of market returns is 20 percent, you do not get much information from knowing that the mean equity premium is 2 percent rather than 6 percent. From an asset-allocation point of view, I doubt that such knowledge would make any difference over a short time horizon—the next one or two years. The only approach that makes sense in this type of analysis is to estimate the equity premium over the very long horizon. The problem of predicting the premium in the short run is as difficult as predicting equity returns in the short run. Even if the conditional equity premium given current market conditions is small (and the general consensus is that it is), that fact, in itself, does not imply either that the historical premium was too high or that the unconditional equity premium has diminished.

Looking into the Future

If this analysis had been done in 1928, what would an exercise similar to what Prescott and I did in 1985 have yielded? Suppose the analysis were done for the period from 1889 to 1928; in 1929, the mean real return on the S&P 500 was 8.52 percent, the mean real return on risk-free assets was 2.77 percent, and thus the observed mean equity premium would have been 5.75 percent. A theoretical analysis similar to Prescott's and mine would have yielded a 2 percent equity premium.

Figure 5. Mean Equity Risk Premium Three Years Ahead and Market Value/National Income, January 1929–January 2000

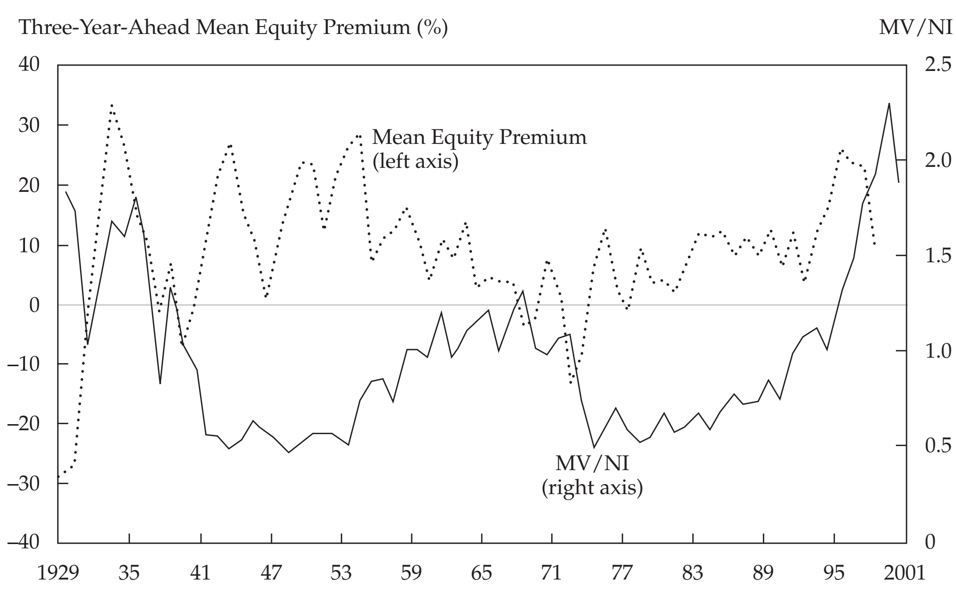
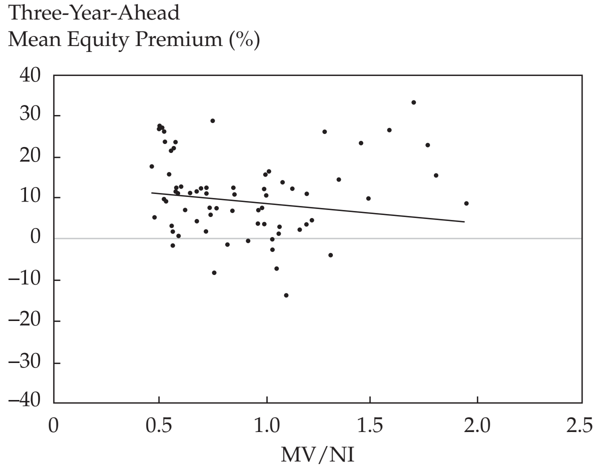


Figure 6. Scatter Diagram: Mean Equity Risk Premium Three Years Ahead versus Market Value/National Income, January 1929–January 2000 Data

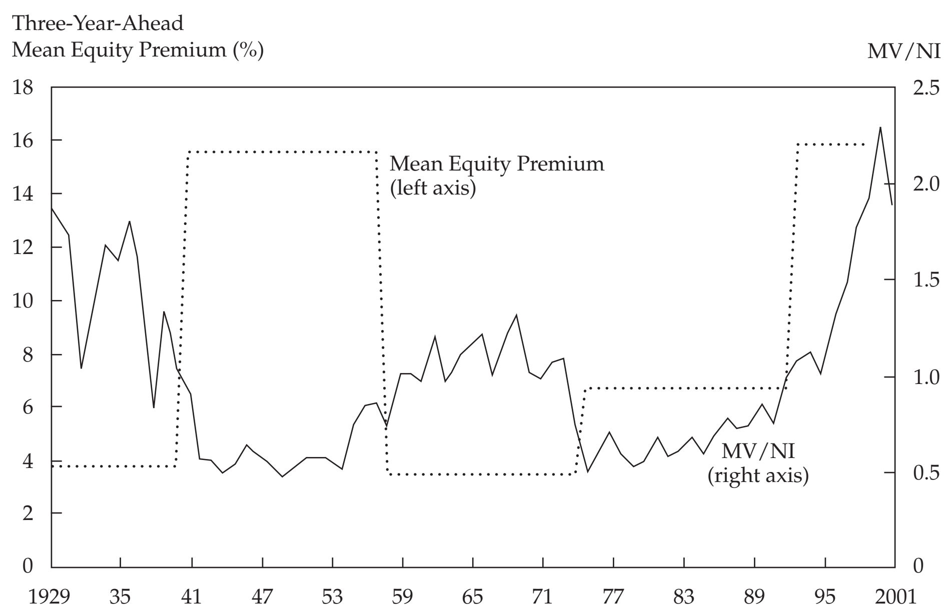


Note: $y = 4.7159x + 13.321$.

What could have been concluded from that information? The premium of 2 percent is the realization of a stochastic process with a large standard deviation. If the investor of 1928 saw any pattern in the stochastic process, optimizing agents would have endogenously changed the prices. That understanding makes it much more difficult to say we have a bubble. What we see is only one realization of a stochastic process. We would ideally like to see the realizations in many different, parallel universes and see how many times we actually came up with 2 percent and how many times we didn't. However, we are constrained by reality and observe only one realization!

The data used to document the equity premium are as good and clean as any economic data that I have seen. A hundred years of economic data is a long time series. Before we dismiss the equity premium, not only do we need to understand the observed phenomena (why an equity risk premium should exist), but we also need a plausible explanation as to why the future is likely to be different from the past. What factors may be important in determining the future premium? Life-cycle and demographic issues may be important, for example; the retirement of aging Baby Boomers may cause asset deflation. If so, then the *realized* equity premium will be low in 2010. But if asset valuations are expected to be low in 2010, why should the premium not be lower now? Perhaps what we are seeing in the current economy is the result of market efficiency taking the aging Baby Boomers into account. Either we will understand why a premium should exist (in which case, it will persist), or if it is a statistical artifact, it should disappear now that economic agents are aware of the phenomenon.

Figure 7. Mean Equity Risk Premium Three Years Ahead by Time Periods and Market Value/National Income, January 1929–January 2000



Note: The equity premium was averaged over time periods in which MV/NI > 1 and MV/NI < 1.

Current Estimates and Prospects for Change II

Rajnish Mehra

Professor of Finance
University of California, Santa Barbara
National Bureau of Economic Research and Vega Asset Management

SUMMARY

by Peter Williamson

Amos Tuck School of Business Administration
Dartmouth College, Hanover, New Hampshire

Rajnish Mehra proposed that analyzing the equity risk premium is an exercise in forecasting that has little to with the academic debate over whether the observed past excess return on equities presents a puzzle. Why is the equity premium a puzzle?

Table 1 shows real returns for long and not-so-long periods of time for the U.S. stock market, a relatively riskless asset, and the risk premium. A real return on equities of about 7 percent characterizes some long time periods, including 1889–1978, a period that did not incorporate the recent bull market. For the 1889–2000 period, the return was 7.9 percent. The standard deviation of annual returns was about 20 percent. Moreover, as **Table 2** shows, other countries have shown similar returns.

Table 1. Real U.S. Equity Market and Riskless Security Returns and Equity Risk Premium, 1802–2000

Period	Mean Real Return on Market Index	Mean Real Return on Relatively Riskless Asset	Risk Premium
1802–1998	7.0%	2.9%	4.1%
1889–2000	7.9	1.0	6.9
1889–1978	7.0[a]	0.8	6.2[b]
1926–2000	8.7	0.7	8.0
1947–2000	8.4	0.6	7.8

[a]Not rounded, 6.98 percent.
[b]Not rounded, 6.18 percent.

Sources: Data for 1802–1998 are from Siegel (1998); for 1889–2000, from Mehra and Prescott (1985).

U.S. T-bills have returned about 1 percent with a 4 percent standard deviation. Why are the returns on T-bills so different from those on equity? We might say we are looking at an aberration, but this time series is the best evidence we have. The difference defies easy explanation by standard asset-pricing models. Is it explained by risk differences? The answer is not clear.

Our theory tells us that assets are priced in such a way that, *ex ante*, the loss in marginal utility incurred by sacrificing current consumption to buy an asset at a certain price is equal to the expected gain in marginal utility contingent on the anticipated increase in consumption when the asset pays off in the future. The emphasis here is on *incremental loss or gain* of utility of consumption, which should be differentiated from incremental consumption because the same amount of consumption may result

Table 2. Real Equity and Riskless Security Returns and Equity Risk Premium: Selected Developed Markets, 1947–98

Country	Period	Mean Real Return on Market Index	Mean Real Return on Relatively Riskless Asset	Risk Premium
United Kingdom	1947–1999	5.7%	1.1%	4.6%
Japan	1970–1999	4.7	1.4	3.3
Germany	1978–1997	9.8	3.2	6.6
France	1973–1998	9.0	2.7	6.3

Sources: Data for the United Kingdom are from Siegel (1998); the remaining data are from Campbell (2002).

in different degrees of well-being at different times. As a consequence, assets that pay off when times are good and consumption levels are high—i.e., when the marginal utility of consumption is low—are less desirable than those that pay off an equivalent amount when times are bad and additional consumption is more highly valued.

This theory is readily illustrated in the context of the capital asset pricing model, in which good times and bad times are captured by the return on the market. Why do high-beta stocks yield a high expected rate of return? A high-beta security tends to pay off more when the market return is high—that is, when times are good and consumption is plentiful. Such a security provides less incremental utility than a security that pays off when consumption is low, is less valuable, and consequently, sells for less. Because rates of return are inversely proportional to asset prices, the former class of assets will, on average, give a higher rate of return than the latter.

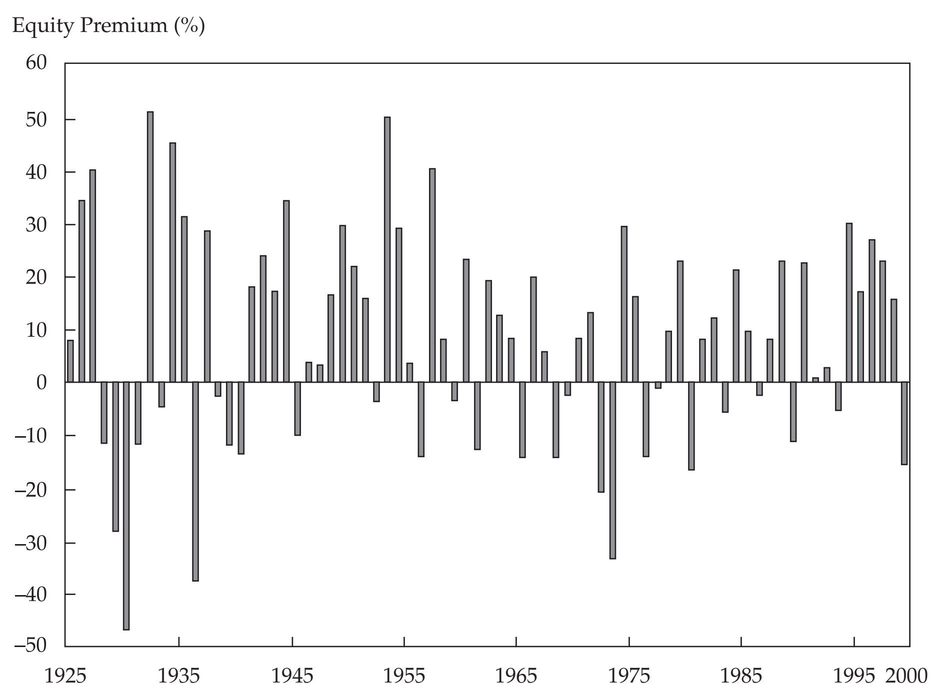
Another perspective emphasizes that economic agents prefer to smooth patterns of consumption over time. Assets that pay off a relatively larger amount at times when consumption is already high "destabilize" these patterns of consumption, whereas assets that pay off when consumption levels are low "smooth" out consumption. Naturally, the latter are more valuable and thus require a lower rate of return to induce investors to hold them. And such assets are purchased despite their very low expected rates of return. Insurance is an example.

What is surprising is that stocks and bonds *pay off in approximately the same states of nature* or economic scenarios. Hence, as Mehra argued earlier, they should command approximately the same rate of return. Using standard theory to estimate risk-adjusted returns, Mehra and Prescott (1985) showed that stocks, on average, should command, at most, a 1 percent (100 bps) return premium over bills. This finding presented a puzzle because the historically observed mean premium on stocks over bills was considerably and consistently higher.

The *ex post* excess return has varied a lot, which is not surprising. Graphs of the annual realized excess return in **Figure 1** and of the excess return for 20-year periods in **Figure 2** show dramatic differences.

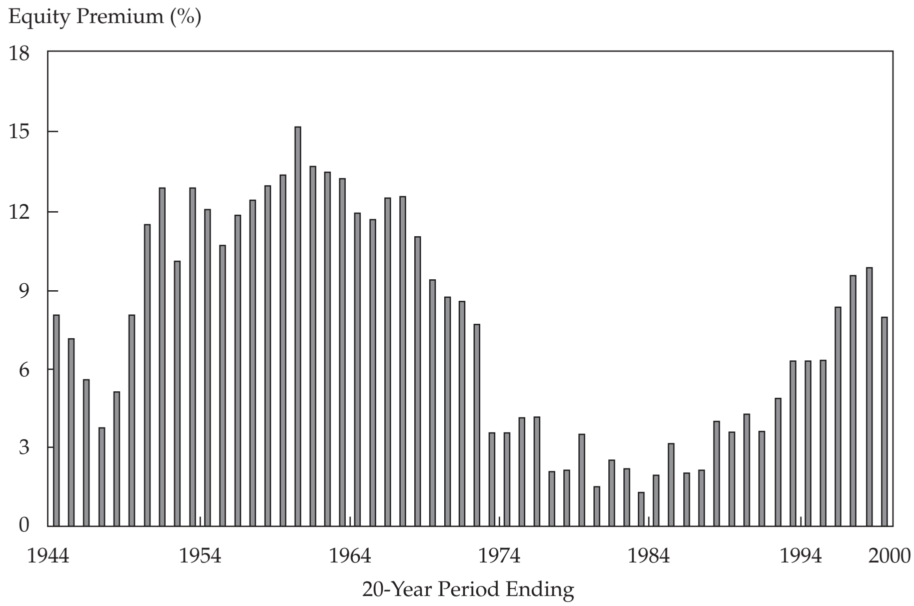
Mehra stressed that we need to distinguish the *ex post* excess return on equity from the *ex ante* risk premium. The expected equity premium *must* be positive. Following a bull market, the *ex post* will be high and the *ex ante* will be low. Over time, they will average out. A conclusion for the future depends on the planning horizon. Mehra was addressing the premium for the very long term—on the order of 50–100 years. In the short term, as in Figure 1, the variance in returns makes it quite impossible to come up with any reliable forecast. Figure 2 for 20-year periods, however, shows something more promising.

Figure 1. Realized Equity Risk Premium per Year, January 1926–January 2000



Source: Ibbotson Associates (2001).

Figure 2. Mean Equity Risk Premium by 20-Year Holding Periods, January 1926–January 2000

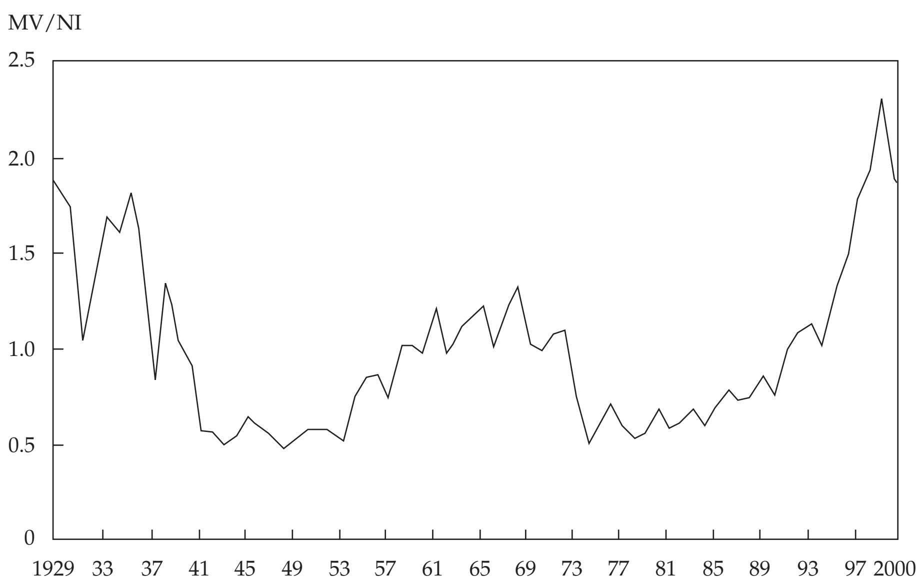


Source: Ibbotson Associates (2001).

Mehra's **Figure 3** showed the ratio of market value of equity (MV) to national income (NI) since 1929, and his **Figure 5** overlaid on that graph the three-year-ahead equity premium.[1] The ratio has ranged from 2 × NI to 0.5 × NI to 2.25 × NI. In **Figure 7**, Mehra split the 1929–2000 period into subperiods—those in which MV as a ratio of NI was greater than 1 and those in which it was less than 1—and overlaid on that graph is the three-year-ahead mean equity premium. Figure 7 shows that we have had two and a half cycles since 1929, and they reveal some predictive ability: On average, when MV/NI is low, the risk premium is high, which is useful as a guide for the very long term.

[1] Table and figure numbers in each Summary correspond to the table and figure numbers in the full presentation.

Figure 3. U.S. Stock Market Value/National Income, January 1929–January 2000



Source: Data updated from Mehra (1998).

Figure 5. Mean Equity Risk Premium Three Years Ahead and Market Value/National Income, January 1929–January 2000

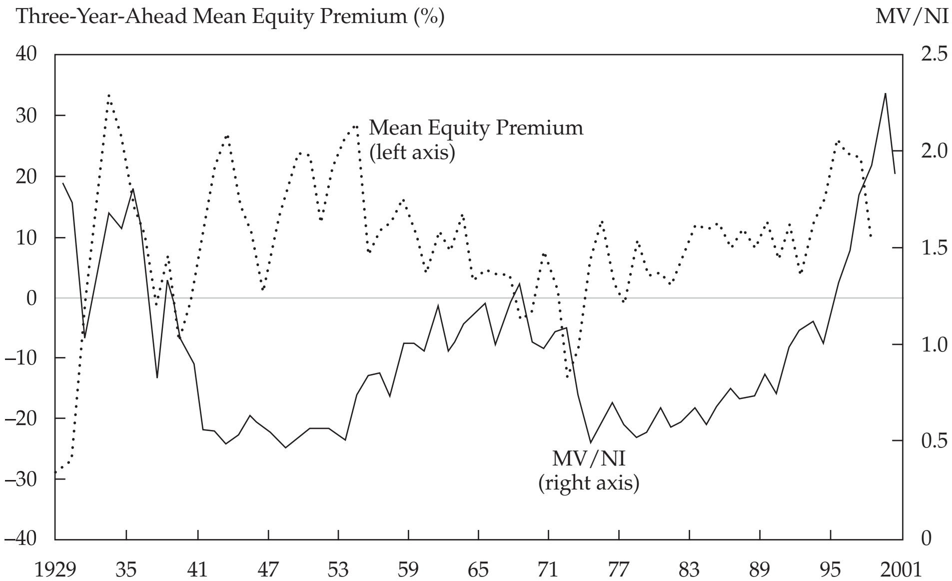
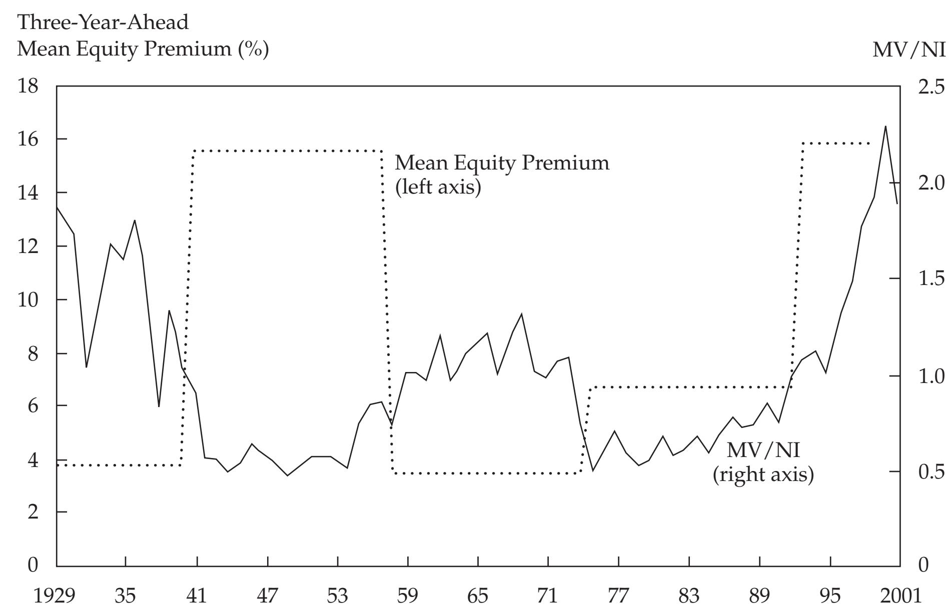


Figure 7. Mean Equity Risk Premium Three Years Ahead by Time Periods and Market Value/National Income, January 1929–January 2000



Note: The equity premium was averaged over time periods in which MV/NI > 1 and MV/NI < 1.

Mehra suggested that individuals who are interested in short-term investment planning will wish to project the conditional equity premium over their planning horizon. But doing so is by no means a simple task. It is isomorphic to forecasting equity returns. Because returns have a standard deviation of 20 percent, the noise dominates the drift. Operationally, how much information comes from knowing that the mean risk premium is 2 percent rather than 6 percent when the standard deviation is 20 percent?

In conclusion, Mehra considered how the world must have looked to an investor at the end of 1928.

The mean real return on the S&P 500 had been 8.52 percent for 1889–1928, and the mean real return on risk-free assets had been 2.77 percent, so the observed mean equity risk premium would have been 5.75 percent (575 bps). An analysis similar to the Mehra–Prescott (1985) analysis, however, would have indicated an *ex ante* premium of 2.02 percent.

Is the future likely to be different from the past? To decide, we need to focus on what factors might make the future different. Demographic changes, for example, could be very important. But, maybe, because of market efficiency, the market has already taken into account the likely changes.

Current Estimates and Prospects for Change: Discussion

John Campbell (*Moderator*)
Ravi Bansal
Bradford Cornell
William Goetzmann
Roger Ibbotson
Martin Leibowitz
Rajnish Mehra
Thomas Philips
William Reichenstein, CFA
Stephen Ross
Robert Shiller
Jeremy Siegel

JOHN CAMPBELL (Moderator)

I'll make a few remarks and then open the discussion. I would like to amplify a distinction that Raj Mehra was making between the *ex post*, realized premium over some past period and the *ex ante* premium that investors are expecting at a single point in time. Over the long run, these premiums have to average out to the same level if the market has any rationality at all, but in the short run, they can move quite differently. For example, a lot of Raj's graphs indicate that the *ex post* and *ex ante* risk premiums might move in opposite directions, and I think that concept is very important to keep in mind. If we go through a period when the *ex ante* premium falls (for whatever reason), that movement will tend to drive prices up for a given cash flow expectation, so we will see a high realized return during a period when the *ex ante* premium has actually fallen. That is the story of the 1990s—that average returns were high, particularly at the end of the decade, because investors were willing to take on more risk, so the required rate of return was declining. Thus, we had a decline in the *ex ante* equity premium at the same moment that we had very high average returns.

Of course, if the equity premium is estimated by use of historical average returns, even over a period as long as 100 years, a few good years can drive up the long-term average considerably. For example, over 100 years, a single good yearly return of 20 percent adds 20 bps to the 100-year average return. This is the problem with estimating the equity premium from historical average returns; there is so much noise, and the average will tend to move in the wrong direction if the true *ex ante* premium is moving.

As a result, the methodology used by many at this forum is to focus on valuation ratios at a single point in time and make adjustments for growth forecasts. The methodology can be applied simply or elaborately. You can simply look at the earnings yield, or you can try to adjust the yield for return on equity being greater than the discount rate equilibrium or Tobin's q being different from 1, which we discussed this morning [in the "Historical Results" session]. I think this approach is the right way to go. If you want to estimate the *ex ante* premium, you start with a valuation ratio that summarizes the current state of the market, make some adjustments based on your best judgment, and back out the *ex ante* premium.

The approach has two difficulties that one has to confront. They arise from the fact that the models we are using are steady-state models that give long-term forecasts in a deterministic setting. The problem with using a deterministic model is that you obliterate any distinction between different kinds of averages. In a random world, however, that distinction matters a lot. It matters to the tune of 1.5–2.0 percentage points.

The second problem is that a forecast from a valuation ratio is really the equivalent of the yield on a long-term bond. The valuation ratio produces an infinite discounted value of future returns. You don't necessarily know the sequence of predicted returns. You don't know the sequence of forward rates or the term structure; you just have a single measure of a long-term yield. So, it's very difficult to construct or generate a view about the actual path that returns might follow.

In my work with Bob Shiller, we argue that, given the level of prices, this long-term yield must be very low. But that argument is consistent with two different views about the time path. One view is that a correction is going to occur in the short or medium term, followed by a return to historical norms. If you hold this view, you have to be bearish in the short term but you are more optimistic about returns in future years. This outlook would be very pessimistic for an investor who has finished accumulating wealth and wants to cash out; it would be a more optimistic outlook for an investor who expects to accumulate assets over the next several decades.

The other view, which I think has some plausibility, is that we might see mediocre returns over the long term because of structural changes—structural changes in that transaction costs have come down, the costs of diversification have come down, investors have learned about the equity premium puzzle, and therefore, the *ex ante* premium is down and will be permanently down. This view is less bearish in the short term than the first view but also less optimistic in the long term.

I think Bob and I differ a little bit on this time-path issue in terms of how to chop up the long-term yield into a sequence of forecasts. Bob is probably closer to the view that returns will be very poor in the short term and then revert to historical norms, and I am closer to the view that there may have been a permanent structural change that will mean mediocre returns in the near term and the longer term.

It is hard for me to imagine a long-run equilibrium with an equity premium relative to U.S. T-bills less than about 1.5 percent geometric (2.5–3.0 percent arithmetic). And I think it may take a further price decline to reach that long-run equilibrium. In other words, we are in for a short period of even lower returns followed by a (geometric) premium of about 1.5 percent for the long term.

MARTIN LEIBOWITZ: One thing we have not talked much about is that if, over time, we have more data on earnings, price movements, and returns, what is going to be the catalyst for moving the risk premium to higher or lower levels—or to a point of acceptance? Of course, one of the really great things about the market is its ambiguity; even if you are earning dismal returns now, the market's volatility always allows you to look back at a recent period when you earned great returns. But what sequence of events and flow of information would wake up market participants to say, "Hey, a 2 percent equity risk premium? I'm not buying for 2 percent. Give me something else. Is there another market I can invest in? Is there another advisor out there?" This possibility is worth thinking about because if we make the rounds and tell our friends and professional colleagues, "Look, we've found out that the nominal, arithmetic equity risk premium is roughly only 3.0–3.5 percent, and that's going to be it, but I can give you some good news: Volatility will be relatively low, so you will really be getting a lot of return for the amount of risk you'll be taking," people will say, "Forget it!" I would not want to be invested in the equity market with that sort of outlook. People would just run away from the equity market. People are thinking, hoping, and dreaming of returns well over an equity premium of 3 percent; they are thinking of a risk premium greater than that. This kind of question is what we need to discuss.

RAJNISH MEHRA: This point is the reason that understanding *why* we have an equity premium is so important. On the one hand, if there is a rational reason for the equity premium—for instance, if investors are scared of recessions and actually demand a 6 percent equity premium, then I would expect a 6 percent premium in the future. On the other hand, if we find out that investors do not actually demand that premium for holding stocks—that they perceive stocks, in some sense, to be not much riskier than bonds—then, the premium will be lower. You seem to be saying that investors *do* perceive stocks to be much riskier than bonds and they *do* want a high premium, in which case they will get it. If investors refuse to own stocks when they get only a 2 percent premium, a repricing of assets will take place.

STEPHEN ROSS: One thing that we all agree on is that there is enormous estimation error in figuring out the risk premium. I find it ironic that the estimation error in the risk premium that we agree on plays no role whatsoever in the models that we use to infer the risk premium. It is somewhat like option pricing, where you assume you know the volatility. You look at the option price, and then you figure out what the volatility must be for that to be the option price. Then, you build models of what the option price should be. But estimating the risk premium is even more complicated, and estimation error is even more damaging.

The estimation error in estimating the risk premium is huge. Over a 100-year period, the standard error alone of the sample estimates is on the order of 2–3 percent. I am not convinced by John Campbell's argument that structural models, which are efforts to get conditional probability estimates and do a better job of conditioning, will improve the situation, because we have about the same volatility on our conditional estimates. I have a very pessimistic view of those models. They introduce other parameters, and where we had 2 percent standard errors on a few parameters, now we have 4 percent because we have more parameters. I'm not convinced that this approach will narrow down the estimate.

I am troubled by the fact that in this world of incredible volatility, and with no real confidence in our estimations of the risk premium, we still go ahead and advise people about what to do with their portfolios. As Rajnish Mehra said, we have a strange disconnect: The uncertainty that we all perceive in these models plays no role in the construction of the models. As a consequence, uncertainty plays no role in our ability to filter from the models better estimates. One of the things we have to think seriously about is estimation error in these models.

THOMAS PHILIPS: I share John Campbell's view that, barring an unforeseen surge in productivity, we are in for a prolonged period of lower returns prior to transaction costs and fees. However, the actual return that will be *realized* by investors net of transaction costs and fees is probably not very different from the return achieved in the past. Don't forget that index funds did not exist in 1926. In those days, transaction costs and fees subtracted 2–3 percent each year from returns; today, costs have fallen by 90 percent.

WILLIAM REICHENSTEIN: A number of models predict returns using a dividend model. In this model, long-run return is the current dividend yield plus long-run expected growth in dividends plus the percentage change in price divided by the dividend multiple, P/D. When predicting returns, analysts tend to drop the last term and predict the capital gains as the long-run growth in dividends. In the corresponding earnings model, predicted return is the current dividend yield plus the capital gains (the long-run growth in earnings) plus the percentage change in P/E. That has to hold; it is a mathematical certainty.

The reason I do not like the dividend model but like the earnings model is that we have no idea where the P/D multiple is going to go. Yet, the predictions from the dividend model assume it will remain constant. I can accept that there is some normal range for the P/E multiple, but I agree with Fisher Black that there is no normal range for the P/D multiple. Black looked at the various arguments to try to explain why companies pay dividends, and in the end, he threw up his hands and said we have no idea. If we have no theory or empirical evidence to explain dividend policy, then we have no reason to believe the P/D multiple is going to be stable. And we have no way of predicting it. That ratio could go to infinity. Therefore, any model that drops out that term, even for a long-run analysis, may be very, very wrong.

BRADFORD CORNELL: The dividend ratio may not be stable. In fact, we are seeing declining dividends, but you may have a constant payout ratio.

REICHENSTEIN: If we wanted to estimate the ending P/E after the next 50 years, whatever we came up with, we might feel reasonably confident it is going to be between 30 and 8.

ROSS: It is higher than 30 now!

REICHENSTEIN: Let's say that something will stop the P/E multiple from going too high or too low. But if you ask what the ending P/D multiple will be, well, if companies keep dropping dividends, it could be a billion.

CORNELL: That is why you might want to include payouts. Wouldn't you think that political pressures would arise to make sure shareholders got a certain fraction, on average, of corporate earnings? If shareholders do not get some share, they will become dissatisfied and companies will not be able to issue equity. Corporations cannot play the game of siphoning off all the earnings indefinitely for executives' perks and options and so forth.

ROGER IBBOTSON: You do not have to get your return through dividends. If the company is bought out, you can get your money out. You can get your money out in lots of ways other than dividends. Speaking for myself, if I had a choice, I would not want to get any of my money out in dividends.

MEHRA: Tandy Corporation, for instance, does not pay out any dividends. It was sued by the U.S. IRS, which charged that it was helping stockholders evade taxes. The company successfully won the case with an argument that it had a diverse group of stockholders and was not acting in the interest of any particular shareholder group. A rational approach would be for shareholders, instead of receiving a dividend payment, to sell shares and pay a capital gains tax when they want cash.

REICHENSTEIN: Yes, we do end up paying taxes. So, if you are only able to tell me that 50 years from now, the P/D multiple could be anywhere from infinity to something much, much lower, then that is a heck of an estimation error.

ROSS: The interesting question being raised is whether price to dividends is the variable you should be looking at or whether we should be asking: Is there stability in price divided by total payout, including stock repurchases, dividends, and Roger Ibbotson's suggestion that there is a constant probability that you will get a cash offer for the holding? So, the totality of all the payouts would be an interesting long-term variable to look at that may well be quite stable.

CORNELL: There are also some monies that go the other way, however, so the effective payout rate is very hard to compute.

REICHENSTEIN: But if you are using a model and put in the current dividend yield to project long-run growth and if dividends come from some historical average, then in a period like the past 20 years (in which we have had this dramatic fall in dividend payout rates and dividend yields), if you don't include repurchases, you have a problem. Past growth is going to be below future growth, and the dividend model predictions miss this point. I think Stephen Ross is saying that dividend payouts are unstable but might be stable if we added back in repurchases. In my view, the dividend model is a questionable framework.

RAVI BANSAL: Both Rajnish Mehra and Bob Shiller commented on the size of the premium but didn't comment on, or make predictions about, the underlying volatility of the market portfolio. From John Campbell's comment, if I am interpreting it correctly, he views the current scenario as a form of a drop in the Sharpe ratio. Has uncertainty fallen or risen? What is happening to the Sharpe ratio?

CAMPBELL: There haven't been any long-term trends in the volatility of the market as a whole. Certainly, marketwide volatility fluctuates. Volatility was unusually low in the mid-1990s and has risen a lot since then, but if you look over decades, you don't see any trend. The result is different when you look at the idiosyncratic volatility measure, however, because then you do see a trend over the last three decades. But looking marketwide, we do not see trends. Actually this lack of trend is a puzzle because of the evidence that the real economy has stabilized. GDP growth seems to be less volatile. So, some people claim that risk has fallen, which would justify the fall in the equity premium. Yet, we don't see that lower volatility when we look at short-term stock returns. The market does not appear to think that the world is any less risky.

JEREMY SIEGEL: Could I suggest something? Because real uncertainty has declined, companies can lever up more, generate higher P/Es. The result is maintenance of equity volatility, but it's because of an endogenous response to the increased real stability of the economy. So, greater leverage and higher P/Es could be generating the same equity volatility, which wouldn't be a puzzle even with the more stable real economy.

CAMPBELL: But if companies have levered up to maintain the same equity volatility, the equity premium should not fall as a result.

SIEGEL: Yes, if you don't take labor income being more stable into account as one of the factors that might determine risk preferences. In fact, some research shows that if there were more stability on the wage side (labor income), that stability would give people more incentive to buy equities.

WILLIAM GOETZMANN: Just a word on dividends: With all the studies that have looked at historical dividend yields, the problem is that we do not know very much about the dividends on which the studies were based. For data before 1926, we have the Cowles Commission (1938) information on dividends, but when you start reading Cowles' footnotes, you see he had a problem figuring out whether he was actually identifying all the dividends that were being paid by the companies.

ROBERT SHILLER: Have you solved this problem? We had the same problem.

GOETZMANN: Well, no, but we found it was a striking problem. We started from the Cowles period and worked back to see if we could collect information on dividends. We have the information back to the 1820s or so, but we could be missing dividends.

SHILLER: You're concerned that you don't have all the information, that you are missing a significant chunk of it?

GOETZMANN: Yes. You have a set of stocks that are similar to each other—their industrial characteristics are similar, for example. One stock may be paying 8 percent dividends for 10 years, but for another stock, you have no dividend information available. Are you to presume that the second stock did not pay any dividends or that your records simply do not show the dividend? So, what we have had to resort to is to report the high number and to report the low number. And we don't think anybody else has ever really been able to get any better information about dividends than we have. So, if we're going to talk about model uncertainty, let's also talk about data uncertainty—particularly as the records go back through time.

SHILLER: Do you think that companies sometimes reported dividends to commercial and financial chronicles and at other times, misreported them or didn't report them at all?

GOETZMANN: Yes, that could be true.

SHILLER: Wouldn't it have to happen on a big scale to affect the aggregate numbers?

IBBOTSON: As you go back in time, it is not clear who or what was getting the reports. For one period of time, there was an official source for the NYSE, but later, that source disappeared. It is hard enough to get actual stock price data, but it is much harder to find out who reported dividends to whom. Therefore, dividend information comes from all sorts of sources.

GOETZMANN: So, for what it's worth, sprinkle some more noise into this whole process. It's a real challenge to focus on valuation ratio regressions. We've been talking about valuation ratio regressions and statistics in one form or another for eight or nine years now, and we have all sorts of details about the econometrics, but the real issue to me is whether we really know what the payouts were as we push backward in time.

IBBOTSON: For the stock price data, we only needed to go to one (or possibly two or three) sources, but for the dividend data, we had to go to many sources, and even after going to many sources, we found we were getting only some of the data. However, when we found the data, companies paid all their earnings out in dividends. They had 100 percent payout ratios in the 19th century. But for the missing data—who knows.

ROSS: In this entire discussion, we are focusing entirely on the risk premium, and we have sort of ignored the other variable, volatility. What is interesting about volatility is that it is the one variable about which we do have confident expectations.

Volatility has two features that are curious. One feature is that we can actually measure volatility with a certain amount of precision; we know what volatility is. Volatility is a lot less ambiguous than the equity risk premium. We need to bring volatility to bear on such questions as long-run portfolio allocation problems. Someone who has great estimation error about the risk premium and cannot quite figure out what it is but who, nonetheless, is taking others' advice as to what to do, would perhaps be informed in this decision by observing that we do know a lot about the pattern of volatility, we have far less estimation error for it, we sort of know what volatility is today, and we have pretty good ability to predict it over fairly long horizons. At least this person should understand the volatility of volatility, which shows up as much in those allocation problems as does expected return.

The second curious feature of volatility is, it seems to me, that we can use this variable in some interesting ways. Implied volatilities have been around now for 20 years. I know that the week before the 1987 crash, implied volatilities went to an annualized rate of about 120 percent. Prior to the current crash, implied volatilities again rose substantially. The cynic would say, well, implied volatility was quite high, but people didn't know whether the market was going up 200 points or down 200 points the next day; they just knew it was going to be a big move. But my guess is that investors figured that the market wasn't going to go up much more; they really thought the market was going to go down. It would be nice for those who are doing the empirical work on the risk premium to have a variable that actually has expectation recorded in it. It might be fun to look at its empirical content for the puzzles we are talking about today.

SIEGEL: I would like to add something to that comment. I think we know short-run volatility because we can measure it using options, most of which are very short term. But the question of long-run volatility depends very much on the degree of mean reversion, which is very important for long-term investors and is, as we all know, subject to great debate.

ROSS: Actually, I suspect long-term volatility is subject to less debate than long-run returns. For short-run volatility, even for an option one year out, with pretty good liquidity, you can start to see reversion—pretty clear reversion—one year out.

SIEGEL: But we don't have 10-year, or 20-year, or 30-year options, which might be very important for longer-term investors.

ROSS: Volatility is a lot better measure than returns, for which we have nothing that tells us anything about the short term or the long term.

SHILLER: I want to remind you of the very interesting discussion in Dick Thaler's talk this morning about *perceived* volatility [See the "Theoretical Foundations" session]. We seem to be forgetting about the distinction between the actual and the perceived risk premium. When Marty Leibowitz was saying that people would not be interested in stocks with an equity premium of 1.5 percent, he may have been assuming that the perceived volatility was very high. Dick was saying that it is the *presentation* to the general public that affects the public's perception of volatility. His research disclosed a very striking result, which is that when you present investors with high-frequency data, they have a much different perception of what the data are saying than when you present them with less-frequent—say, annual—data. And the way the data are being presented is changing. When I walk down the street now, I can look up at a bank sign that alternates between time, temperature, and the Nasdaq.

LEIBOWITZ: I have a couple of comments. First, if you had a volatility estimate that you could live with and you had actual asset allocations that were stable and common—most asset allocations, at least by institutional investors, are surprisingly stable and common—you could (theoretically) clearly back out from those variables the implied risk premium. No big challenge. At least, you could back out mean–variance estimates. Of course, the question is: What kind of time horizon would you be looking at? The horizon would be the critical ingredient. If you were looking over a long enough time horizon, the risk premium could be 0.1 percent. If you were looking over a short horizon, the risk premium could be something enormous.

Robert Merton wanted me to introduce along these lines the Zvi Bodie construct.[1] Bodie says that the kind of option you would have to buy as you go out to very long horizons is very different, in terms of the Sharpe ratio, from a short-horizon option; it is a very expensive option. That reality has to tell you something.

The other thing that I want to mention is that the issue of equilibrium payout ratios is very important. The question is: When an equilibrium is reached, at which point earnings are growing at either the growth rate of the economy or near that rate (i.e., that rate is your stable equilibrium view), then in terms of dividends, how much of a company's aggregate earnings have to be put back into the company to sustain that growth? This is the critical question. All else would then follow from the answer. It's surprising that this issue has not been much addressed, as far as I know, even from a macro level.

[1] Robert Merton was invited to attend but could not.

PHILIPS: There is a pragmatic solution to the question that Stephen Ross and Jeremy Siegel raised. We have about 20 years of option data, so you might construct the volatility data going back 20 years, and you could explore the fact that as you sample faster and faster, the estimates of volatility get sharper and sharper. Just take a perfect-foresight model: Assume it's 1920, and you're going to assume that the world is rational and that the forecasted volatility would have been the volatility that was actually realized over 1921, or 1921–1925, or whatever years you want to use. From those data, you could impute a data series going back in time and then try to do the appropriate tests. Cliff Asness has a very nice paper in the *Financial Analysts Journal* that explores this approach (2000b). Cliff looks at historical volatility and then backs out future returns as a function of historical volatility.

EQUITY RISK PREMIUM FORUM, NOVEMBER 8, 2001

Implications for Asset Allocation, Portfolio Management, and Future Research I

Robert D. Arnott

First Quadrant, L.P.
Pasadena, California

A practitioner's empirical approach to estimating prospective (expected) equity risk premiums does not bode well for finding alpha through conventional U.S. equity allocations. In the United States and the United Kingdom, real earnings and real dividends have been growing materially slower than real GDP. Based on empirical evidence, if today's dividend yield is 1.7 percent and growth in real dividends is about 2.0 percent, cumulative real return on stocks will be about 3.7 percent. With a 3.4 percent real yield on bonds available, the *ex ante* risk premium all but disappears. Perhaps most troubling in the empirical evidence is the 60 percent negative correlation between payout ratios and subsequent 10-year earnings growth. With current payout ratios close to 40 percent, the implication for earnings growth over the coming decade is a rate of about –2 percent. When an assumed negative earnings growth rate is combined with an assumed zero risk premium, we have a serious problem.

I have to begin by offering profuse apologies. You are seasoned, very capable academics, and I'm not. I'm just a practitioner and an empiricist. So, we're going to focus on practice and empiricism in this presentation and stay far away from the theory related to the equity risk premium.

History versus Expectations

First, I want to emphasize an observation that a number of speakers have made: Much of the dialogue about the risk premium is very confused because the same term, "risk premium," is used for two radically different concepts. One is the historical excess return of stocks relative to bonds or cash, and the other is the prospective risk premium for stocks relative to bonds on an *ex ante* basis, without any assumptions about changes in valuation levels. The two concepts are totally different, should be treated separately, and, I think, should carry separate labels. Excess returns measure past return differences. The risk premium measures prospective return differences. I wish the industry would migrate to using different terms for these two radically different concepts.

A quick observation: If you are a bond investor and you see bond yields drop from 10 percent to 5 percent, and in that context, you have earned a 20 percent return, do you look at those numbers and say, "My expectation of 10 percent was too low. I have to ratchet my expectation higher. I'll expect 12–15 percent"? Of course not. The reaction by the bond investor is, "Thank you very much for my 20 percent returns; now, I'll reduce my expectation to 5 percent." If the earnings yield on stocks falls from 10 percent to 5 percent, however, what is the investment community's response when they see the 20 percent return? They say, "Our expectations were too low! Let's raise our expectations for the future."

My impression of the discussion we have been having today is that the reaction in this room would be absolutely unanimous in saying the portion of return attributable to the drop in the earnings yield (earnings to price) or the drop in the dividend yield can and should be backed out of the historical return in shaping expectations. I haven't heard a lot of discussion of the fact—and I think it is a fact—that a drop in the earnings yield should have a second-stage impact. The first stage is to say 10 percentage points (pps) of the return came from falling earnings yields; therefore, let's back that out. The second stage is that the fall in the earnings yield should produce a haircut in future expectational returns. I don't hear this concept out in the marketplace, and I don't hear it much in the academic community either.

Strategic Implications of Lower Returns

Let's begin with the hypothesis that the risk premium, the forward-looking premium, on U.S. stocks is now zero. Please accept that supposition for the next few minutes. If the risk premium is zero, what is the implication for asset allocation policy? In the past, the policy allocation to stocks and fixed income was the king of asset management decisions. It was the number one decision faced by any U.S. institutional investor—indeed, any investor in general. The reason was that more stocks meant more risk and more return.

The fiduciary's number one job was to gauge the risk tolerance of the investment committee and to push the portfolio as far into stocks as that risk tolerance would permit. If that job was done correctly, the fiduciaries had succeeded in their primary responsibility. But if stock, bond, and cash real returns are similar, if the risk premium is approximately zero, then it doesn't matter whether you have a 20/80 equity/debt or an 80/20 equity/debt allocation. It does affect your risk and your year-by-year returns, but it doesn't affect your long-term returns. So, if the risk premium is zero, this fundamental policy decision is radically less important than it has ever been in the past.

As for rebalancing, the empirical data support the notion that rebalancing can produce alpha, but we do not have a lot of empirical data to support the notion that rebalancing adds value. History suggests that rebalancing boosts risk-adjusted returns, but it sometimes costs money. Rebalancing produces alpha by reducing risk, and in the long term, it typically adds some value in addition to risk reduction. Now, suppose we are in a world in which there is no risk premium and in which stocks and bonds have their own cycles, their own random behavior. If that behavior contains any pattern of reversion to any sort of mean, rebalancing suddenly can become a source not only of alpha but also of actual added value—spendable added value.

In the past, tactical asset allocation (TAA) provided large alpha during periods of episodic high returns but did not necessarily provide large added value. So, the actual, live experience of TAA in the choppy, see-saw market of the 1970s was awesome. In the choppy bull market of the 1980s, value added from TAA was not awesome but was still impressive. In the relentless bull market of the 1990s, the value added from TAA was nonexistent. Alpha was certainly still earned in the 1990s (a fact overlooked by many), but it came mostly from reduced risk. If we are moving into markets like those of the 1970s, then TAA certainly merits another look.

What about the strategic implications of lower returns for pension funds? If conventional returns lag actuarial returns, then funding ratios are not what they seem. I did a simple analysis of funding ratios for the Russell 3000 Index and found that for every 1 pp by which long-term returns fall short relative to actuarial returns, the true earnings of U.S. pension assets fall by $20 billion. If, as I believe is the case, long-term returns are going to be about 3 pps below long-term actuarial assumptions, pension fund earnings will be $60 billion less than what is being reported, and this shortfall will need to be made up at some later date.

In a world of lower returns, if you don't believe in efficient markets, alpha matters more than ever before. If you do believe in efficient markets, the avoidance of negative alpha by not playing the active management game matters more than ever.

Now, a truism would be that conventional portfolios will produce conventional returns. That is fine if conventional returns are 15 percent a year, as they were for the 18 years through 1999. In a market environment of 15 percent annual returns, another 1 pp in the quest for alpha doesn't matter that much to the board of directors, although it does make a material difference to the health of the fund. However, if the market environment is producing only 3–4 percent real returns for stocks *and* bonds, another 1 pp matters a lot.

What investments would be expected to consistently add value in a world of lower expected returns? "Conventional" alternative investments may or may not produce added value. Private equity and venture capital rely on a healthy equity market for exit strategies. They need a healthy equity market to issue their IPOs (initial public offerings). Without a healthy equity market, private equity and venture capital are merely high-beta equity portfolios that can suffer seriously in the event of any sort of reversion to the historical risk premium. International equities and bonds may have slightly better prospects than U.S. equities and bonds, but not much better.

Strategies well worth a look are the elimination of slippage, through the use of passive or tactical rebalancing, and cash equitization. If the equity risk premium is lost, then alternative assets whose returns are uncorrelated with the U.S. equity market will absolutely produce added value. Uncorrelated alternatives include TIPS,[1] real estate, REITs (real estate investment trusts), natural resources, and commodities. Absolute return strategies (market-neutral or long–short strategies and other hedge fund strategies) will also absolutely produce added value—if you can identify strategies that *ex ante* have an expectation of alpha. These approaches are, more than anything else, bets on skill and bets on inefficient markets. So, the investment strategies that will work in a world of lower returns differ greatly from the conventions that are driving most institutional investing today.

These reflections are from the vantage point of a practitioner. Much of what I've said makes the tacit assumption that markets are quite meaningfully inefficient, so these comments might be viewed with a jaundiced eye by a group that accepts market efficiency. Now, let's turn from practice to empiricism.

Empirical Experience

The Ibbotson data going back 75 years show about an 8 percent cumulative real return for stocks (see Ibbotson Associates 2001). Starting at the end of 1925 with a 5.4 percent dividend yield, the valuation attached to each dollar of dividends quadrupled in the 75-year span. That increase translated into nearly a 2 percent a year increase in the price/dividend valuation multiple—hence, 2 pp of the 8 percent real return. I think nearly everyone in this room would feel comfortable backing this number out of the returns in shaping expectations for the future. Over the 75-year period, real dividends grew at a rate of 1 percent a year. So, over the past 75 years, stocks produced an 8.1 percent real return. The real yield at the start of this period was 3.7 percent. (I say "real" yield because the United States was still on a gold standard in 1925; inflation expectations were thus zero. Bonds yielded 3.7 percent, and bond investors expected to earn that 3.7 percent in real terms.) Bonds depreciated as structural inflation came onto the scene. So, stocks earned a cumulative 4.7 percent real return in excess of the real return earned by bonds over the same period.

What does the future have in store for us from our vantage point now in the fall of 2001? **Table 1** contains the Ibbotson data and our analysis of the prospects from October 2001 forward. We'll start with a simple model to calculate real returns for stocks:

Real stock return = Dividend yield
+ Dividend growth
+ Changes in valuation levels.

In October 2001, the dividend yield is roughly 1.7 percent. If we assume that stock buybacks accelerate the past growth in real dividends, we can double the annual growth rate in real dividends observed over the past 75 years to 2 percent. Those two variables give us a 3.7 percent expected annual real return. TIPS are currently producing a 3.4 percent annual real return. Thus, the expected risk premium is, in this analysis, 0.3 pp, plus or minus an unspecified uncertainty, which I would argue is meaningful but not huge.

Why was the historical growth in real dividends (from 1926 through 2000) only 1 percent a year? Did dividends play less of a role in the economy? Were corporate managers incapable of building their companies in line with the economy? I don't believe either was the reason. The explanation hinges on the role of entrepreneurial capitalism as a diluting force in the growth of the underlying engines for valuation—that is, earnings and dividends of existing enterprises. The growth of the economy consists of growth in existing enterprises and the creation of new enterprises. A dollar invested in the former is not invested in the latter. **Figure 1** shows real GDP growth, real earnings per share (EPS) growth, and real dividends per share (DPS) growth since January 1970. Over the past 30 years, until the recent earnings downturn, real earnings have almost kept pace with real GDP

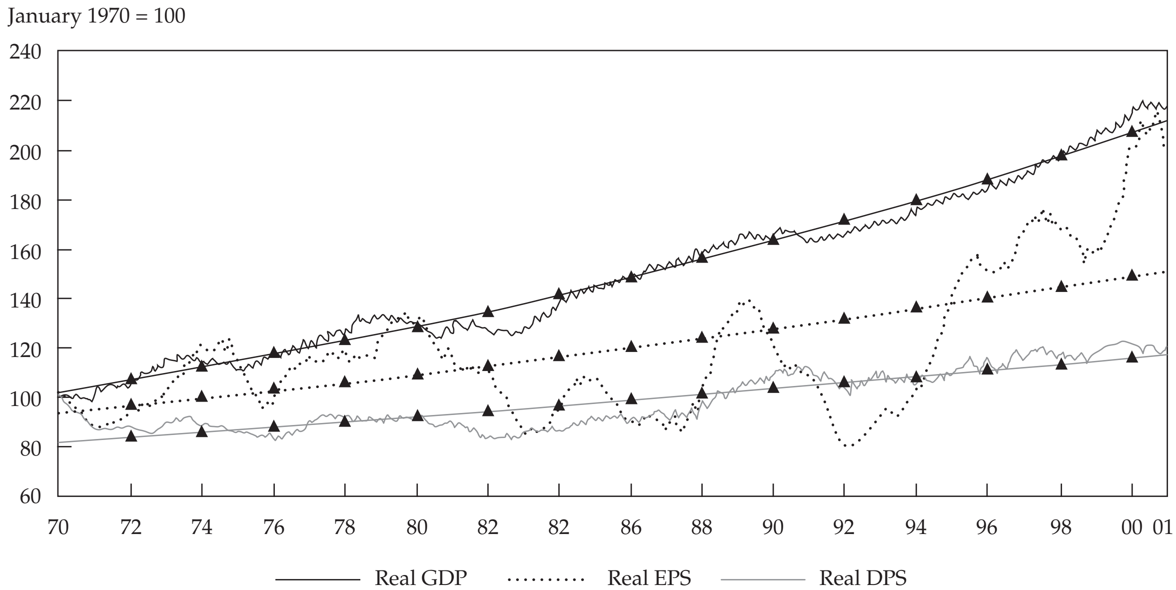
Table 1. The Ibbotson Data Revisited and Prospects for the Future

Component	75 Years Starting December 1925	Prospects from October 2001
Starting dividend yield	5.4%	1.7%
Growth in real dividends	1.0	2.0
Change in valuation levels[a]	1.7	???
Cumulative real return	8.1	± 3.7
Less starting bond real yield	3.7[c]	3.4[d]
Less bond valuation change[b]	–0.4	???
Cumulative risk premium	4.7	± 0.3

[a] Yields went from 5.4 percent to 1.4 percent, representing a 2.1 percent increase in the price/dividend valuation level.
[b] Bond yields went from 3.7 percent to 5.5 percent, representing a 0.3 percent annualized drop in long bond prices.
[c] A 3.7 percent yield, less an assumed 1926 inflation expectation of zero.
[d] The yield on U.S. government inflation-indexed bonds.
Source: Based on Ibbotson Associates (2001) data.

[1] TIPS are Treasury Inflation-Protected Securities; these securities are now called Treasury Inflation-Indexed Securities.

Figure 1. GDP, EPS, and DPS: United States, January 1970–January 2001



Note: Triangles identify exponentially fitted lines.

Source: Data from Organization for Economic Cooperation and Development (OECD).

growth. However, this pattern has occurred in the context of earnings as a share of the macroeconomy rising from below historical norms to above historical norms, including a huge boom in the 1990s. From the line of best fit, we can see that the growth trend in real earnings and real dividends is materially slower than the growth in the economy.

Is the picture different in Canada? Yes, it is. **Figure 2** illustrates that real earnings and real dividends on an indexed portfolio of Canadian equities have actually shrunk while real GDP has grown, producing a bigger gap between the series than we find in the United States. Why did this happen? In Canada, the fundamental nature of the economy has evolved in the past 30 years from resource driven to information and services driven.

The experience of the United Kingdom, where real earnings and real dividends grew materially slower than real GDP, has been similar to that of the United States. The experience of Japan has been rather more like Canada's. Japan, like Canada, is a fundamentally restructured economy. The result is that over the past 30 years, entrepreneurial capitalism in Japan has had a larger dilutive effect on shareholders in existing enterprises than it has in the United States.

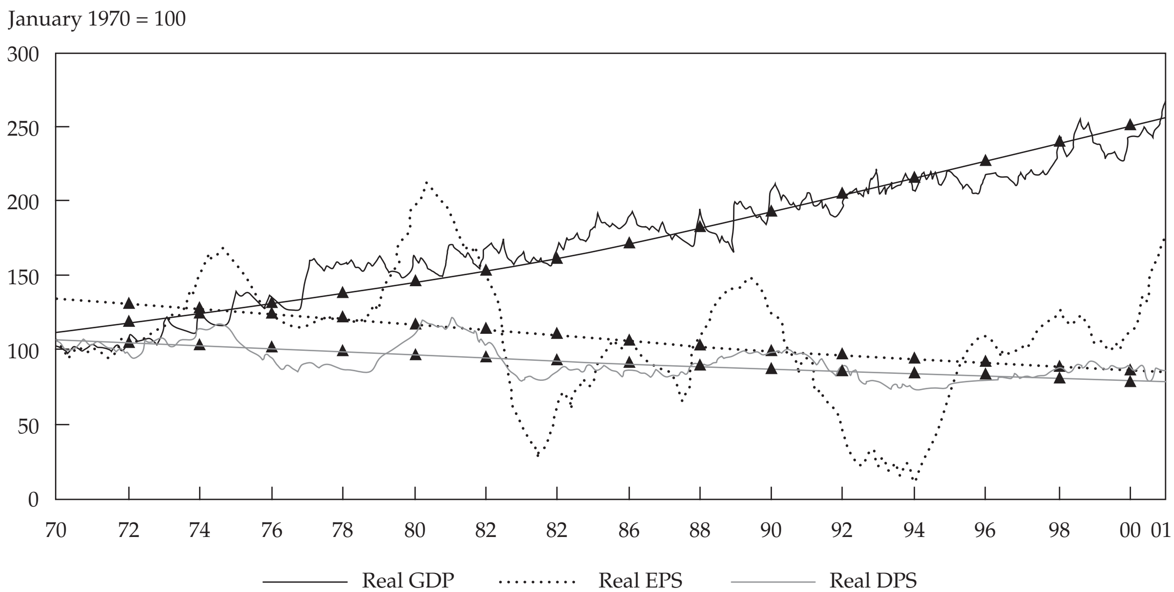
Table 2 shows, for the period from 1970 through 2000, the average growth of the four countries in real GDP, real EPS, real DPS, and average real EPS plus real DPS; Table 2 also shows the combined averages for each country and for all four countries grouped together. The general pattern is clear: Entrepreneurial capitalism is the dominant source of GDP growth, so it dilutes the growth of earnings for investors in existing enterprises.

We can look back over a much longer span for the U.S. market, from 1802 to 2001. **Figure 3** graphs the growth of $100 invested in U.S. stocks at the beginning of the 200-year period. Assuming dividends are reinvested, the $100 would have grown to more than $600 million by December 2001—a nice appreciation in any portfolio. By removing the effects of inflation and reinvestment of dividends, we can isolate the internal growth delivered by the existing companies. When the effect of inflation is removed, the ending value drops to $30 million. And when the assumption of reinvested dividends is removed, the ending value is reduced to a mere $2,000.

Figure 4 illustrates the link between real growth in stock value and economic growth. Real GDP growth increased 1,000-fold over the 1802–2001 period, real stock prices increased some 20-fold, and real per capita GDP growth similarly increased about 20-fold.

We can now assess the underlying engines of valuation. We'll examine the real dividend (you could do the same thing with real earnings). As **Figure 5**

Figure 2. GDP, EPS, and DPS: Canada, January 1970–January 2001



Note: Triangles identify exponentially fitted lines.
Source: OECD.

Table 2. Growth in GDP, EPS, DPS, and EPS + DPS, January 1970–January 2001

Measure	Canada	Japan	United Kingdom	United States	Average
Real GDP	2.7%	3.1%	2.4%	2.0%	2.5%
Real EPS	–1.4	–3.8	1.3	1.3	–0.6
Real DPS	–0.8	–1.6	2.0	1.0	0.1
Average real EPS + real DPS	–1.1	–2.7	1.6	1.1	–0.3
Average EPS + DPS growth as a percentage of GDP	–41.0	–87.0	67.0	57.0	–11.0

Source: OECD; Morgan Stanley Capital International.

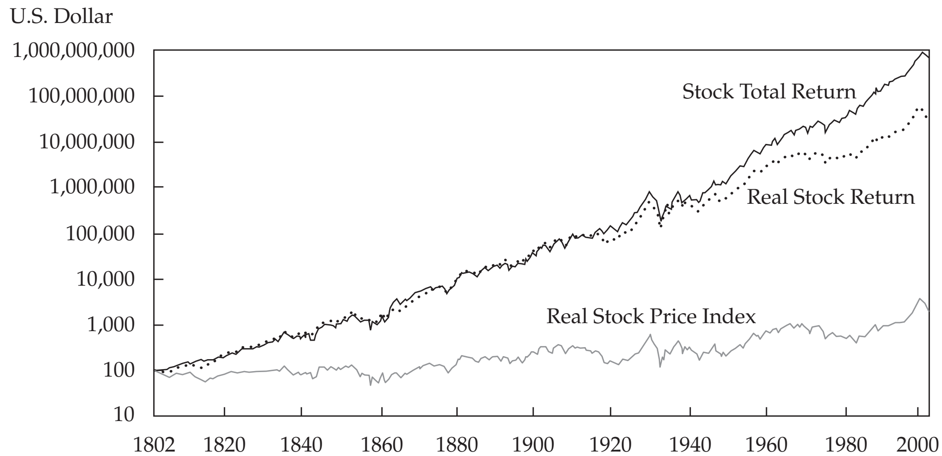
shows, real dividend growth matches very closely the growth in real per capita GDP. The implication is that the internal growth of a company is largely a matter of productivity growth in the economy and is, in fact, far slower than the conventional view—that dividends grow at the same rate as GDP.

Now we are ready to model and estimate real stock returns. In **Figure 6**, the dashed line represents the dilution of GDP growth in the growth of dividends. Growth in dividends tracks growth in real per capita GDP (the dotted line) remarkably tightly; the standard deviation is very modest—only 0.5 percent. This relationship is astonishingly stable. On a 40-year basis, the deviation is never above +0.1 percent and never below –1.6 percent. Moreover, current experience is in line with historical norms, despite anecdotal opinions that companies are delivering less in dividends than ever before.

A model that estimates real stock returns is useful only if its estimates actually fit subsequent experience. **Figure 7** is a scattergram providing the correlation between estimated and subsequent actual 10-year real stock returns. The correlation between the two is approximately 0.46 for the full period and far higher since World War II. The current figure for the real stock return is down in the 2–4 percent range. Of course, what the subsequent actual real return will be is anybody's guess, but I am not optimistic.

The same type of modeling can be done to estimate the real bond return. An inflation estimate can be subtracted from the nominal bond yield to arrive at an estimated real bond return. How do the

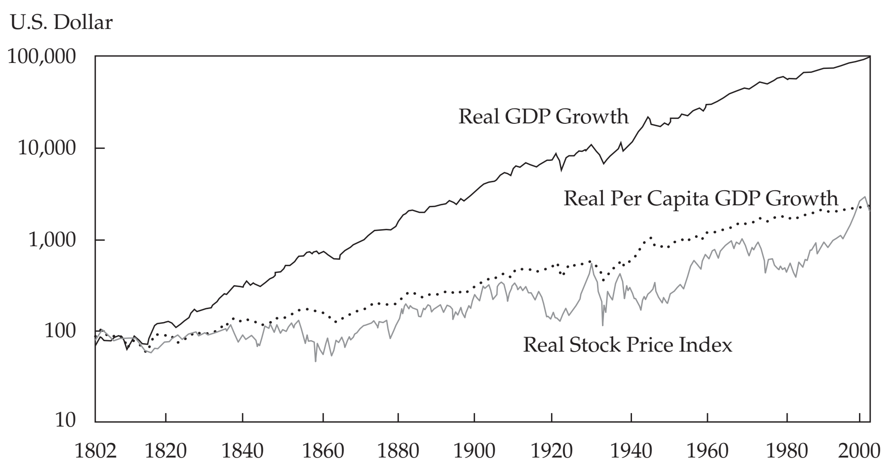
Figure 3. Return from Inflation and Dividends, 1802–2001



Notes: The "Real Stock Price Index" is the internal growth of real dividends—that is, the growth that an index fund would expect to see in its own real dividends in the absence of additional investments, such as reinvestment of dividends.

Source: Arnott and Bernstein (2002).

Figure 4. The Link between Stock Prices and Economic Growth, 1802–2001



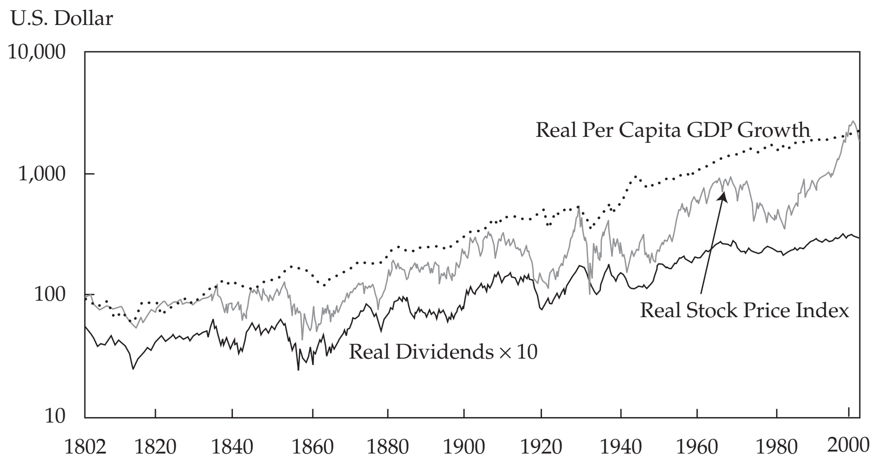
Source: Arnott and Bernstein (2002).

estimates calculated by this model fit with the subsequent real bond returns? As **Figure 8** shows, over a 200-year span, they fit pretty darned well. The loops off to the left relate to wartime. In several periods—the Civil War, World War I, World War II—investors were content to receive a negative expected real return for bonds, which can perhaps be attributed to patriotism. The country survived, so the real returns exceeded the expectations.

By taking the difference between the estimated real stock return and the estimated real bond yield, you get an objective estimate of what the forward-looking equity risk premium might have been for investors who chose to go through this sort of straightforward analysis at the various historical points in time. As shown in **Figure 9**, the *ex ante* risk premium of 5 percent, considered normal by many in the investment business, actually appears only during major wars, the Great Depression, and their aftermaths.

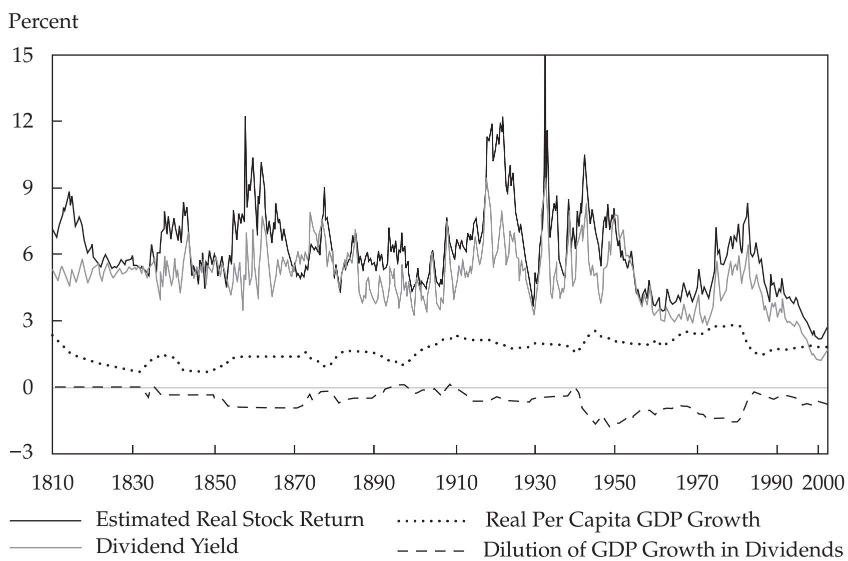
How good is the fit between this estimated risk premium and subsequent 10-year excess returns of

Figure 5. Dividends and Economic Growth, 1802–2001



Notes: Real dividends were multiplied by 10 to bring the line visually closer to the others; the result is that on those few occasions when the price line and dividend line touch, the dividend yield is 10 percent.
Source: Arnott and Bernstein (2002).

Figure 6. Estimating Real Stock Returns, 1810–2001

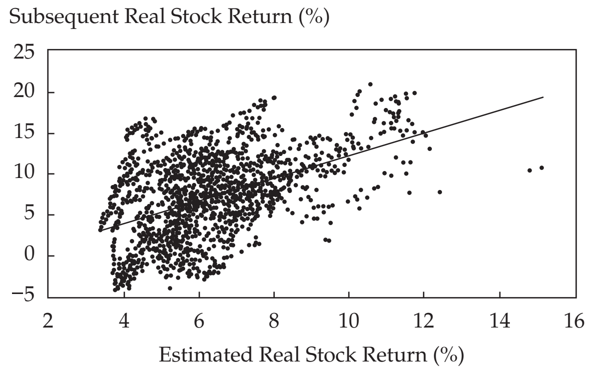


Notes: Based on rolling 40-year numbers. Real stock return = Dividend yield + Per capita GDP growth – Dividend/GDP dilution. The line "Dilution of GDP Growth in Dividends" indicates how much less rapidly dividends (and earnings) on existing enterprises can grow than the economy at large.
Source: Arnott and Bernstein (2002).

stocks over bonds? **Figure 10** shows that the fit is fairly good, which is worrisome in light of the poor current outlook. The current point on the *x*-axis (when this particular formulation is used) is about –0.5 percent. The implications for forward-looking 10-year real excess returns of stocks relative to bonds are worrisome—if this model holds in the future, if things are not truly different this time.

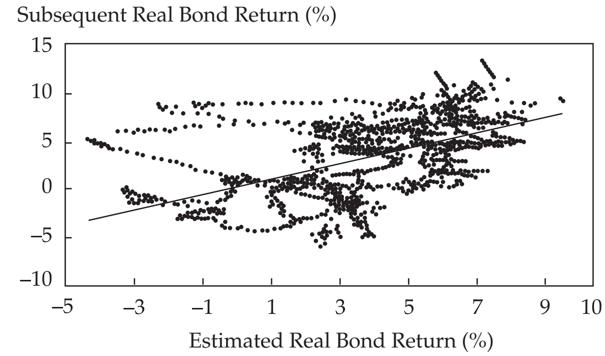
Figure 11 is a scattergram that relates the payout ratio to subsequent 10-year earnings growth from 1950 through 1991. This information ties in with Cliff Asness's talk [in the "Theoretical Foundations"

Figure 7. Estimated and Subsequent Actual Real Stock Returns, 1802–2001



Source: Arnott and Bernstein (2002).

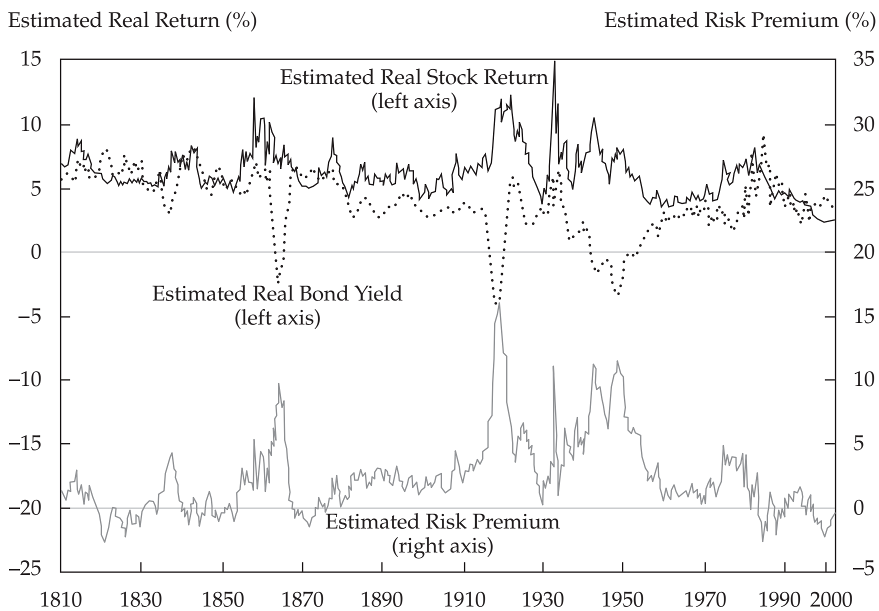
Figure 8. Estimated and Subsequent Actual Real Bond Yields, 1802–2001



Source: Arnott and Bernstein (2002).

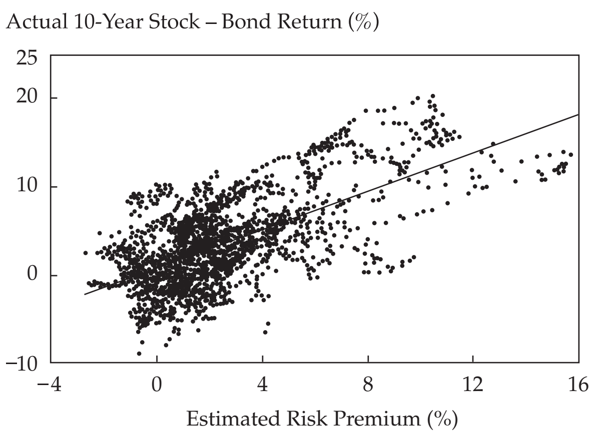
session]. Modigliani and Miller would suggest that if payout ratios are low (see Modigliani and Miller 1958), the reinvestment averaged across the market should produce the same market return that one could get by receiving those dividends and reinvesting them in the market. The tangible evidence is not encouraging. (Keep in mind that the M&M focus is cross-sectional, not intertemporal, so what I've just said is a variant of Modigliani and Miller's work, but it is a widely cited variant. M&M's work is frequently referred to in making the case that earnings growth is going to be faster than ever before.) Based on Figure 11, the correlation between payout ratios and subsequent 10-year earnings growth is a *negative* 0.60—which is worrisome. With recent payout ratios well below 40 percent, the implication for earnings growth is a rate of about –2 percent or worse, from the 2000 earnings peak, over the coming decade. If we combine an assumed negative earnings growth rate with an assumed zero risk premium, I believe that we have a serious problem.

Figure 9. Estimating the Equity Risk Premium, 1810–2001



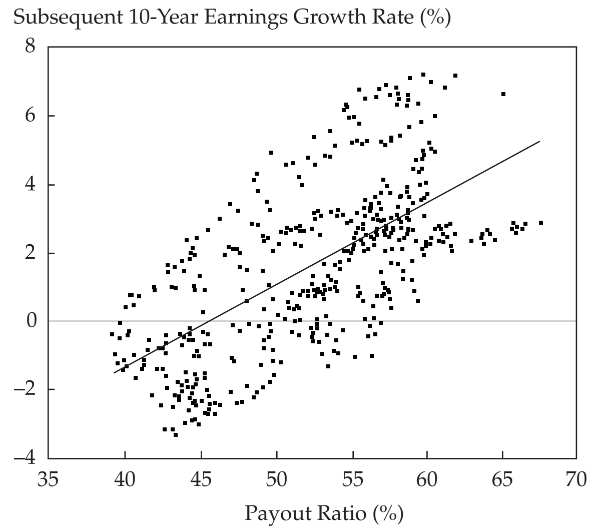
Source: Arnott and Bernstein (2002).

Figure 10. Risk Premium and Subsequent 10-Year Excess Stock Returns: Correlations, 1810–1991



Source: Arnott and Bernstein (2002).

Figure 11. Payout Ratio and Subsequent 10-Year Earnings Growth, 1950–91



Implications for Asset Allocation, Portfolio Management, and Future Research I

Robert D. Arnott
First Quadrant, L.P.
Pasadena, California

SUMMARY

by Peter Williamson
Amos Tuck School of Business Administration
Dartmouth College, Hanover, New Hampshire

Robert Arnott began with an emphasis on practice and empiricism, as opposed to theory. He urged the use of the terms "equity excess return" for the past and "equity risk premium" for the future.

We have seen a decline in bond yields. Does this decline portend an increase or a decrease in bond returns? And we have seen a decline in stock earnings yields (earnings to price). Does this decline portend an increase or decrease in stock returns? The participants in the Equity Risk Premium Forum would all, he believes, when shaping expectations, back out the portion of return attributable to the drop in earnings or dividend yield from the historical return. But he had not heard much discussion of the fact that a drop in earnings yield should have a second-stage impact—a haircut in expected returns accompanying the fall in earnings yield.

Arnott estimated an *ex ante* risk premium at the present time of zero. In this case, the old policy of balancing risk and return no longer works. Rebalancing used to recognize that more stock meant more risk and more return. So, fiduciaries gauged the risk tolerance of the investment committee and pushed the portfolio as far into stocks as that risk tolerance would permit. If the return expectations for stocks and bonds are similar, the policy asset allocation matters in terms of risk but not in terms of returns and the allocation decision is far less critical than it was in the past.

Strategic Implications

Historically, rebalancing has produced an alpha by reducing risk. Over long periods, it produced a little extra return. Now, with no risk premium, with any pattern of reversion to a mean for stocks and for bonds, rebalancing can boost returns.

Tactical asset allocation achieved episodic returns that conveyed a large alpha in the turbulent 1970s and 1980s but did not necessarily add value in the roaring bull market of the 1990s, although it could reduce risk. If the U.S. market is headed for a repeat of the 1970s, then TAA may be especially worthwhile in the near future.

What about strategic implications for pension funds? If conventional returns lag actuarial estimates, which is likely, then current funding ratios are misleading, contributions will have to catch up, and alpha matters. In a world of lower returns, an emphasis on such alternative investments as private equity may be appealing, but to the extent that this emphasis relies on a strong equity market for an exit strategy, it may not be so attractive. International stocks and bonds may be attractive, but the expected returns there will also be low. Rebalancing and cash equitization are worth a look. Uncorrelated alternatives such as TIPS, real estate, REITs (real estate investment trusts), and commodities will be promising.[1] Absolute return strategies may be seen as more important in inefficient markets. There will be increased searching for inefficiencies by active managers and increased searching for avoidance of negative alpha by those who believe in market efficiency.

Empirical Results

Turning from practice to empiricism, Arnott's **Table 1** showed the Ibbotson data together with the prospects based on our current situation. Starting with a dividend yield of 5.4 percent, the U.S. equity market has seen an approximately 8 percent compounded real return on stocks over the past 75 years. The change in the price/dividend valuation ratio added 1.7 percent, which should be backed out of the returns for forecasting purposes. Note that real dividends grew at a scant 1 percent. The initial real bond yield in 1925 was 3.7 percent, and because it

[1] TIPS are Treasury Inflation-Protected Securities; these securities are now called Treasury Inflation-Indexed Securities.

Table 1. The Ibbotson Data Revisited and Prospects for the Future

Component	75 Years Starting December 1925	Prospects from October 2001
Starting dividend yield	5.4%	1.7%
Growth in real dividends	1.0	2.0
Change in valuation levels[a]	1.7	???
Cumulative real return	8.1	±3.7
Less starting bond real yield	3.7[c]	3.4[d]
Less bond valuation change[b]	−0.4	???
Cumulative risk premium	4.7	±0.3

[a] Yields went from 5.4 percent to 1.4 percent, representing a 2.1 percent increase in the price/dividend valuation level.

[b] Bond yields went from 3.7 percent to 5.5 percent, representing a 0.3 percent annualized drop in long bond prices.

[c] A 3.7 percent yield, less an assumed 1926 inflation expectation of zero.

[d] The yield on U.S. government inflation-indexed bonds.

Source: Based on Ibbotson Associates (2001) data.

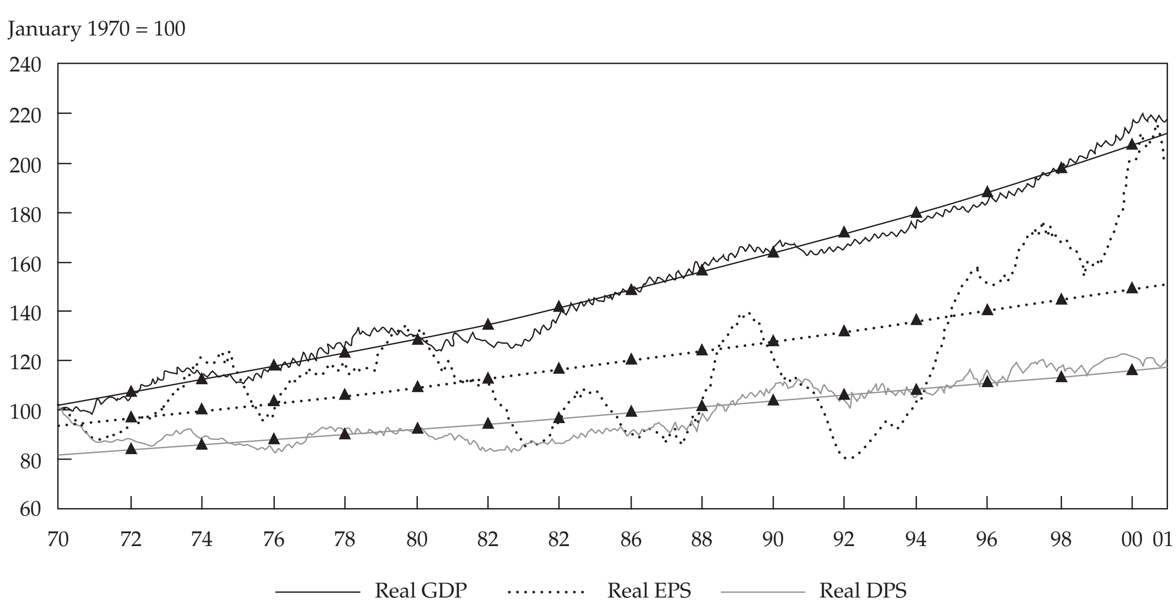
was the quoted bond yield, investors had no reason to expect that inflation would matter. So, the excess return of equities over bonds was close to 5 percent. Now, we are looking at a 1.7 percent starting dividend yield, roughly a 2 percent growth in real dividends, and probably no increase in valuation levels—for a total prospective real return of about 3.7 percent. Subtracting a 3.4 percent real bond yield (e.g., the TIPS yield) produces a 0.3 percent (30 bps) cumulative risk premium plus or minus some small standard deviation.

Why did dividends grow at only 1 percent in the past? Looking at the **Figure 1** graph of real GDP, real EPS, and real dividends per share (DPS), we can see that earnings have almost kept pace with GDP growth—but in the context of going from a small share of the national economy to a large share. Entrepreneurial capitalism dilutes the growth experienced by investors in existing enterprises. The trend in dividend growth is well below that of GDP. Over the period January 1970 to January 2001, real GDP growth was fairly steady. Real earnings growth and real dividend growth followed slower trends and were quite irregular, with relatively high earnings growth since about 1995. The relative growth in GDP, equity earnings, and dividends has been similar in the United Kingdom to that in the United States. In Canada and Japan, however, the trend in earnings and dividends has been down, not up, over the past 30 years.

Turning to the 200-year history beginning in 1802, Arnott's **Figure 3** indicated that $100 invested in stocks in 1802 would have grown, with dividends reinvested, to nearly $1 billion in 200 years.[2] In real

[2] Table and figure numbers in each Summary correspond to the table and figure numbers in the full presentation.

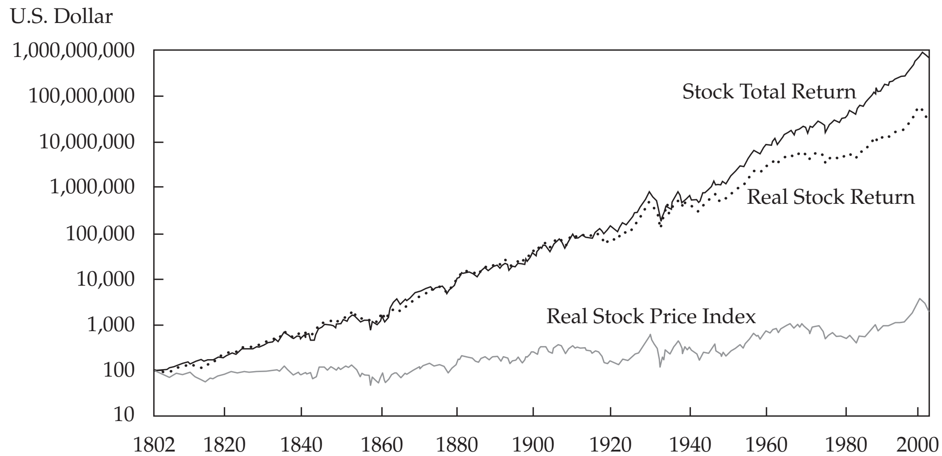
Figure 1. GDP, EPS, and DPS: United States, January 1970–January 2001



Note: Triangles identify exponentially fitted lines.

Source: Data from Organization for Economic Cooperation and Development (OECD).

Figure 3. Return from Inflation and Dividends, 1802–2001



Notes: The "Real Stock Price Index" is the internal growth of real dividends—that is, the growth that an index fund would expect to see in its own real dividends in the absence of additional investments, such as reinvestment of dividends.

Source: Arnott and Bernstein (2002).

terms, however, the ending amount is $30 million, and when we look at the index alone, without dividend reinvestment, the $100 rose barely above $1,000.

Real dividends have trailed per capita GDP growth. **Figure 4** indicated that, in this time frame, an index of real stock prices tracked real per capita GDP growth rather well in the United States, although the index persistently trailed aggregate GDP growth for the 200 years.

Figure 6 provided a basis for modeling and estimating real stock returns. Real per capita GDP growth and dilution of GDP growth in dividends are both remarkably stable and closely parallel. The note to Figure 6 provides Arnott's equation for estimating real stock returns. This equation can also be used for the more recent subperiod of 1950–2001 to forecast future real stock returns. A similarly simple model can be used to estimate future real bond returns.

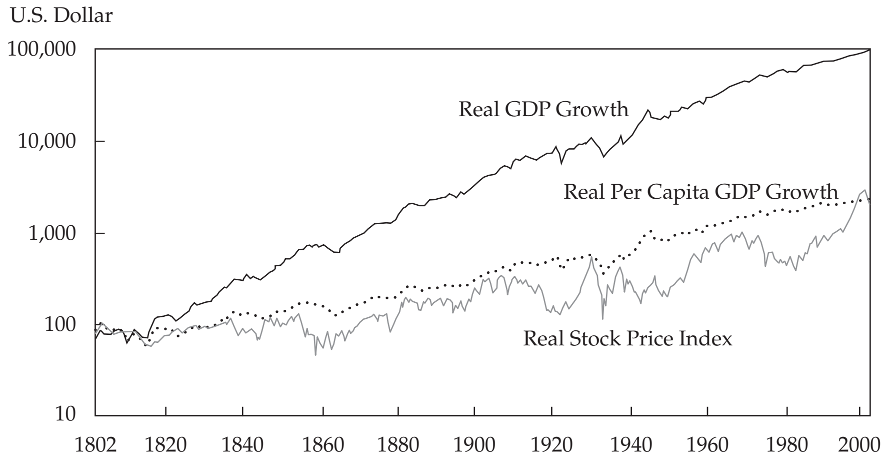
Figure 9 showed the results of using these simple models to estimate the real stock return, real bond yield, and equity risk premium (what might be called the "objective risk premium") year-by-year from 1810 to 2001. The risk premium rarely rose above 5 percent, only at the times of the Civil War, World War I, the Great Depression, and World War II. The premium is currently at or below zero.

During previous discussion of the Miller and Modigliani propositions, Arnott had commented that empirical evidence was not consistent with M&M. In this presentation, he showed the **Figure 11** plot of the payout ratio against subsequent 10-year earnings growth. Noting that M&M dealt with cross-sectional, not time-series, propositions and that he was showing time-series evidence, Arnott pointed out that high earnings retention (low payout) led *not* to higher earnings growth but to *lower* growth, a source of some concern.

Summary Implications

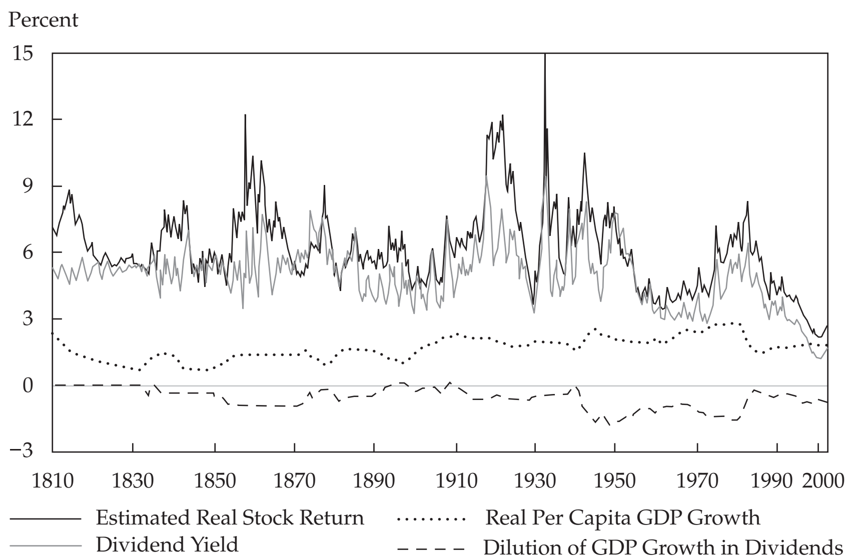
The implications of lower expected returns for policy allocation are as follows: In the past, the choice between stocks and fixed income was the essence of the policy asset-allocation decision. More stocks meant more risk and more return. For the future, with prospective stock and bond returns similar, policy allocation is no longer "king." If real earnings fall, as the empirical evidence on payout ratios suggests, or if valuation ratios "revert to the mean," then the situation is even worse.

Figure 4. The Link between Stock Prices and Economic Growth, 1802–2001



Source: Arnott and Bernstein (2002).

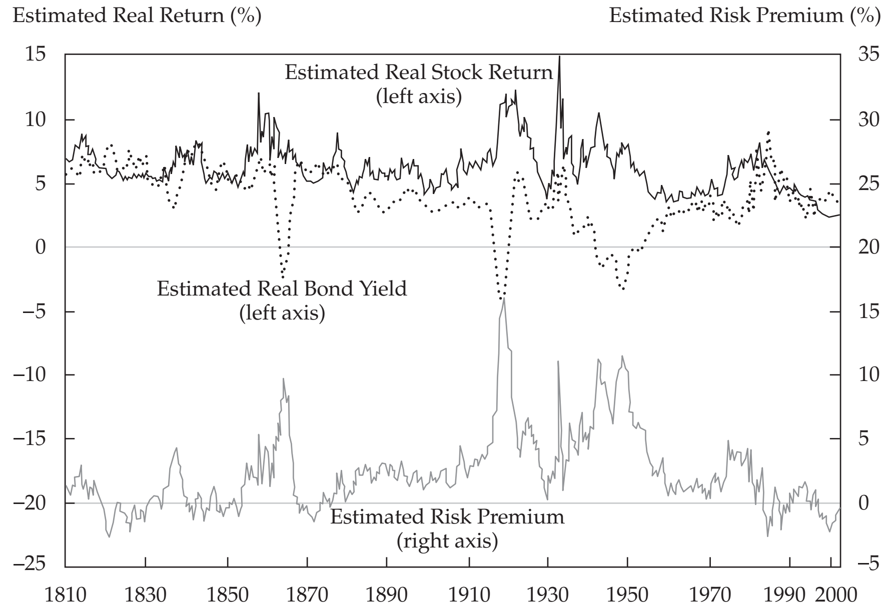
Figure 6. Estimating Real Stock Returns, 1810–2001



Notes: Based on rolling 40-year numbers. Real stock return = Dividend yield + Per capita GDP growth – Dividend/GDP dilution. The line "Dilution of GDP Growth in Dividends" indicates how much less rapidly dividends (and earnings) on existing enterprises can grow than the economy at large.

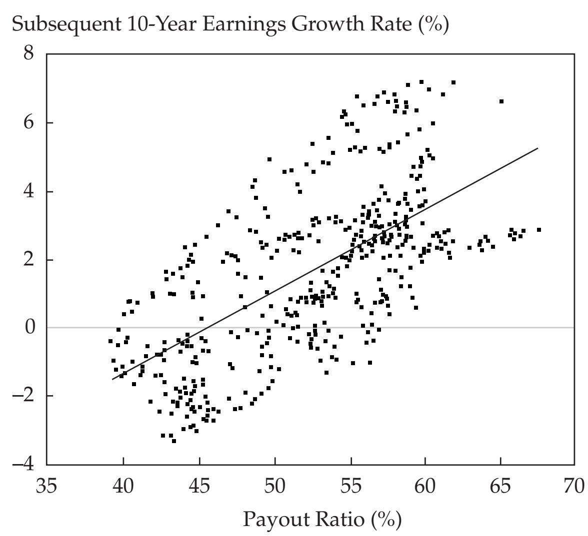
Source: Arnott and Bernstein (2002).

Figure 9. Estimating the Equity Risk Premium, 1810–2001



Source: Arnott and Bernstein (2002).

Figure 11. Payout Ratio and Subsequent 10-Year Earnings Growth, 1950–91



Implications for Asset Allocation, Portfolio Management, and Future Research II

Campbell Harvey

Duke University, Durham, North Carolina
National Bureau of Economic Research, Cambridge, Massachusetts

The reported survey of chief financial officers of U.S. corporations makes a unique contribution to the measurement of the expected equity risk premium and market volatility. Beginning with the second quarter of 2000, the research team has been conducting an ongoing, multiperiod survey of CFOs about their estimates of future equity risk premiums and equity market volatility. Results of the survey indicate the following: Return forecasts are positively influenced by past returns, which constitutes a type of "expectational momentum"; expected volatility is negatively related to past returns; the respondents seem to be very confident in their forecasts; and time horizon makes a big difference, in that a positive relationship was found between risk and expected return only for long-horizon forecasts.

After everything that has been said today, it is a challenge to make a unique contribution. We have heard how difficult it is to get a measure of expectations in terms of the equity risk premium, and what I am going to present is an approach to measuring expectations that is different from those that have been discussed.

For the past five years, John Graham and I, in conjunction with Financial Executives International, have been conducting a survey of chief financial officers of U.S. corporations about their estimates of future equity risk premiums and volatility.[1] Beginning in the second quarter of 2000 and, so far, extending into the third quarter of 2001, we have analyzed the more than 1,200 responses from the CFOs. Only 6 observations will appear in the graphs, but each observation is based on approximately 200 observations.

We know from other surveys that have been done that CFOs do actually think about the risk premium problem. We know that 75 percent of corporate financial executives—treasurers and CFOs—admit to using a CAPM-like or multifactor model. Therefore, we believe that the CFOs we are surveying are a reasonable sample of the population to question about the equity risk premium. I believe it is a sample group superior to that of economists surveyed—for example, by the Federal Reserve Bank of Philadelphia. The Philadelphia Fed's survey contains unreliable data (which I know from directly examining these data). I also think our survey has advantages over the survey of financial economists reported by Ivo Welch (2000) because our respondents are making real investment decisions. Finally, it is well known that the forecasts by financial analysts are biased. So, the survey we are conducting should provide some benefit in our search for *ex ante* risk premiums.

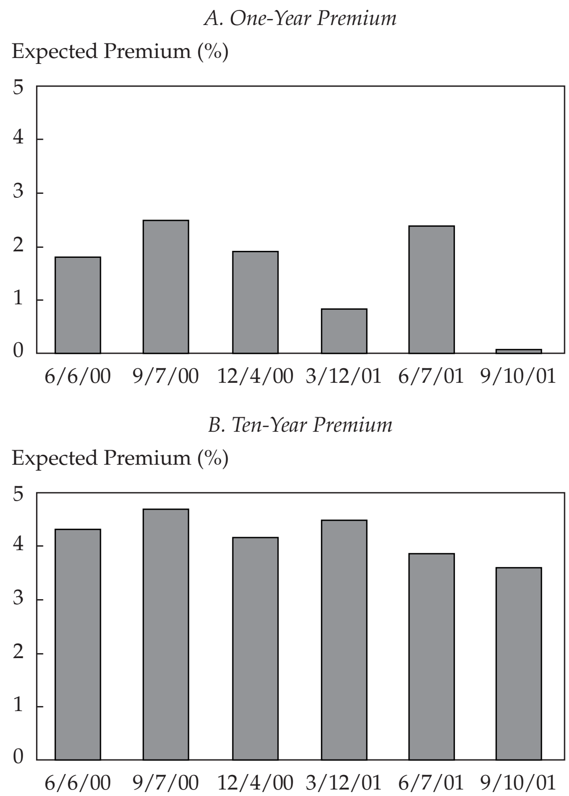
Survey of CFOs

Our survey has a number of components; it does not simply ask what the respondent thinks the risk premium is today. First, our survey is a multiperiod survey that shows us how the expectations of the risk premium change through time. Second, we ask about forecasts of the risk premium over different horizons. We have not talked much today about the effect of the investment horizon on the expected risk premium, but in our survey, we are asking about risk premium expectations for a 1-year horizon and a 10-year horizon. A third piece of information that we get in the survey is a measure of expected market volatility. Finally, we can recover from the responses a measure of the asymmetry or skewness in the distribution of the risk premium estimates.

[1] For a complete description of the study reported here, see Graham and Harvey (2001a).

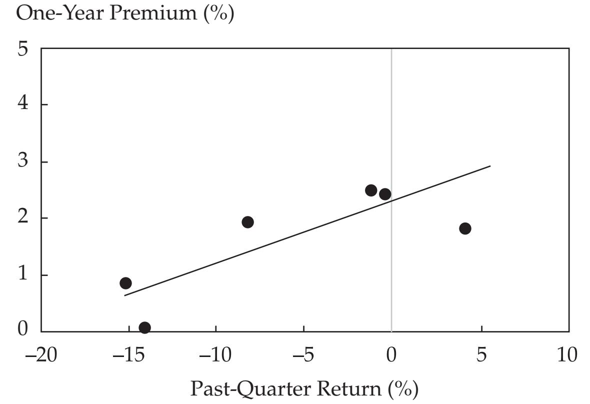
The first result I want to show you is striking. Panel A of **Figure 1** indicates that the CFOs' one-year *ex ante* risk premiums (framed in the survey as the excess return of stocks over U.S. T-bills) vary considerably over time. The last survey, finished on September 10, 2001, indicates the CFOs were forecasting at that time a one-year-ahead risk premium of, effectively, zero. The 10-year-horizon *ex ante* risk premium, given in Panel B, is interesting because it is higher than the 1-year-horizon forecast and is stable from survey to survey at about 4 percent (400 bps). Note that the September 10, 2001, forecast is 3.6 percent.

Figure 1. Survey Respondents' One-Year and Ten-Year Risk Premium Expectations



One of the first aspects we investigate is whether the CFOs' expectations about future returns are influenced by past returns. That is, if the market has performed poorly in the immediate past, does this performance lead to lower expected returns? **Figure 2** is a simple plot of the expected one-year equity risk premium against the previous quarter's return. (As we go through the analysis, please keep in mind that one can really be fooled by having so few observations. Indeed, this problem is exactly the reason we chose to present most of the results graphically. By eyeballing the data, you can see whether one observation is driving the relationship.) Figure 2 shows a fairly reliable positive relationship between past return and future near-term expected risk premium. Also, we found that you can pull out any of these observations and the fit is still similar. Apparently, a one-year-horizon forecast carries what Graham and I call "expectational momentum." Therefore, negative returns influence respondents to lower their forecast of the short-term future premium.

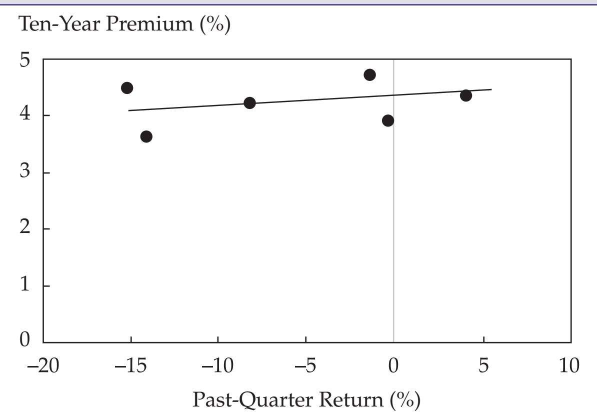
Figure 2. One-Year Risk Premium and Recent Returns



Notes: $y = 0.1096x + 2.3068$; $R^2 = 0.7141$.

Figure 3 plots the same variables for the 10-year horizon. There is a slight positive relationship between the past quarter's return and the *ex ante* 10-year-horizon risk premium, but it is not nearly as positive as the relationship observed for the 1-year horizon.

Figure 3. Ten-Year Risk Premium and Recent Returns

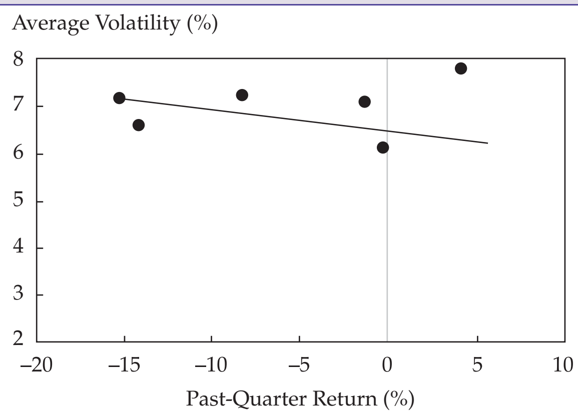


Notes: $y = 0.0179x + 4.3469$; $R^2 = 0.1529$.

We measured expected market volatility by deducing each respondent's probability distribution. We asked the respondents to provide a high and a low forecast by finishing two sentences: "During the next year, there is a 1-in-10 chance the S&P 500 return will be *higher* than ______ percent" and "During the next year, there is a 1-in-10 chance the S&P 500 return will be *lower* than _______ percent." The expected market volatility is a combination of the average of the individual expected volatilities (which I will refer to in the figures as "average volatility") plus the dispersion of the risk premium forecasts (referred to as "disagreement").[2]

Figure 4 shows that (annualized) average expected volatility for the one-year horizon is weakly negatively related to the past quarter's return. In fact, if one observation were pulled out, we might find no relationship whatsoever. And **Figure 5** shows the (annualized) disagreement component—basically, the standard deviation of the risk premium forecast—for the one-year horizon. The disagreement component for the one-year horizon is strongly related to the past quarter's return. A bad past return suggests a higher disagreement volatility. Even with so few data points, this relationship appears to be strong.

Figure 4. Average (One-Year-Horizon) Volatility and Recent Returns



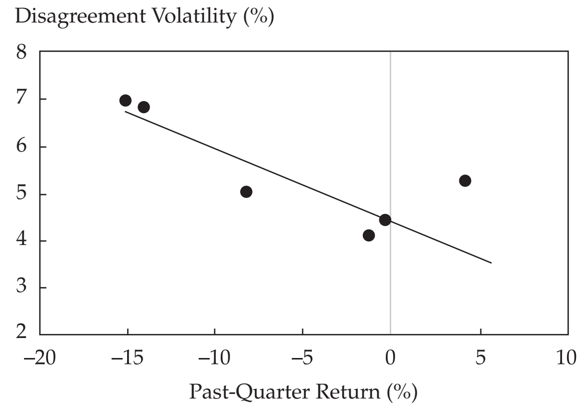
Notes: $y = -0.0452x + 6.4722$; $R^2 = 0.1282$.

One thing to keep in mind is that these points on Figures 4 and 5 are annualized. When you examine the individual volatilities, you find that these respondents are extremely confident in their assessments. The result is a 6–7 percent annualized volatility in the one-year-horizon *ex ante* risk premium. This volatility is much smaller than typical market estimates, such as the Chicago Board Options Exchange VIX (Volatility Index) number on the S&P 100 option, which averages around 20 percent.

We also found that our measure of asymmetry is positively related to the past quarter's return. Given that we get the tails of the distribution, we can look at the mass above and below the mean and compare them, which gives us an *ex ante* measure of skewness. If past returns are negative, we find more negative *ex ante* skewness in the data.

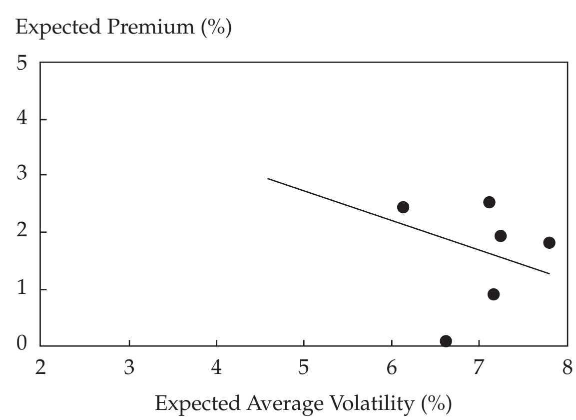
Instead of looking at the relationship of the forecasted risk premium to past return, **Figure 6** relates the forecasted (*ex ante*) risk premium to expected (*ex ante*) volatility. Many papers in academic finance have examined the relationship between expected risk and expected reward. Intuitively, one would expect the

Figure 5. Disagreement (One-Year Horizon) Volatility and Recent Returns



Notes: $y = -0.153x + 4.3658$; $R^2 = 0.7298$.

Figure 6. Expected Average Volatility and Expected Risk Premium: One-Year Horizon



Notes: $y = -0.5178x + 5.2945$; $R^2 = 0.2538$.

[2] Market volatility was measured as

$$\mathrm{var}[r] = E[\mathrm{var}(r|Z)] + \mathrm{var}[E(r|Z)],$$

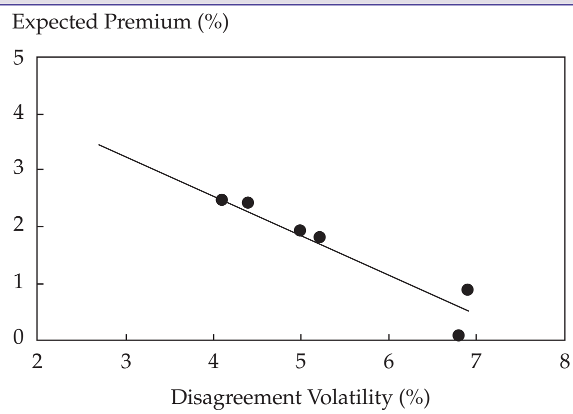
where r is the market return, Z is the information that the CFOs are using to form their forecasts, $[E(r|Z)]$ is the expected risk premium conditional on the CFO's information, $E[\mathrm{var}(r|Z)]$ is the average of each CFO's individual volatility estimate, and $\mathrm{var}[E(r|Z)]$ is disagreement volatility or the variance of the CFOs' forecasts of the premium. Individual volatilities were measured as

$$\mathrm{var} = \left[\frac{x(0.90) - x(0.10)}{2.65}\right]^2,$$

where $x(0.90)$ is the "one in ten chance that the return will be higher than" and $x(0.10)$ is the "one in ten chance that the return will be lower than." The equation for individual volatilities is from Davidson and Cooper (1976).

relationship to be positive, but the literature is actually split. Indeed, many papers have documented a negative relationship, which is basically what we see for the one-year-horizon predictions. In Figure 6, the *ex ante* premium and the *ex ante* average volatility appear to be weakly negatively related. **Figure 7** plots the one-year-horizon expected risk premium against disagreement about the expected premium. The result is a strongly negative relationship: The higher the disagreement, the lower the expected premium over one year. Again, almost any observation could be pulled out without changing the degree of fit.

Figure 7. Disagreement Volatility and Expected Risk Premium: One-Year Horizon

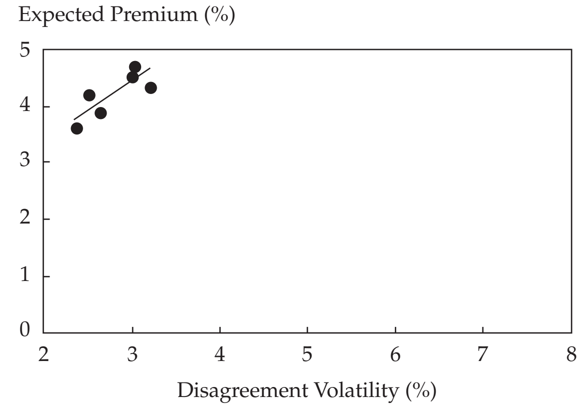


Notes: $y = -0.6977x + 5.3410$; $R^2 = 0.9283$.

Using the same variables as in Figure 7 and keeping the scale the same, **Figure 8** shows the data for the 10-year horizon. The fit is again strikingly good, but the relationship is positive. Notice that the disagreement is much smaller for the 10-year horizon than for the 1-year horizon. This positive relationship between the *ex ante* premium and *ex ante* volatility is suggested by basic asset-pricing theory.

The latest survey documented in Figures 2–8 is June 1, 2001, plus data returned to us by September 10, 2001. We just happened to fax our most recent quarterly survey to the survey participants at 8:00 a.m. on the morning of September 10. I did not include observations from the surveys returned on September 11 because the survey might have been completed on either September 10 or 11, and classification of the responses as pre- or post-September 11 was not possible. The response data we received on September 12 or later we maintained and analyzed separately. **Table 1** provides a comparison of pre- and post-September 11 data for the 1- and 10-year horizons. Although the size of the sample is small (33 observations), one can see the impact of September

Figure 8. Disagreement Volatility and Expected Risk Premium: Ten-Year Horizon



Notes: $y = 0.9949x + 1.4616$; $R^2 = 0.6679$.

11. The 1-year-horizon mean forecasted premium decreases after September 11, but volatility—both disagreement and average—increases. For the 10-year horizon, the mean forecasted premium and disagreement volatility increase. I'll be the first to admit that these results are not statistically significant, but the data tell an interesting story. After September 11, perceived risk increases—which is no surprise. In the short term, participants believe that market returns will be lower. In the long term, however, premiums increase to compensate for this additional risk.

Table 1. Impact of September 11, 2001: Equity Risk Premium and Volatility

Measure	Before	After
Observations	127	33
1-year premium		
Mean premium	0.05 %	–0.70 %
Average volatility	6.79	9.76
Disagreement volatility	6.61	7.86
10-year premium		
Mean premium	3.63 %	4.82 %
Disagreement volatility	2.36	3.03

Implications of Results

So, what have we learned from this exercise? First, expectations are affected, at least in the short term, by what has happened in the recent past—an expectational momentum effect. Second, these new expectational data appear to validate the so-called leverage effect—that negative returns increase expected volatility. Third, the individual volatilities (at 6–7 percent) seem very low, given what we would have expected. And fourth, there is apparently a positive relationship between risk and expected return (or the risk premium) only at longer horizons. So, the horizon is critical.

How should we interpret these results, what are the outstanding issues, and where do we go from here? The CFOs in the survey are probably not using their one-year expected risk premiums for one-year project evaluations. What CFOs think is going to happen in the market is different from what they use as the hurdle rate for an investment. I do think that the 10-year-horizon risk premium estimates we are getting from them are close to what they are using. An interesting paper being circulating by Ravi Jagannathan and Iwan Meier (2001) makes some of these same arguments—that higher hurdle rates are probably being used for a number of reasons: the scarcity of management time, the desire to wait for the best projects, and financial flexibility. Corporate managers want to wait for the best project, and with limited management time, a hurdle rate that is higher than what would be implied by a simple asset-pricing model allows that time.

Another angle is that the premium *should* be high in times of recession. Indeed, a lot of research documents apparently countercyclical behavior in the premium. Such behavior implies that today's one-year-horizon investment should have a high hurdle rate.

Further Research

We hope our research sheds some light on the measure of expectations. I believe in asset-pricing models based on fundamentals, but it is also enlightening to observe a direct measure of expectations. Our data may not be the true expectations, but they supply additional information about the *ex ante* risk premium in terms of investment horizon, expected volatility, and asymmetry.

Our next step is to conduct interviews in the first week of December 2001 with a number of the CFOs participating in the multiperiod survey. We have already carried out a few preliminary interviews, and we find it extraordinary how much thought CFOs have given to these issues. The main question we want to ask in December is the reason (or reasons) for the difference between their risk premium forecasts for a one-year horizon and the actual internal hurdle rates they use to evaluate one-year-horizon projects. How do CFOs use the *ex ante* risk premium in terms of making real allocation decisions? I will keep you updated on the progress of our research project.

Implications for Asset Allocation, Portfolio Management, and Future Research II

Campbell Harvey

Duke University, Durham, North Carolina
National Bureau of Economic Research, Cambridge, Massachusetts

SUMMARY

by Peter Williamson

Amos Tuck School of Business Administration
Dartmouth College, Hanover, New Hampshire

The presentation made by Campbell Harvey was unique, in that it was based essentially on surveys of investor expected risk premiums. What he had heard from the previous speakers was how difficult it is to get a measure of investor expectations.

Harvey's surveys, over time, of chief financial officers offered what he considered to be a less biased sample than the surveys that have been made of economists or financial analysts. CFOs are known to be concerned about a measure of their cost of capital for investment planning purposes and have no reason to favor high or low forecasts. He stated that, although he does not see the survey results as a replacement for the kind of analyses presented by previous speakers, he does believe that the surveys add valuable information.

The survey questions and responses were for 1-year and 10-year time horizons, which provided an opportunity to compare short-term with long-term expectations. The surveys elicited information not only on the expected premiums but also on the probability distributions of the respondents' forecasts. Harvey considered two components of expected market volatility: the average of the individual expected volatilities (from each individual's probability distribution) and the disagreement over the risk premium forecasts (the standard deviation of the risk premium forecasts).

Figure 1 shows the results of six surveys asking for a 1-year risk premium estimate and a 10-year estimate. The 10-year forecasts show little variation, whereas the 1-year forecasts vary widely through time. The 10-year forecasts are also consistently higher than the 1-year forecasts.

Figure 1. Survey Respondents' One-Year and Ten-Year Risk Premium Expectations

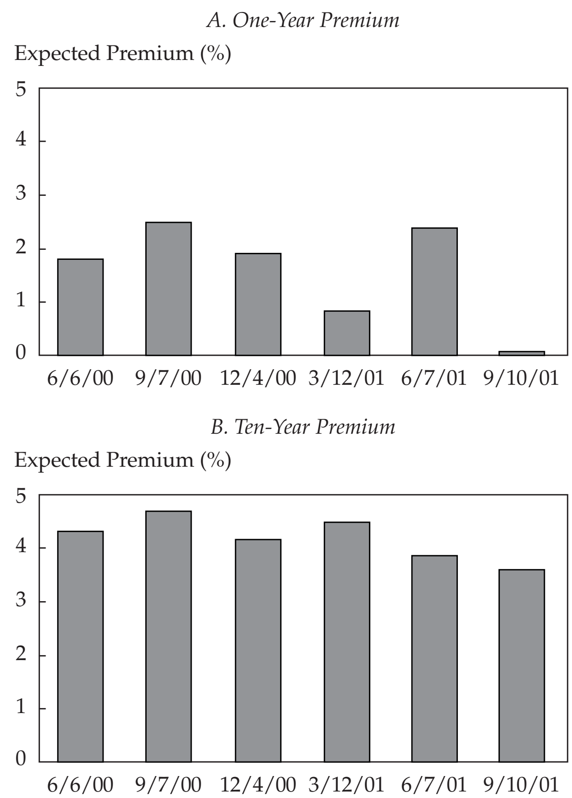
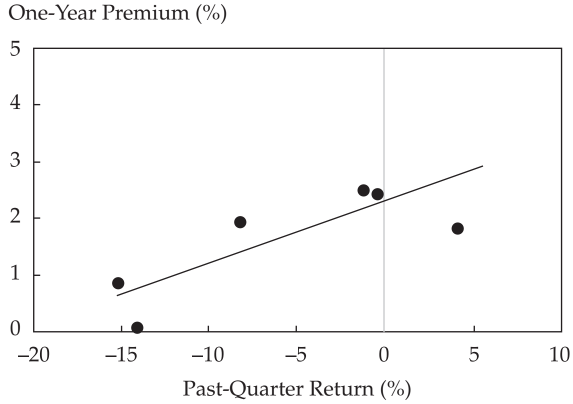


Figure 2 shows the influence of past returns on forecasts of 1-year premiums, and **Figure 3** does the same for 10-year premiums. Past returns had a positive impact on 1-year forecasts and a very slight positive effect on 10-year forecasts. Past returns also had a weak negative effect on expected 1-year average volatility and a strong negative effect on disagreement. They had a strong positive effect on expected skewness. Negative returns led to more negative skewness in the forecasts.

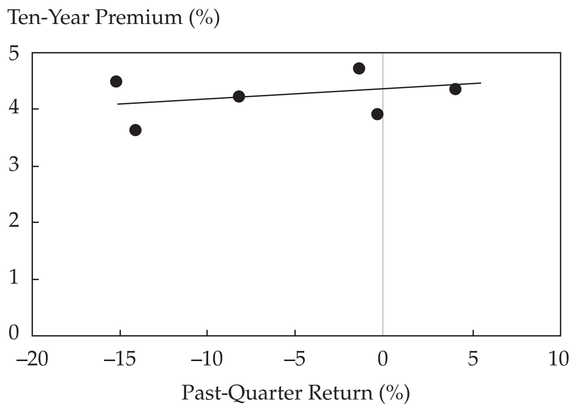
Turning to the effect of expected rather than past returns, Harvey showed in **Figure 6** that the average

Figure 2. One-Year Risk Premium and Recent Returns



Notes: $y = 0.1096x + 2.3068$; $R^2 = 0.7141$.

Figure 3. Ten-Year Risk Premium and Recent Returns



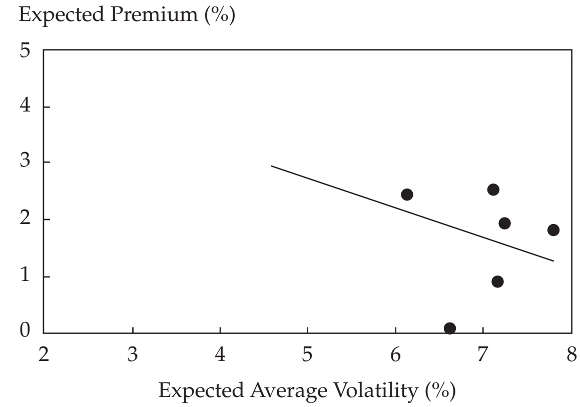
Notes: $y = 0.0179x + 4.3469$; $R^2 = 0.1529$.

of individual volatilities is weakly *negatively* related to expected 1-year returns.[1] One-year expected returns were found to be strongly negatively related to disagreement volatility, as shown in **Figure 7**. This finding may seem counter to the usual risk–expected return theories, but the finding is for very short term forecasts. For the 10-year horizon shown in **Figure 8**, however, expected returns are strongly *positively* related to disagreement—which is consistent with the way we usually think about risk and expected reward.

Harvey reported the impact of the events of September 11, 2001, in **Table 1**. After the crisis, the CFOs revised expected returns for the 1-year forecasts downward. For both the 1-year and the 10-year forecasts, expected volatility increased after the crisis.

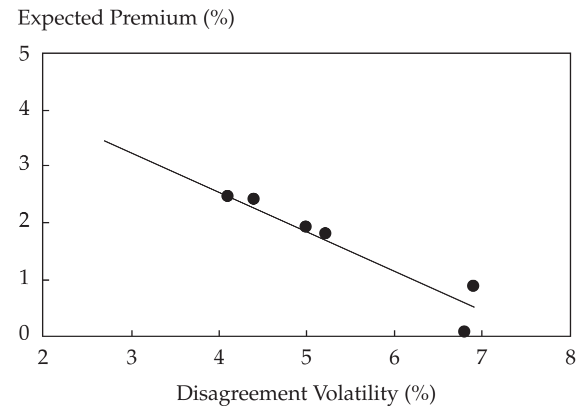
[1] Table and figure numbers in each Summary correspond to the table and figure numbers in the full presentation.

Figure 6. Expected Average Volatility and Expected Risk Premium: One-Year Horizon



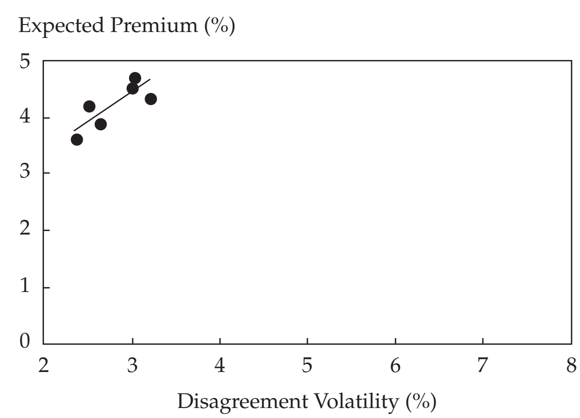
Notes: $y = -0.5178x + 5.2945$; $R^2 = 0.2538$.

Figure 7. Disagreement Volatility and Expected Risk Premium: One-Year Horizon



Notes: $y = -0.6977x + 5.3410$; $R^2 = 0.9283$.

Figure 8. Disagreement Volatility and Expected Risk Premium: Ten-Year Horizon



Notes: $y = 0.9949x + 1.4616$; $R^2 = 0.6679$.

Table 1. Impact of September 11, 2001: Equity Risk Premium and Volatility

Measure	Before	After
Observations	127	33
1-year premium		
Mean premium	0.05%	–0.70%
Average volatility	6.79	9.76
Disagreement volatility	6.61	7.86
10-year premium		
Mean premium	3.63%	4.82%
Disagreement volatility	2.36	3.03

Summarizing, Harvey presented the following conclusions:

- Survey measures of expectations provide useful alternatives to statistical measurements.
- Return forecasts are positively influenced by past returns—what John Graham and Harvey (2001a) call "expectational momentum."
- Expected volatility is negatively related to past returns.
- Individual volatilities seem very low; the respondents seem very confident in their forecasts.
- Time horizon makes a big difference. There is a positive relationship between risk and expected return but only for long-horizon forecasts.

In closing, Harvey expressed doubt that the CFOs were actually using their 1-year forecasts for hurdle rates in 1-year project evaluations. He suggested that there is a difference between what CFOs believe will happen to the market next year and the rate of return they would accept for a new project. The 10-year forecasts are probably closer to what the CFOs are using for the cost of capital.

Implications for Asset Allocation, Portfolio Management, and Future Research: Discussion

Roger Ibbotson (*Moderator*)
Robert Arnott
John Campbell
Bradford Cornell
William Goetzmann
Campbell Harvey
Martin Leibowitz
Thomas Philips
William Reichenstein, CFA

ROGER IBBOTSON (Moderator)

I was particularly pleased to see Campbell Harvey's paper because we have seen surveys of financial analysts, individuals, and economists (such as Welch's 2000 survey of financial economists), but the Graham and Harvey (2001a, 2001b) survey breaks new ground by surveying a particularly astute group. The results of their survey bring fresh information to the table. The survey was also well designed, which gives us confidence in the data.

I think each of us understands that we are concerned with equity risk premiums looking forward, but the distance we are looking ahead, our horizons, may differ. And today we have had both discussions—looking short term and looking out long term. The differences between the short-run and the long-run risk premium were certainly brought out by Rajnish Mehra [in the "Current Estimates and Prospects for Change" session] and are highlighted in the Graham and Harvey work.

I would like to present a few ideas from a paper that Peng Chen and I wrote (Ibbotson and Chen 2002) that uses much of the same data that Rob Arnott used but interprets the data almost completely differently. One of the reasons for the lack of overlap in interpretations is that Rob's primary focus is a short-run prediction of the market.

Figure 1 is yet another P/E chart—this one based on the Wilson and Jones (forthcoming 2002) data because their earnings data match the S&P 500 Index earnings data. The S&P 500 had very low, not negative but very low, earnings in the 1930s, and the actual maximum P/E is off the chart for that period. Figure 1 begins with a P/E, calculated as price divided by prior-year earnings, of 10.22 in 1926 and ends with a P/E of 25.96 at year-end 2000 (the October 2001 P/E, excluding extraordinary earnings, is 21); that growth from about 10 to the most recent P/E is an important consideration in the forecast I will discuss.

The forecast that Peng and I are making is based on the real drivers of P/E growth. We focus on the contribution of earnings to P/E growth and on GDP. **Table 1** shows the historical average nominal return for stocks over the 75-year period of 1926 through 2000 to be 10.70 percent. We can break that nominal stock return into its contributing components: about 3 percentage points (pps) inflation, and so forth. The P/E growth rate from a multiple of about 10 in 1926 to a multiple of almost 26 in 2000 amounts to 1.25 percent a year. When we make our forecasts, we remove that historical growth rate because that P/E jump from 10 to 26, in our opinion, will not be repeated. The "Earnings Forecast" column in Table 1 shows what history was without the P/E growth rate; that is, the forecasted return is 1.25 pps less than the historical return.

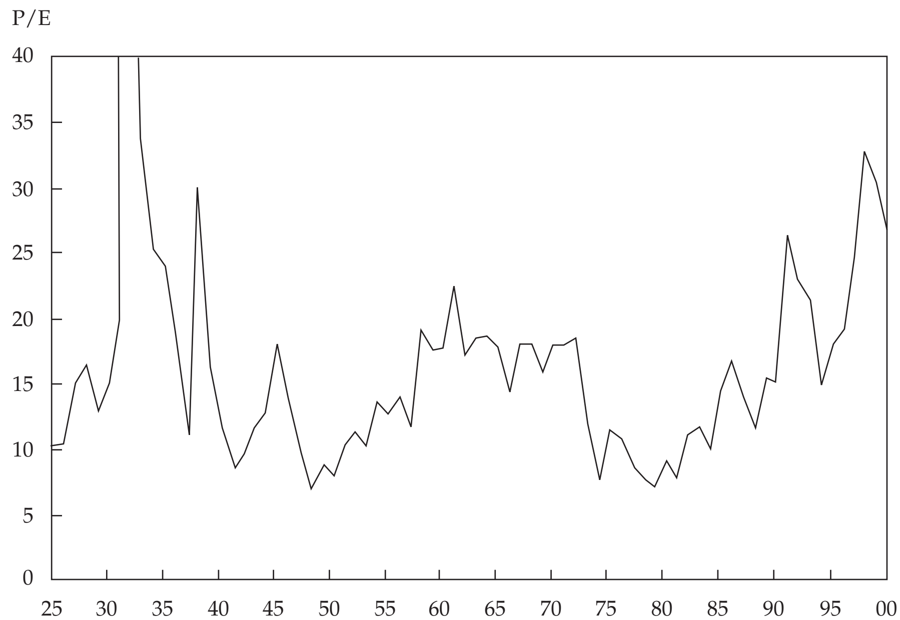
Table 1. Historical and Forecasted Components of Stock Returns, 1926–2000

Component	Historical[a]	Earnings Forecast
Income	4.28 pps	4.28 pps
P/E growth	1.25	—
Earnings growth	1.75	1.75
Inflation	3.08	3.08

[a]Total historical return for the period is 10.70 percent; data do not sum to that total because of the geometrical mathematics used.

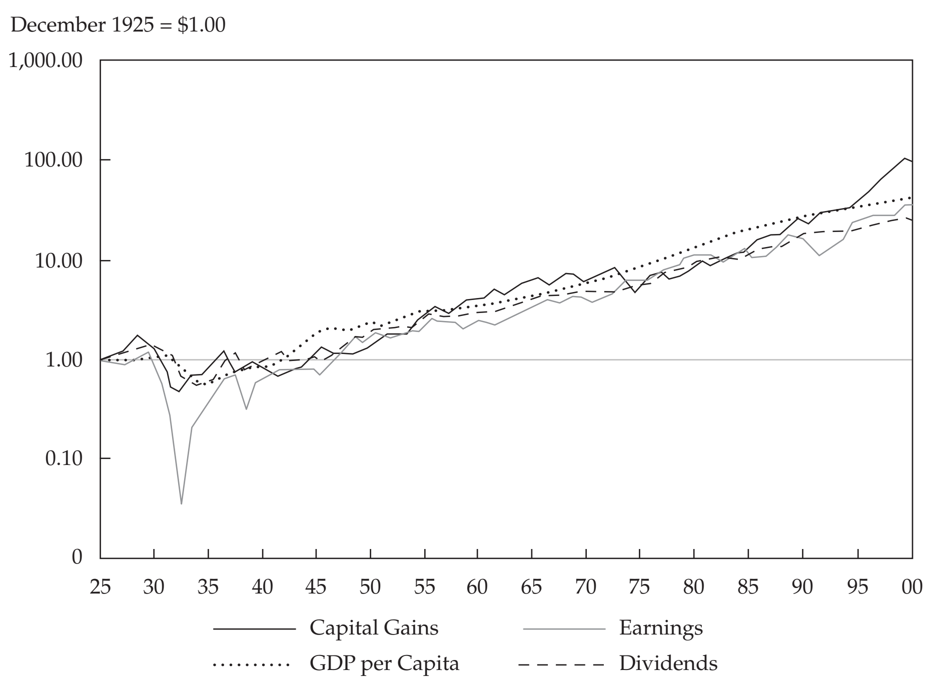
Figure 2 provides the historical growth of per capita GDP and of earnings, dividends, and capital gains on a per share, not aggregated, basis. All are indexed to $1 at the end of 1925. The capital gains grow to about $90 at the end of 2000—the most growth of any of the measures shown. Earnings are less because of the increase in the P/E multiple. The $90 is the $36 multiplied by 2.5, which was the P/E

Figure 1. The P/E, December 1925–December 2000



Note: The P/E for December 1932 was 136.5.

Figure 2. Historical Growth of per Capita GDP and of per Share Earnings, Dividends, and Capital Gains, December 1925–December 2000



Note: At end date, capital gains were $90.50, GDP per capita was $44.10, earnings were $35.60, and dividends were $24.20.

change from 10 to 26. The line for GDP per capita shows that the economy (on a per capita basis) has outgrown earnings by a small amount over the entire period. And finally, the growth in dividends trails the pack. So, I very much agree with the comment that Bill Reichenstein made earlier today that dividends are not a good forecasting tool; they grow the most slowly and even distort the picture for earnings growth [see "Current Estimates and Prospects for Change: Discussion"].

I am struck by how tied together each data series is—how the stock market is related to the economy, which is related to earnings, which are related to dividends. Although the link between earnings and dividends is a little less close than the other links, it is still there. One of the reasons Peng and I wanted to carry out this type of analysis is that the economy *should* be reflected in the stock market. And in fact, the separation in their behaviors is solely the result of the changing P/E, which we have thus removed from our forecasts. The P/E rose from 1926 to 2000 for a reason, but that reason will not continually recur in perpetuity. For that annual growth rate in the P/E multiple of 1.25 percent a year to continue, to assume that it will replicate, would mean that in another 75 years, the P/E will have grown to 62.

Figure 3 shows why dividends are not a good tool for forecasting the future. Dividend yields started the period at 5.15 percent and averaged 4.28 percent over the past 75 years; if you include the data for the 19th century, the historical average dividend yield is much higher. Every time we found a dividend for the 19th century, it seemed to be 100 percent. The dividend yield has now dropped to 1.10 percent (the most recent year would push it up somewhat). Thus, a long-run secular decline has occurred in the dividend yield, which was largely caused by the decreasing payout ratio. As **Figure 4** shows, the payout ratio, which began the period at 46.68 percent and averaged almost 60 percent over the 1926–2000 period, is now 31.78 percent.

Several reasons could explain the trend toward lower payout ratios. We interpret the trend as an issue of trust and changing attitudes about trust. As investors place more trust in the companies in which they invest and in the financial market system, shareholders no longer require that the companies pay all of their earnings to the shareholders; the discipline that dividends were designed to impose on corporations is gradually falling by the wayside. Another possible reason for the trend toward lower payout ratios is that, of course, dividends and capital gains (the fruit of reinvested corporate earnings) are taxed differently—providing an incentive for shareholders to relax their desire for company earnings to be paid out as dividends. Moreover, today, earnings can be taken out in many forms, such as share repurchases, buyouts in a merger or acquisition, or investment in internal projects of a company. I predict that these myriad forms of paying out earnings will remain. A

Figure 3. Dividend Yield, December 1925–December 2000

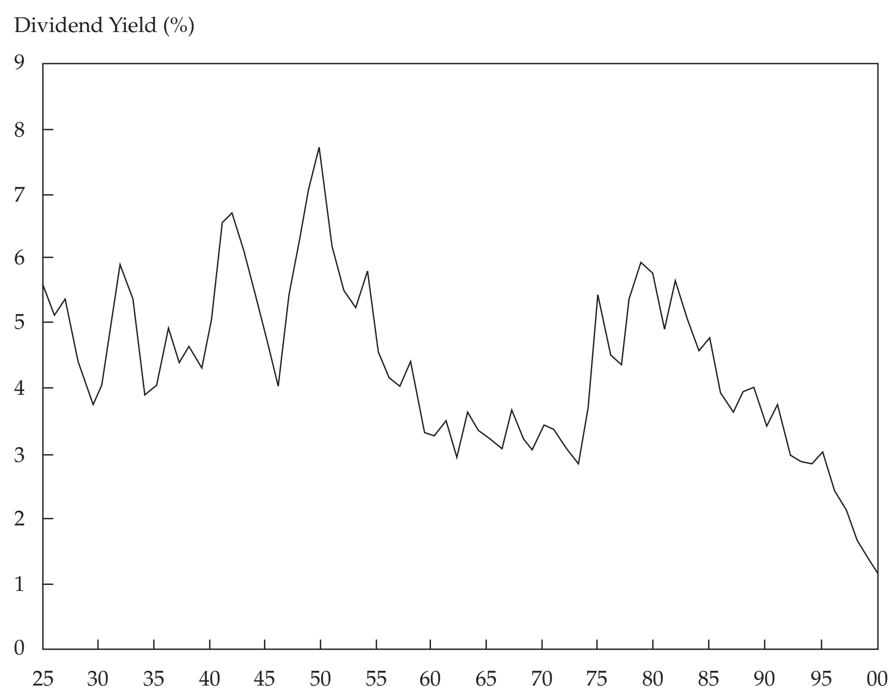
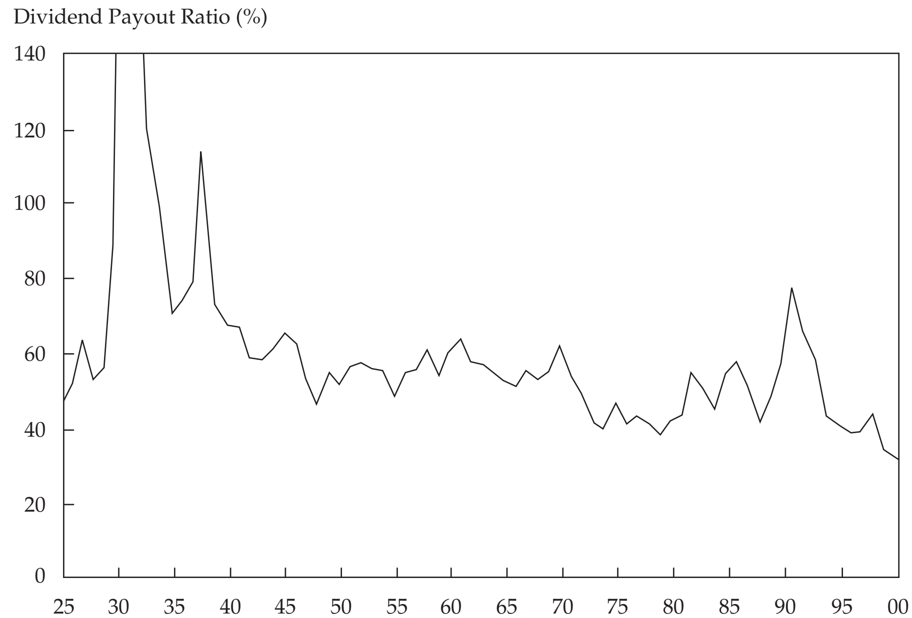


Figure 4. Dividend Payout Ratio, December 1925–December 2000



Note: The payout ratio as of December 1931 was 190.52 percent; as of December 1932, it was 929.12 percent.

larger and larger portion of companies in the market are not paying earnings out in the form of dividends. For example, the technology companies do not pay out any of their earnings as dividends. Thus, the payout ratio is not stable, and we may see it continue to fall.

A contender in the race to be a reliable forecasting tool (one that a number of people have already discussed today) is the dividend yield model in one of its many forms. If you could accept the dividend yield model by itself and with its purest assumptions—that is, the dividend yield plus dividend growth, assuming constant growth—the model would be a forecast of the stock market. But there are three problems with the pure dividend yield model that we must make adjustments for if the model is to be useful for forecasting. The first two problems are potential violations of Modigliani and Miller theory.

I am assuming that M&M holds true. (Despite what some of you have said about how dividend payouts do not seem to be reinvested in anything at all, I am clearly on the other side of that argument. If there is any truth to that supposition, however, that theory needs further investigation.) So, the first problem with some forms of the dividend yield model is that they violate M&M because they assume you can add the current dividend yield (which is now 1.10 percent) to historical dividend growth. Historical dividend growth underestimates historical earnings growth, however, because of the decrease in the payout ratio. Dividends have run slowest in the growth race because the payout ratio has continually dropped.

The second problem with using the dividend yield model as a forecasting tool (and it is, again, a violation of M&M) is that if the low payout ratios of today (31.8 percent) were reflected in the historical series, the percentage of earnings retained would have been higher and, therefore, historical earnings would have grown faster than observed. In short, the first problem is that dividend growth has been too slow historically, and the second problem is that with further earnings retention, historical earnings growth would have been potentially faster than observed.

The third problem with the dividend yield approach is the high P/E multiple observed today—over 25. Unlike some of you, I am going to assume efficient markets, which in this case I take to mean that the current high P/E implies higher-than-average future EPS growth.

My estimate of the average geometric equity risk premium is about 4 percent relative to the long-term bond yield. It is, however, 1.25 percent lower than the pure sample geometric mean from the risk premium of the Ibbotson and Sinquefield study (Ibbotson Associates 2001).

We have had some debate today on future growth rates—specifically for the 10-year horizon. Data that Peng and I are studying provide some support for the tie between high P/Es and high future growth. One of the problems with the 10-year horizon is that 10 years is not really long enough to encompass many independent events.

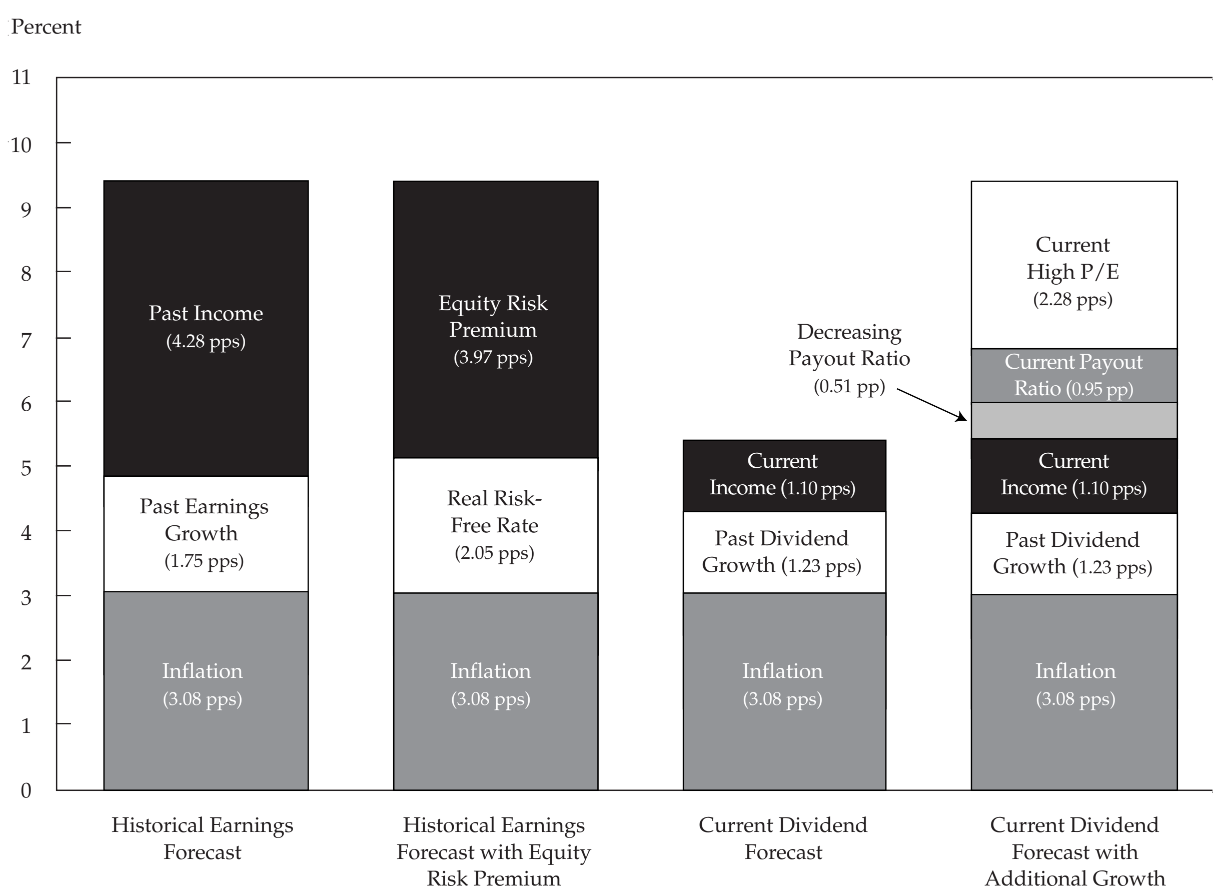
The extreme end of the spectrum of proponents of the dividend yield model would support using past dividend growth to forecast future dividend growth, then add current income. (Of course, that method almost wipes out the risk premium, and in some ways, it is actually similar to what Rob Arnott presented.)

In our response, we make three adjustments to the dividend yield model shown in the third column ("Current Dividend Forecast") of **Figure 5**. These are shown in the fourth column ("Current Dividend Forecast with Additional Growth"). We add 0.51 pp so that historical dividend growth matches historical earnings growth, we add an additional 0.95 pp because of the extra retention associated with the current record low payout rate, and finally we add 2.28 pps to future earnings growth to reflect the current high P/E that we assume forecasts higher earnings growth.

What about long-term earnings growth? Corporate America is likely to proceed in the next quarter century as it did in the previous 75 years. Corporate cash will be used for projects, investments, share repurchases, and acquisitions, but less and less will it be used for dividend payouts. Future earnings growth will be higher than past growth because of lower dividend payouts and the high current P/E. For the next 25 years, I predict (1) stocks will outperform bonds, (2) increased earnings growth will offset future low dividend yields, (3) the P/E jump from 10 to 26 will not repeat, and (4) the stock market return will provide more than 9 percent a year over the 25-year period.

JOHN CAMPBELL: When you make the adjustments, aren't you assuming not only efficient markets but also a constant discount rate? If so, you are assuming the answer. We are trying to find out what the discount rate is, but you assume the discount rate in your calculation. If so, aren't you bound to come up with an answer for the end that is the same as historical norms going in?

Figure 5. Historical versus Forecasts Based on Earnings and Dividend Models



IBBOTSON: True. In addition to assuming an efficient market (M&M), we are not assuming that the discount rate is dynamic. We are assuming it to be unknown, and we are searching for the single discount rate that best describes history. The presumption is that history can be extrapolated forward. It could be considered a reconciliation between the two approaches. Certainly, our quest is debatable.

BRADFORD CORNELL: I have some questions for Campbell Harvey. Are CFOs really not using their one-year-horizon market forecasts in evaluating their internal investments? Maybe the one-year market forecast they provide you is just a throw-away number; they are so uncertain about it that they do not incorporate it into any decision they make. If they really believe that the equity risk premium is zero today, shouldn't they be issuing stock?

CAMPBELL HARVEY: I think this survey gives us respondents' guesses of what is going to happen in the market; it does not necessarily map into what they are going to do in terms of their real project evaluations at a one-year horizon. In a recent working paper by Jagannathan and Meier (2001), which is based on some older work by McDonald and Siegal (1986), they say people tend to have higher hurdle rates than what the capital asset pricing model (CAPM) would suggest. CFOs are looking for the best projects, internal investments that throw off the best return, and there is no way they are going to accept a project with a rate of return equal to the T-bill rate—even if they expect next year's market return to be basically the same as the T-bill's return. So, what the data suggest to me is that there is a big difference between the short-horizon expectation of return and the hurdle rate one would actually use in terms of project evaluation. Of course, I want to go deeper into this problem by asking the survey participants for more details.

ROBERT ARNOTT: One would assume that to arrive at the estimated required return of any new commitment, a "credibility" hurdle rate is added on top of the cost-of-capital hurdle rate. Those cost-of-capital hurdle rates are always optimistic, so the credibility rate is added and is part of where the reported hurdle rate in the responses comes from.

MARTIN LEIBOWITZ: Just one clarification: How did your 10-year risk premium, 4.5 percent, relate to the hurdle rate? Do you have any evidence of what that longer-term hurdle rate is?

HARVEY: For the 10-year horizon, the risk premium reported is closer to the hurdle rate for internal projects than for the 1-year horizon. We don't have much information about the longer-term hurdle rate, but the next phase of my research with John Graham will be interviewing the CFO participants to shed additional light on these issues.

WILLIAM GOETZMANN: I was very excited to see Campbell Harvey's paper—to see more interesting data about dispersion of opinion. I know that in one of your earlier papers—the one on the market-timing ability of investment newsletter writers (Graham and Harvey 1996)—you unexpectedly found dispersion of opinion that had some forecasting ability. Cragg and Malkiel (1982) also found some dispersion in analysts' forecasts in relation to risk. Also, Massimo Massa and I have been finding some information about dispersion related to price effects and so forth (Goetzmann and Massa 2001). What particularly strikes me in looking at your results is the consistent message that this dispersion of opinion is having interesting effects that we ought to explore. If you are going to be talking to these CFOs, it would be great to find out more about the basis for the dispersion. It is an interesting potential area of research.

HARVEY: We have a lot of data on earnings forecasts, but I am more interested in the dispersion than the actual forecasts. An older paper by Frankel and Froot (1990) looked at dispersion of beliefs in terms of currency forecasting. It is very impressive. So, I agree that this area is worthy of more research.

THOMAS PHILIPS: I want to address the question about forecasts versus hurdle rates by describing an experience that I had. When I talk to our corporate clients, I often ask if they need help estimating their cost of capital (which, of course, is the same as the expected return) and I ask how they do it currently. Some tell me that they use the CAPM, while others say they use a more complicated factor model. But one answer stands out for its simplicity and its brilliance. At National Service Industries, an executive told me that his cost of capital was 10 percent. I asked him how he knew that it was 10 percent. He replied that he did not *know* that it was 10 percent. So, I queried further: "Why, then, do you assert that it is 10 percent?" He replied, "In my world, the cost of capital is not very important in terms of making new investment decisions. We have a hurdle rate to make that type of decision. The cost of capital is important to us because the lines of business that we are in are not fabulously profitable, and the simplest mistake we can make is to squander the capital we have invested in them. The one thing I want to do is to have every employee understand that capital is a real input and that it is incredibly easy to squander. When I use 10 percent as the cost of capital, everyone from the janitor to the CEO can apply it. They can move a decimal point; they can divide by 10. So, I can explain to them in simple terms that $1 million worth of equipment sitting idle represents $100,000 of real money going down the tubes every year. And that ability is much more important to me and to the company than having the right answer." Theoretically, he has the wrong answer, but in spite of that, his answer and approach are absolutely brilliant.

The other comment that I want to make is an observation on the difference in earnings growth rates. Roger Ibbotson is showing it growing close to per capita GDP.

ARNOTT: No, he has it growing faster than GDP.

PHILIPS: Roughly the same rate.

IBBOTSON: Historically, it is the same.

ARNOTT: But now the payout ratio is lower, so earnings would have to grow faster. Earnings growth is going to gain on GDP on a per share basis, not necessarily on an aggregate basis as Bradford Cornell was talking about.

WILLIAM REICHENSTEIN: Going back to what Rob Arnott said about taking another look at tactical asset allocation. Let's say that over the next 10 years, stocks, bonds, and cash will all produce a 10 percent rate of return. It seems to me the 10-year return should not make any difference; the asset-allocation decision is relatively insignificant at that point.

ARNOTT: Correct, the policy asset allocation decision is insignificant. For rebalancing to add value, for tactical asset allocation to add value, the absolutely crucial premise is that reversion to the mean will occur in at least a weak form.

REICHENSTEIN: That is when you pick up your alpha?

ARNOTT: Right. The presumption is based on a long-term historical record for live TAA experience. Even when it did not add value (in the 1990s), it did produce alpha. If there were not some weak reversion to the mean at work in the 1990s, it would not have produced an alpha.

LEIBOWITZ: Why do you say policy allocation is invariant? Even if you have zero difference in returns, you still have volatility.

ARNOTT: I am assuming geometric, not arithmetic, returns. If we assume arithmetic returns are the same, then the volatility differences carry a cost. If we assume the geometric returns are the same, then the return-maximizing portfolio is the risk-minimizing portfolio, which would probably have an allocation of only 10–20 percent equities. But the difference in returns would be tiny, so whether the allocation was 20/80 or 80/20 would not make much difference in the return.

LEIBOWITZ: But you would not have much in equities?

ARNOTT: This message is not welcomed with open arms by investors or investment practitioners. It has not been good for First Quadrant's business for me to publish this sort of stuff. Some consultants are annoyed because we are saying, basically, that the assumptions they are endorsing are wrong. Clients don't want to hear it because we've been correct for the last year and a half, and the losses hurt. When we first proposed the idea, it was viewed as slightly flaky, but since then, it's been on target—which has made some people even angrier.

GOETZMANN: I'm a bit confused. Are you talking about just *your* track record or evidence about TAA in general? I haven't seen any empirical evidence indicating that, on average (or even in the tails), any tactical allocators have been successful.

ARNOTT: I am speaking on the basis of our track record and what little information I can garner about competitors' track records. The comparative studies, like the one that Tom Philips did (Philips, Rogers, and Capaldi 1996), have dwindled to next to nothing because no one is interested in TAA. Our founding chairman was fond of saying, "Don't buy what's easy to sell. Do buy what's tough to sell." Well, TAA is tough to sell right now. I think it is an interesting idea that has fallen from favor in a circumstance where, prospectively, it is probably going to produce the kind of results that we had in the 1970s, which were breathtaking, just breathtaking.

PHILIPS: Let me comment on that. In the paper of mine that Rob Arnott is referring to, I took the actual live track records of every domestic TAA manager (about a dozen of them, and they had 95 percent of the assets under management in TAA at the time) and performed Henriksson–Merton and Cumby–Modest tests for timing skills. I found that in the 1970s, TAA was very successful. Then, in the 1980s, the results become a little mixed. If you include the period up to and including the crash of 1987, *all* the TAA managers added value; after the crash, no one added value. But here's an interesting twist to the story: Let's say a genie came to you once a quarter or once a month, take your choice, from 1980 onwards, and whispered "buy stocks" or "buy bonds" in your ear—and the genie was never wrong. And let's say you can make the appropriate portfolio changes without transaction costs. By how much did the genie outperform a simple 60/40 mixture of stocks and bonds? It turns out that the genie's outperformance went down enormously from the precrash to the postcrash period. It dropped from about 24 percent a year to about 15 percent a year. In effect, the genie got a lot less prosperous after 1987, so it's not surprising that TAA managers found themselves in trouble.

Summary Comments

Robert Arnott
John Campbell
Peng Chen, CFA
Bradford Cornell
William Goetzmann
Brett Hammond
Campbell Harvey
Roger Ibbotson
Martin Leibowitz
Rajnish Mehra
Thomas Philips
William Reichenstein, CFA
Robert Shiller
Kevin Terhaar, CFA
Peter Williamson

MARTIN LEIBOWITZ: I think it might be interesting to just go around the table for any last comments on our topic, the equity risk premium, or for any comments on any of the papers presented today.[1]

BRETT HAMMOND: I would like to hear more discussion from Roger Ibbotson and Rob Arnott. As I have listened to the presentations today, I have been trying to decide what we could say if we were charged as a group with coming to some consensus. I'm going to assume the role of the naive observer, and in that role, I can say I have learned that in some areas, we are talking past each other and in other areas, once we clarify the definitions (or what is being measured and how), we are closer together. That understanding is useful, but what is the next step in educating our colleagues and practitioners? What would we want to tell them about their problem, which is, of course, estimating the equity risk premium looking forward? I have been wanting to ask this question all day, so now I will: What would you tell them about the equity risk premium?

[1]For Martin Leibowitz's summary of academic and practitioner research on the equity risk premium, see the Webcast of his presentation to "Research for the Practitioner: The Research Foundation Pre-Conference Workshop" held in conjunction with the AIMR 2002 Annual Conference. The Webcast is available in summer 2002 at aimr.direct.org.

ROGER IBBOTSON: What you say is to the point. First, we see a need for clarification of what we mean by the equity risk premium: I think all of us in this room see it as an expectation, not a realization; if we look at realizations, it's to help us understand expectations. But not everybody outside the room understands this distinction.

The second issue is the use of "arithmetic" versus "geometric." Every time we make a forecast, we should say whether the forecast is arithmetic or geometric and which risk-free rate we are using—U.S. T-bills, the long bond, or TIPS.

Third, we need to distinguish between yields and returns. Jeremy Siegel, for example, used realized returns, whereas others today used realized yields.

Fourth, we should always specify the forecast horizon—whether we are talking about a short or a long horizon. The risk premium for a short horizon is basically about timing, an attempt to judge whether the market is currently over- or undervalued; the risk premium for the very long horizon provides a more stable concept of what the risk premium is—namely, the long-term extra return that an investor is expected to get for taking risks, assuming the market is fairly valued.

If we could at least get these definitions delineated and clarified and let everybody know what the definitions are, it would help identify the differences among us. We are actually much more of one mind than some might think. And the theoretical analyses actually come closer to the empirical results I might have imagined before this conference.

The 4 percent (400 bps) equity risk premium forecast that I have presented here today is a geometric return in excess of the long-term government bond yield. It is a long-term forecast, under the assumption that today's market is fairly valued.

WILLIAM REICHENSTEIN: I want to make a comment in terms of asset allocation based on the geometric difference between future stock and future bond returns. Let's say that the real return on stocks is expected to be 4 percent. Of course, the numbers would depend on the assumptions used; if you use the dividend model, the real return might be 2.5 percent, and with the earnings model, it might increase to 4 percent, but in either case, we are talking about a number well below the historical 7 percent real return on stocks. If we are looking at a real return on stocks of 4 percent and a real return on bonds of 3 percent, the equity risk premium is about 1 percent, which is much lower than in the past. So, the expectation for future equity real returns is down. But for a 50/50 stock/bond portfolio, if you use the historical Ibbotson numbers of 7 percent for stocks and 2 percent for bonds, then your historical real return on a 50/50 portfolio is 4.5 percent. How much worse off are you today at an estimate of 4 percent real return on stocks and 3 percent real return on bonds? That 50/50 portfolio has 3.5 percent real return instead of 4.5 percent, and that is only a 1 percentage point difference. Part of the reason the equity risk premium is lower, it seems to me, is because the real returns on bonds are up.

ROBERT ARNOTT: That's a very good point. The 4.5 percent versus the 3.5 percent expected *portfolio* return invites the question: Why is the actuarial community allowing sponsors to use 6.5 percent as an actuarial real return assumption for their aggregate balanced pension funds? The average nominal return is 9.3 percent, and the average inflation assumption is 2.8 percent. I would say that assuming a 6.5 percent real return is irresponsible and dangerous regardless of whether the reasonable expectation for real return going forward is 4.5 percent or 3.5 percent.

KEVIN TERHAAR: I think of the risk premium as most appropriately viewed as a discount rate element corresponding to a long horizon and relative to a risk-free rate, *commensurate with the asset's risk*. The risk premium issues that we have been discussing today are not unique to the U.S. equity market. Equities or bonds, or any other asset class for that matter, should be discounted in light of the risks that the asset entails. Although there seems to be some agreement on definition and, to a lesser extent, expectations, we are still left with a question that is one step removed from the equity risk premium: What is the appropriate price of risk as we look to the future? Even if we can agree that risk is more stable and thus more easily forecastable than return, and we are able to develop agreed-upon and reasonable forward-looking risk estimates, the issue of the appropriate *price of risk* still exists. Ultimately, it is this price of risk that determines the risk premium, not only of U.S. equities, but also of any other asset class. The risk premium on the domestic equity market should not and cannot be viewed in isolation.

LEIBOWITZ: In response to Brett Hammond, I'm very impressed by the level of consensus on the view that earnings can grow only at a somewhat slower rate than GDP per capita and that no one seems to feel it can grow much more—except Roger Ibbotson, who thought EPS could grow faster than GDP because of extra earnings retention and the implicit growth estimate inherent in the high recent price-to-earnings ratio. The fact that we're basically in agreement that earnings are tightly bound to the growth in the economy has, I think, a lot of implications. Also, I think we can agree that the distinction between arithmetic and geometric is important in terms of the way these concepts are discussed and analyzed. Another important point is that the term structure that is being used to analyze the risk premium must be defined. We also need to keep in mind that the estimation error over the short term is very, very high. So, our views, at least our expectations, may be more convergent over time, but the differences still remain.

Another thing that is surprising is the disconnect between the low growth assumption and the risk premium we tend to believe in, or at least corporate executives tend to believe in. Historically, the risk premium has been more than 5 percent, which may be tough to get in the future with the earnings growth numbers that have been cited today. I think we've come to some important agreements here.

I am troubled, however, by one aspect we haven't explored: Given the growth rate of GDP (the rate of all the corporate profits—including all the entrepreneurial profits that are not captured in the public market, all the free enterprise profits in the economy), how much of the earnings has to be reinvested to sustain that growth? That's a critical equilibrium question. Roger is the only person who addressed it, which he did in terms of his historical study. I think this point is worthy of a lot more thought.

ARNOTT: In terms of the lessons learned today, a tidy way to look at the whole returns picture is to hearken back to the basic notion that the real return on stocks has just three constituent parts—changes in valuation levels, growth, and income (whether income is dividends or dividends plus buybacks). We typically know the yield, so much of the discussion gets simplified to a reexamination of two key issues: (1) Is current pricing wrong? Should valuation levels change? (2) What growth rate is reasonable to expect? As you saw in the rather sharp dichotomy between my formulation for growth and Roger Ibbotson's formulation for growth, there's plenty of room for dialogue—in fact, immense room for dialogue.

A related aspect I think is interesting to observe is that, although there are a whole host of theories relating to finance, some of them elegant, brilliantly crafted, and sensible formulations of the way the world *ought* to work—the capital asset pricing model and Modigliani and Miller being two vivid examples—comparatively few people believe that the world actually works in exact accord with any such theories. We've seen tangible evidence that M&M, while a fine theory, doesn't necessarily work intertemporally. And we know that the CAPM in its raw form doesn't fit the data very well. This doesn't make it a bad theory; it's a wonderful theory and a wonderful formulation of the way the world ought to work. Similarly, the notion that higher P/Es should, in an efficient market, imply faster future permanent growth makes sense. It's an intuitive theory. Does it stand up to historical testing? No.

A similar lesson I think we can take away from today is that the theory and the reality of the risk premium puzzle differ. There are a host of theories that relate to the risk premium puzzle and, from our views on the risk premium, relate to the asset allocation decision, but the theories don't stand up to empirical tests. A very interesting area of exploration for the years ahead will be to try to find a theoretically robust construct that fits the real world.

CAMPBELL HARVEY: I was struggling through the morning just with the vocabulary related to the risk premium: It depends on the horizon; it depends on the risk-free rate; it's a moving target through time; it's conditional; it's unconditional. I now have a better understanding of these concepts and the difficulties in defining them. It is extraordinary that, given the importance of the definitions of these variables, there is so much disagreement in terms of approach. Indeed, I have to teach this material, and it is a difficult topic for the students. We talk in class about the risk premium, but we also have to take a step back and define risk, which is extraordinarily difficult to do.

We have talked today about the current state-of-the-art models. There is a burgeoning literature on different measures of risk, and we are learning a lot from the new behavioral theories. So, we are moving forward in our understanding of the risk premium. Indeed, some of the foremost contributors to this effort are in this room. And I think more progress will be made in the future. It is somewhat frustrating that we are not there yet. I cannot go into the classroom or into the corporate world and say with some confidence, "*This* is the risk premium."

ROBERT SHILLER: I was thinking about the ambiguity of our definitions of the equity risk premium and about what we mean by expectations. We tend to blur the concepts of our own expectations with the public's expectations and with rational expectations. And the interpretations we give to the concept of expectations have changed through time. The history of thought about expectations is interesting. I remember a 1969 article by Conard and Frankena about the term structure—before the rational expectations revolution—that asserted that there is no objective way to specify expectations in a testable model but by assuming perfect foresight. They wrote this after Muth (1961) wrote the first treatise on rational expectations but before it had any impact on the profession. Without access to the theoretical framework proposed by Muth, there was no concept at all of rational expectations. That was then, and now, today, 30 years later, we economists often seem to think that the word "expectations" has no other meaning than "rational expectations."

Economists today think expectation is the summation of P_iX_i, where P is the probability, but that is a very abstract concept that we've been taught. We can trace the word "probability" very far back in time, but it didn't always have all the associations that it has today. The word "probability" didn't even have the meaning that we attach to it now until the mid-1600s, when it seemed to suddenly explode on the intellectual scene. Before then, the word "probability" existed, but it meant "trustworthiness" and had no connection at all to our modern concept of probability. Suddenly, Blaise Pascal and others got people talking about probability, which led naturally to the concept of mathematical expectation.

Just as "probability" is not a natural concept, I think "expectations" is not a natural concept. When you do surveys and you ask people for their expectations, should we expect them to give us some calculation of mathematical expectations? In fact, their reaction to questions about their expectations often seems a sort of a panic: What are these people asking for? What kind of number do they want? I have to come up with a number fast! (Incidentally, a lot of people don't remember that John Maynard Keynes' first claim to fame was a 1921 book about probability in which he argued that people really don't have probabilities as we think of them today.[2])

With all of these ambiguities, one starts to wonder what the equity risk premium is measuring. When I was surveying individual and institutional investors about their outlook for the market, I found that if I asked investors what they thought the DJIA would do in the next year, the average answer was + 5 percent. But the PaineWebber/Gallup survey taken at the same time found that investors thought the DJIA would rise by 15 percent. That's quite a big discrepancy. So, I called Gallup and asked them if we could figure out the reason for such different results. As it turned out, the different survey responses were a function of the wording of the questions. The Gallup

[2] This work can be found in Keynes (1973).

poll was conducted by randomly telephoning people at the dinner hour. Their question was (more or less): What return do you expect on the stock market in the next year in percentage terms? My survey was conducted through a written questionnaire, and the specific question about the market was (more or less): "What do you think the DJIA is going to do in the next 12 months? Put a plus mark if you think it's going to go up and put a minus mark if you think it's going to go down."

The critical difference is that I mentioned the possibility that the market might go down, so about one-third of my answers were negative. I called Gallup and asked them what fraction of their respondents said "Down." And they said that there were so few down responses that they rounded them to zero. So, I was trying to figure out why they got so few negative responses. Well, the Gallup respondents were called at dinnertime, and maybe the person who called was somewhat intimidating, so respondents had to have some courage to say they thought the market return was going to be negative. In my survey, however, I brought up in writing a possible negative choice, and I got a lot of negative responses. So, I think reported expectations are very fragile.

In the investment profession, we've learned to have respect for psychologists and the concepts they use because they've learned a lot by studying how people frame their thinking and decision making. The concepts arising from this knowledge can be very helpful to us in our work. And psychologists deal with other attitudes related to expectations—aspiration, hope, regret, fear, and the salience of stories. All of these parameters are constantly changing through time. So, when you ask someone about their expectations, the answer they give will be very context sensitive.

With surveys, we've learned you need to ask exactly the same questions in exactly the same order on each questionnaire. Even so, you don't know quite what you're really getting because expectations have so many different definitions.

RAJNISH MEHRA: I want to make two quick comments. My first point is that valuation models help us structure the problem, but what breathes life into a valuation model are the forecasts, and these forecasts have huge conditional errors. Not many of the estimates for the equity premium that were given today were accompanied by the standard deviation of that estimate. That standard deviation is too important to be missing. For example, in my data relating the expected mean equity risk premium to national income, the standard deviation around that mean is huge. Just giving a point estimate is not enough. The omission of the conditional error worries me.

My second point is that profound demographic shifts are going to be occurring in the United States, in terms of the Baby Boomers retiring, about which Ed Prescott and I wrote (1985). That phenomenon is going to lead to asset deflation, which has profound implications for the *ex ante* equity premium.

THOMAS PHILIPS: I have been very interested to see two broad strands of thought discussed today. One of these strands, exemplified by Rajnish Mehra, is the line of thinking in which the basic model involves human economic behavior, whether that behavior is utility maximizing or motivated by something else, and the effects of that behavior in the capital markets. The second strand is more empirical—constructing a point estimate for the equity risk premium—and it is exemplified by Rob Arnott's and Roger Ibbotson's work. I see two somewhat different challenges for these two strands, and ultimately, they have to meet in the middle so that we can build a unified theory.

For the economist, the challenge I see is related to Richard Feynman's argument about why scientific imagination is so beautiful: It must be consistent. You cannot imagine just anything; it has to be consistent with classical mechanics, with quantum mechanics, with general relativity, and so on and so forth. Within this set of constraints, beautiful ideas are born that tie neatly into a powerful edifice. I see the challenge for financial economists as not simply explaining the equity risk premium but explaining a fairly wide range of economic phenomena within a unified framework. Instead of a patchwork of models, financial economics needs to look more like physics.

The challenge for the second group of people, those who provide the point estimates, is (as Rajnish Mehra correctly points out) to estimate some of the errors in our estimates and to be able to communicate all this information in a language that is accessible to the person on the street. In particular, we need to dissuade investors from using the sample mean as the best estimator of the true mean.

So, the two challenges are different, but the overarching challenge is to somehow unify the two approaches in a clean way that answers the question of what the equity risk premium is and makes tactical predictions.

BRADFORD CORNELL: I like to think more in terms of valuation and expected returns than in terms of the equity risk premium. The salient feature to me in that regard is that corporate profits after tax seem to be closely tied to GNP, particularly if the market is measured properly, in the aggregate and not limited to the S&P 500 Index, so that what we have to value is not all that uncertain. However, the way we value earnings, as Rajnish Mehra pointed out, has changed quite a bit. Stock market value in the United States has varied over time from half of GNP to twice GNP, which is about where it is now. To say that earnings are twice GNP, we either have to say that the expected returns are low and are expected to remain low for the long term or that the market has simply made a mistake. The one point that I would make to practitioners, fund managers, and so forth, is that they cannot maintain a 6.5 percent actuarial assumption in light of these data.

PENG CHEN: I think there are probably two types of data: One type is what the companies and the economy reveal—the analysis that Roger Ibbotson and I are working on—and the other type is drawn from the investor's point of view—how much the investor expects from a project or a security. What I think is really interesting is that the answers are going to lie between these two dynamics. How people adjust to the dynamics, how the dynamics change people's behavior, and how that behavior affects the market are very important to observe. I think the reason we see the valuation of the market rise and fall is not necessarily because the entire investment community believes the actual risk premium has fallen or gone up or that risk rose or fell but because of this dynamic. Not all investors have to change their minds to affect market value. Maybe the dynamic affected only a small number or a certain group of investors; only a marginal number of investors have to change their minds. So, it would be interesting to see how the two sides work together dynamically.

PETER WILLIAMSON: One of the most interesting aspects of our discussion today is the areas of agreement and of disagreement. The benefit of identifying areas of disagreement is that it can lead to the search for the reason for the disagreement. It is fascinating to me how all of the findings or theory might be implemented. Can you imagine an active manager turning to his clients and saying, "You must understand that the growth in earnings of your portfolio can't exceed GDP growth"? The client wouldn't believe it, and the manager wouldn't believe it. An active manager can't afford to believe it. Or can you imagine a firm that sells S&P 500 indexed funds sending a letter to all of the shareholders saying that they must realize earnings cannot grow faster than GDP? I can't imagine that message going out. So, what impact does all of the discussion we have had today make on the actual allocation of assets, the actual management of money? I don't know. I don't know whether investors ever have to *really* understand the equity risk premium, whether it's even in their best interest to understand it.

As for allocation, my sense is that different sectors of the investment community will do very different things in terms of asset allocation on the strength of the same expected risk premium. I think that the CREF participant who's 25 years old—looking ahead 40 years to retirement, saving money—versus the investor who is 66 years old—in the process of "dissaving," consuming now—given the same expected rate of return on equity, might do very different things with their money.

Richard Thaler and I deal with the problem of college and university endowment funds. One would think that endowment funds should all be thinking very long term, but the decisions are made by people—who don't live centuries and who, in fact, can be very embarrassed if the endowment has even one very poor quarter. For example, I am on the investment committee of a prep school, and years ago, the trustees agreed that the school should be much more heavily invested in equities, that the school should be thinking long term—but not yet. And each year, the suggestion is repeated, but the decision is: not yet.

It's very, very difficult for people to think long term. Yet, to a large extent, what we've been talking about today is what's sensible for the long term. Well, if people simply cannot think long term, then we are reduced to decisions for the short term. And the asset allocation implications may be very different for investors who cannot think much beyond the next quarter from the implications for those who, in theory at least, ought to be thinking about the next 50 years.

In short, I'm really puzzled about where all that we have discussed goes in terms of making any impact on investment behavior and on asset allocation.

JOHN CAMPBELL: My starting point is that we live in a world in which the forward-looking, *ex ante* equity premium that you might expect if you're a thoughtful investor trying to be rational changes over time, and those changes have implications for the methods used to estimate the premium. We've discussed these estimation methods today, and I think we have quite a consensus that past returns can be very misleading so it is probably better to start with valuation ratios and adjust them for growth expectations.

If we live in a world in which these numbers—the real interest rate, the equity premium, and so forth—change over time, that has a big impact on asset allocation. So, I can't resist plugging my forthcoming book with Luis Viceira (2002), *Strategic Asset*

Allocation: Portfolio Choice for Long-Term Investors. Brad Cornell's colleagues at UCLA coined the term "strategic asset allocation" to contrast with tactical asset allocation (Brennan, Schwartz, and Lagnado 1997). TAA is myopic; it looks at the next period, at the risk–return in one period. The idea behind strategic asset allocation is that if risk premiums are changing over time, the risks of different asset classes may look different for different horizons. You wouldn't get such an effect if returns were identically and independently distributed, but it can become quite important if the stock market is mean reverting or if real interest rates change over time.

I'm a little more optimistic than Peter Williamson is. I think there is some hope of influencing the practical world to think about these issues, because many of the rules of thumb that financial planners have used for years have this flavor. That is, the rules make more sense in a dynamically changing world than they would in an i.i.d. world. So, there's been a mismatch between academic research and practitioners' rules of thumb. We can close that gap if we accept in our models of asset allocation that investment opportunities change over time. So, we might, with some additional work, be able to narrow the gap between how practitioners think and how academics think.

WILLIAM GOETZMANN: The thing that struck me about our discussion today is that, with the exception of Campbell Harvey's paper, almost everything we're doing is an interpretation of history—whether it's historical valuation ratios, arithmetic means, or what have you. That basis for argument is exciting but has its limitations. History, after all, is a series of accidents; the existence of the time series since 1926 might itself be an accident. So, I'm more convinced than ever that we've got to find a way out of the focus on U.S. historical data if we want to solve some of these questions and to reassure ourselves, if indeed we can, that the equity premium is of a certain magnitude.

LEIBOWITZ: Thank you all.

References and Recommended Readings

Abel, Andrew B. 1990. "Asset Pricing under Habit Formation and Catching Up with the Joneses." *American Economic Review*, vol. 80, no. 2:38–42.

Ang, Andrew, and Geert Bekaert. 2001. "Stock Return Predictability: Is it there?" National Bureau of Economic Research (NBER) Working Paper 8207 (April).

Arnott, Robert D., and Clifford S. Asness. 2002. "Does Dividend Policy Foretell Earnings Growth?" Unpublished working paper (January).

Arnott, Robert D., and Peter L. Bernstein. 2002. "What Risk Premium is 'Normal'?" *Financial Analysts Journal*, vol. 58, no. 2 (March/April 2002):64–85.

Arnott, Robert D., and James N. von Germeten. 1983. "Systematic Asset Allocation." *Financial Analysts Journal*, vol. 39, no. 6 (November/December):31–38.

Arnott, Robert D., and Ronald Ryan. 2001. "The Death of the Risk Premium: Consequences of the 1990s." *Journal of Portfolio Management*, vol. 27, no. 3 (Spring):61–74.

Asness, Clifford S. 2000a. "Bubble Logic or How to Learn to Stop Worrying and Love the Bull." AQR working paper (June).

———. 2000b. "Stocks versus Bonds: Explaining the Equity Risk Premium." *Financial Analysts Journal*, vol. 56, no. 2 (March/April):96–113.

Bagwell, Laurie Simon, and John B. Shoven. 1989. "Cash Distributions to Shareholders." *Journal of Economic Perspectives*, vol. 3, no. 3 (Summer):129–140.

Bansal, Ravi, and Amir Yaron. 2000. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." NBER Working Paper 8059 (December).

Bansal, Ravi, Robert F. Dittmar, and Christian T. Lundblad. 2001. "Consumption, Dividends, and the Cross-Section of Equity Returns." Working paper, Duke University (October).

Barberis, N., M. Huang, and T. Santos. 1996. "Prospect Theory and Asset Prices." Working paper, University of Chicago.

Benartzi, S., and R. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics*, vol. 110, no. 1 (February):73–92.

Bernstein, Peter L. 1993. *Volatility: Two Views, Two Signals*. Newsletter. (November 15).

———. 1996. "Are Stocks the Best Place to Be in the Long Run? A Contrary Opinion." *Journal of Investing*, vol. 5, no. 2 (Summer):6–9.

———. 1997. *Stock/Bond Risk Perceptions and Expected Returns*. Newsletter. (February 1).

———. 2001a. "Ending Endism." *Economics and Portfolio Strategy* (April 30).

———. 2001b. "The Real Economy and the Unreal Economy." *Economics and Portfolio Strategy* (May 15).

Black, Fisher. 1976. "Studies of Stock Price Volatility Changes." *Proceedings of the 1976 Meetings of the American Statistical Society, Business and Economics Statistics Section*:177–181.

Bogle, John C. 1991. "Investing in the 1990s: Occam's Razor Revisited." *Journal of Portfolio Management*, vol. 18, no. 1 (Fall):88–91.

———. 1995. "The 1990s at the Halfway Mark." *Journal of Portfolio Management*, vol. 21, no. 4 (Summer):21–31.

Boudoukh, Jacob, and Matthew Richardson. 1994. "The Statistics of Long-Horizon Regressions Revisited." *Mathematical Finance*, vol. 4, no. 2 (April): 103–119.

Brennan, Michael J., Eduardo S. Schwartz, and Ronald Lagnado. 1997. "Strategic Asset Allocation." *Journal of Economic Dynamics and Control*, vol. 21, no. 8–9 (June):1377–1403.

Campbell, J.Y. Forthcoming 2002. "Consumption Based Asset Pricing." In *Handbook of the Economics of Finance*. Edited by G.M. Constantinides, M. Harris, and R. Stulz. Amsterdam, Netherlands: North-Holland.

Campbell, John Y., and J.H. Cochrane. 1995. "By Force of Habit: A Consumption-Based Explanation of

Aggregate Stock Market Behavior." NBER Working Paper 4995.

———. 1999. "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." *Journal of Political Economy*, vol. 107 (April):205–251.

Campbell, John Y., and Robert J. Shiller. 1988. "Stock Prices, Earnings, and Expected Dividends." *Journal of Finance*, vol. 43, no. 3 (July):661–676.

———. 1998. "Valuation Ratios and the Long-Run Stock Market Outlook." *Journal of Portfolio Management*, vol. 24, no. 2 (Winter):11–26.

———. 2001. "Valuation Ratios and the Long-Run Stock Market Outlook: An Update." NBER Working Paper 8221 (April).

Campbell, John Y., and Luis M. Viceira. 2002. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford, U.K.: Oxford University Press.

Chancellor, Edward. 1999. *Devil Take the Hindmost*. New York: Farrar, Straus, and Giroux.

Claus, James, and Jacob Thomas. 2001. "Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stocks." *Journal of Finance,* vol. 56, no. 5 (October):1629–66.

Cole, Kevin, Jean Helwege, and David Laster. 1996. "Stock Market Valuation Indicators: Is This Time Different?" *Financial Analysts Journal*, vol. 52, no. 3 (May/June):56–64.

Conard, Joseph W., and Mark W. Frankena. 1969. In *Essays on Interest Rates*. Edited by Jack M. Guttentag and Philip Cagan. New York: NBER and Columbia University Press.

Constantinides, George C. 1990. "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy*, vol. 98 (June):519–543.

Constantinides, George M., John B. Donaldson, and Rajnish Mehra. 1998. "Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle." NBER Working Paper 6617 (June).

Cornell, Bradford. 1999. *The Equity Risk Premium*. New York: John Wiley & Sons.

Cowles, Alfred. 1938 (1939, 2nd edition). *Common Stock Indexes: 1871–1937*. Bloomington, IN: Principia.

Cragg, John G., and Burton G. Malkiel. 1982. *Expectations and the Structure of Share Prices*. Chicago, IL: University of Chicago Press.

Damodaran, Aswath. Forthcoming 2002. "Choosing the Right Valuation Model." In *Equity Valuation Techniques*. Charlottesville, VA: AIMR.

Davidson, L.B., and D.O. Cooper. 1976. "A Simple Way of Developing a Probability Distribution of Present Value." *Journal of Petroleum Technology* (September):1069–78.

Davis, S.J., and Rajnish Mehra. Forthcoming 2002. *Stocks and Bonds in the Portfolio Life-Cycle*. Saddle River, NJ: Prentice Hall.

DeLong, B., A. Shleifer, L. Summers, and R. Waldmann. 1990. "Noise Trader Risk in Financial Markets." *Journal of Political Economy*, vol. 98, no. 4 (August):703–738.

Diermeier, Jeffrey, Roger Ibbotson, and Laurence B. Siegel. 1984. "The Supply of Capital Market Returns." *Financial Analysts Journal*, vol. 40, no. 2 (March/April):74–80.

Dimson, Elroy, Paul Marsh, and Mike Staunton. 2001. *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton, NJ: Princeton University Press.

Epstein, L., and S. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica,* vol. 57, no. 4 (July):937–969.

Fama, Eugene F., and Kenneth R. French. 1988. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, vol. 22, no. 1 (October):3–26.

———. 1989. "Business Conditions and Expected Returns on Stocks and Bonds." *Journal of Financial Economics*, vol. 25, no. 1 (July):23–50.

———. 2000. *The Equity Risk Premium*. Center for Research in Security Prices Working Paper 522 (April).

Fama, Eugene F., and Kenneth R. French. 2001. "The Equity Premium." Center for Research in Security Prices Working Paper No. 522 (April).

Fisher, Irving. 1930. *The Theory of Interest*. New York: MacMillan.

Frankel, Jeffery, and Kenneth Froot. 1990. "Chartists, Fundamentalists, and Trading in the Foreign

Exchange Market." *American Economic Review*, vol. 80, no. 2 (May):181–185.

French, Kenneth R., G. William Schwert, and Robert F. Stambaugh. 1987. "Expected Stock Returns and Volatility." *Journal of Financial Economics*, vol. 19, no. 1 (January):3–29.

Garber, Peter M. 2000. *Famous First Bubbles: The Fundamentals of Early Manias*. Cambridge, MA: MIT Press.

Glassman, James K., and Kevin A. Hassett. 1999. *Dow 36,000*. New York: Random House.

Goetzmann, William N., and Philippe Jorion. 1993. "Testing the Predictive Power of Dividend Yields." *Journal of Finance*, vol. 48, no. 2 (June):663–679.

Goetzmann, William N., and Massimo Massa. 2001. "Dispersion of Opinion and Stock Returns: Evidence from Index Fund Investors." Yale ICF Working Paper No. 00-55 (August).

Goyal, Amit, and Ivo Welch. 1999. "Predicting the Equity Premium." Working paper, University of California at Los Angeles.

———. 2001. "Predicting the Equity Premium with Dividend Ratios." Yale ICF working paper, University of California at Los Angeles (December).

Graham, B., and D.L. Dodd. 1934. *Security Analysis*. New York: McGraw-Hill.

Graham, John, and Campbell R. Harvey. 1996. "Market Timing Ability and Volatility Implied in Investment Newsletters' Asset Allocation Recommendations." *Journal of Financial Economics*, vol. 42 (January):397–422.

———. 2001a. "Expectations of Equity Risk Premia, Volatility and Asymmetry from a Corporate Finance Perspective." Working paper, Duke University.

———. 2001b. "The Theory and Practice of Corporate Finance: Evidence from the Field." *Journal of Financial Economics*, vol. 60, no. 2–3 (May):187–243.

———. 2002. "How Do CFOs Make Capital Budgeting and Capital Structure Decisions?" Working paper, Duke University (March).

Granger, C.J.W., and P. Newbold. 1974. "Spurious Regressions in Econometrics." *Journal of Econometrics*, vol. 2:111–120.

Hicks, Sir John Richard. 1946. *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory*. Oxford, UK: Oxford University Press.

Ibbotson, Roger, and Peng Chen. 2002. "Stock Market Returns in the Long Run: Participating in the Real Economy." Yale ICF Working Paper (March).

Ibbotson, Roger, and Rex A. Sinquefield. 1976a. "Stocks, Bonds, Bills and Inflation: Year-by-Year Historical Returns (1926–74)." *Journal of Business*, vol. 49, no. 1 (January):11–47.

———. 1976b. "Stocks, Bonds, Bills and Inflation: Simulations of the Future (1976–2000)." *Journal of Business*, vol. 49, no. 3 (July):313–338.

Ibbotson Associates. 2001. *Stocks, Bonds, Bills, and Inflation: 2000 Yearbook*. Chicago, IL: Ibbotson Associates.

Ilmanen, Antti. 1995. "Time-Varying Expected Returns in International Bond Markets." *Journal of Finance*, vol. 50, no. 2 (June):481–506.

Jagannathan, Murali, Clifford P. Stephens, and Michael S. Weisbach. 2000. "Financial Flexibility and the Choice between Dividends and Stock Repurchases." *Journal of Financial Economics*, vol. 57, no. 3 (September):355–384.

Jagannathan, Ravi, and Iwan Meier. 2001. "Do We Need the CAPM for Capital Budgeting?" Working paper, Northwestern University.

Jagannathan, Ravi, Ellen R. McGrattan, and Anna Scherbina. 2000. "The Declining U.S. Equity Premium." *Quarterly Review*, Federal Reserve Bank of Minneapolis, vol. 24, no. 4 (Fall):3–19.

Janis, Irving. 1982. *Groupthink*. Boston, MA: Houghton Mifflin.

Jones, Charles. 2001. "A Century of Stock Market Liquidity and Trading Costs." Paper presented to the NBER Summer Institute, Asset Pricing Meetings.

Kahneman, Daniel, and Amos Tversky. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science*, vol. 185, no. 4157 (September):1124–31.

———. 1979 (and 1974). "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, vol. 46, no. 2 (March):171–185.

Kane, Alex, Alan J. Marcus, and Jaesun Noh. 1996. "The P/E Multiple and Market Volatility." *Financial Analysts Journal*, vol. 52, no. 4 (July/August):16–24.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money*. New York: Harcourt Brace and World.

———. 1973. *Treatise on Probability* [1921]. In *The Collected Writings of John Maynard Keynes*. London, U.K.: MacMillan, for the Royal Economic Society.

Kirby, Chris. 1997. "Testing the Predictable Variation in Stock and Bond Returns." *Review of Financial Studies*, vol. 10, no. 3 (Fall):579–630.

Knight, Frank. 1964. *Risk, Uncertainty and Profit*. New York: Augustus M. Kelley.

Lander, Joel, Athanasios Orphanides, and Martha Douvogiannis. 1997. "Earnings Forecasts and the Predictability of Stock Returns: Evidence from Trading the S&P." *Journal of Portfolio Management*, vol. 23, no. 4 (Summer):24–35.

Lintner, John. 1956. "Distribution of Incomes of Corporations among Dividends, Retained Earnings, and Taxes." *American Economic Review*, vol. 46:97–113.

Macaulay, Frederick R. 1938. *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States since 1856*. Washington, DC: National Bureau of Economic Research.

McDonald, R., and Dan Siegal. 1986. "The Value of Waiting to Invest." *Quarterly Journal of Economics*, vol. 101, no. 4:707–728.

McGrattan, Ellen R., and Edward C. Prescott. 2000. "Is the Stock Market Overvalued?" *Federal Reserve Bank of Minneapolis Quarterly Review*, vol. 24, no. 4 (Fall):20–40.

McGrattan, Ellen R., and Edward C. Prescott. 2001. "Taxes, Regulations, and Asset Prices." Unpublished paper, Federal Reserve Bank of Minneapolis (March).

Mehra, Rajnish. 1998. "On the Volatility of Stock Prices: An Exercise in Quantitative Theory." *International Journal of Systems Science*, vol. 29, no. 11:1203–11.

———. Forthcoming 2002a "Equity Premium Puzzle." In *Mastering Investments*. Edited by James Pickford. Saddle River, NJ: Prentice Hall.

———. Forthcoming 2002b. "The Equity Premium: Why Is It a Puzzle?" *Financial Analysts Journal*.

Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economies*, vol. 15, no. 2 (March):145–161.

———. 1988. "The Equity Risk Premium: A Solution?" *Journal of Monetary Economics*, vol. 22 (July):133–136.

———. Forthcoming 2002. "The Equity Premium Puzzle in Retrospect." In *Handbook of the Economics of Finance*. Edited by G.M. Constantinides, M. Harris, and R. Stulz. Amsterdam, Netherlands: North Holland.

Merton, R.C. 1980. "On Estimating the Expected Return on the Market." *Journal of Financial Economics*, vol. 8, no. 4 (June):323–361.

Miller, Merton, and Franco Modigliani. 1961. "Dividend Policy, Growth, and the Valuation of Shares." *Journal of Business*, vol. 34, no. 4 (October):411–433.

Modigliani, F., and R. Cohn. 1979. "Inflation, Rational Valuation, and the Market." *Financial Analysts Journal*, vol. 35, no. 2 (March/April):24–44.

Modigliani, Franco, and Merton H. Miller. 1958. "The Cost of Capital, Corporation Finance, and the Theory of Investment." *American Economic Review*, vol. 48, no. 3 (June):261–297.

Muth, John F. 1961. "Rational Expectations and the Theory of Price Movements." *Econometrica*, vol. 29:315–335.

Philips, Thomas K. 1999. "Why Do Valuation Ratios Forecast Long-Run Equity Returns?" *Journal of Portfolio Management*, vol. 25, no. 3 (Spring):39–44.

Philips, T.K., G.T. Rogers, and R.E. Capaldi. 1996. "Tactical Asset Allocation: 1977–1994." *Journal of Portfolio Management*, vol. 23, no. 1 (Fall):57–64.

Poterba, James M., and Lawrence H. Summers. 1986. "The Persistence of Volatility and Stock Market Fluctuations." *American Economic Review*, vol. 76, no. 5 (December):1142–51.

Reichenstein, William. Forthcoming 2002. "What do Past Stock Market Returns Tell Us about the Future?" *Journal of Financial Planning*.

Samuelson, Paul A. 1994. "The Long-Term Case for Equities and How It Can Be Oversold." *Journal of Portfolio Management*, vol. 21, no. 1 (Fall):15–26.

———. 1979. "Why We Should Not Make Mean Log of Wealth Big Though Years to Act Are Long." *Journal of Banking and Finance*, vol. 3, no. 4:305–308.

Schwert, G. William. 1989. "Why Does Stock Market Volatility Change over Time?" *Journal of Finance*, vol. 44, no. 5 (December):1115–54.

Schwert, G. William. 1990. "Indexes of United States Stock Prices from 1802 to 1987." *Journal of Business*, vol. 63, no. 3 (July):399–426.

Shafir, E., I. Simonson, and A. Tversky. 2000. "Reason-Based Choice." In *Choices, Values, and Frames*. Edited by D. Kahneman and A. Tversky. Cambridge, U.K.: Cambridge University Press.

Shiller, Robert J. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review*, vol. 73, no. 3 (June):421–436.

———. 1989. *Market Volatility*. Boston, MA: Massachusetts Institute of Technology Press.

———. 1990. "Market Volatility and Investor Behavior." *American Economic Review*, vol. 80:58–62.

———. 2000. *Irrational Exuberance*. Princeton, NJ: Princeton University Press.

———. 2002. "Bubbles, Human Judgment, and Expert Opinion." *Financial Analysts Journal*, vol. 58, no. 3 (May/June):18–26.

Shleifer, Andrei, and Nicholas Barberis. 2000. "Style Investing." NBER Working Paper No. 8039 (December).

Siegel, Jeremy J. 1998. *Stocks for the Long Run*. 2nd. ed. New York: McGraw Hill.

———. 1999. "The Shrinking Equity Premium." *Journal of Portfolio Management*, vol. 26, no. 1 (Fall):10–17.

———. 2002. *Stocks for the Long Run*. 3rd ed. New York: McGraw Hill.

Smith, Edgar Lawrence. 1924. *Common Stocks as Long-Term Investments*. New York: MacMillan.

Sorensen, Eric H., and Robert D. Arnott. 1988. "The Risk Premium and Stock Market Performance." *Journal of Portfolio Management*, vol. 14, no. 4 (Summer):50–55.

Stambaugh, Robert F. 1984. "Bias in Regressions with Lagged Stochastic Regressors." Manuscript.

Statman, Meir. 2002. "Lottery Players/Stock Traders." *Financial Analysts Journal*, vol. 58, no. 1 (January/February):14–21.

Thaler, R., A. Tversky, D. Kahneman, and A. Schwartz. 1997. "The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test." *Quarterly Journal of Economics*, vol. 112, no. 2 (May):647–661.

Welch, Ivo. 2000. "Views of Financial Economists on the Equity Premium and on Financial Controversies." *Journal of Business*, vol. 73, no. 4 (October):501–537.

Wilson, Jack W., and Charles P. Jones. Forthcoming 2002. "An Analysis of the S&P 500 Index and Cowles Extensions: Price Indexes and Stock Returns, 1870-1999." *Journal of Business*.

Note on Value Drivers[1]

Value-based management assumes that value creation should be a primary consideration in managerial decision making. It requires a thorough understanding of what creates value and why as well as the ability to measure value accurately. The goal of this note is to highlight the determinants of equity value and, in doing so, provide a framework for making financial, strategic, and investment decisions. In particular, the note describes three value drivers: profitability, advantage horizon, and reinvestment. Using both a theoretical model and a numerical example, it shows how each value driver affects equity value and explains why. It also presents empirical evidence to support the relation between the value drivers and value creation.

Theoretical Equity Valuation Model

Discounted cash flow (DCF) analysis translates future cash flows into current market values. For example, given a stream of equity cash flows (ECF) and a discount rate equal to the cost of equity (K_E), the market value of equity (E_{MV}) is the present value of future equity cash flows:

$$E_{MV} = ECF_1/(1+K_E) + ECF_2/(1+K_E)^2 + \ldots \tag{1}$$

When the equity cash flows and discount rate are constant over time, this series is a stable perpetuity which can be written as:

$$E_{MV} = ECF/K_E \tag{2}$$

Assuming that the equity cash flows are equal to the accounting return on equity (ROE) times the book value of equity (E_{BV}) at the beginning of the period, then equation 2 can be rewritten as:

$$E_{MV} = [(ROE)*(E_{BV})]/K_E \tag{3}$$

$$\text{where ROE} = \text{Net Income}/E_{BV}$$

While the assumption that equity cash flows are equal to accounting earnings is convenient for expositional reasons, this assumption is clearly not valid except in very special circumstances. For example, non-cash items such as depreciation or deferred taxes, and cash-items that do not flow through the income statement such as changes in working capital and fixed assets both cause cash

[1] Much of the material in this note appears in Fruhan (1979), chapter 1.

Professor Benjamin C. Esty prepared this note as the basis for class discussion.

flows to deviate from reported net income. Nevertheless, this assumption is not a bad approximation and, as will be shown in the next section, seems to generate reasonable empirical predictions.

After dividing each side of equation 3 by the book value of equity, the left side of the equality becomes the market-to-book ratio (the market value of equity divided by the book value of equity):

$$\text{Market/Book} = E_{MV} / E_{BV} = \text{ROE}/K_E \qquad (4)$$

Equation 4 says that a firm's market-to-book ratio equals the ratio of its return on equity to its cost of equity. This simple valuation model, or variations of it, can be used to analyze the relation between profitability, growth, and value.

Profitability

The first value driver, profitability, is immediately clear from equation 4. For a given industry, more profitable firms—those able to generate higher returns per dollar of equity—should have higher market-to-book ratios. Conversely, firms which are unable to generate returns in excess of their cost of equity should sell for less than book value.

Profitability	Value
If ROE > K_E	then Market/Book > 1
If ROE = K_E	then Market/Book = 1
If ROE < K_E	then Market/Book < 1

One implication of this model is that firms can increase equity value by increasing their return on equity. The Du Pont formula decomposes ROE into three components and provides some guidance on how to increase it:

$$\begin{aligned} \text{ROE} &= (\text{Net Income/Equity}) \\ &= (\text{Net Income/Sales}) * (\text{Sales/Assets}) * (\text{Assets/Equity}) \\ &= (\text{Profit Margin}) * (\text{Asset Turnover}) * (\text{Financial Leverage}) \end{aligned}$$

For example, increasing the profit margin through higher prices or lower costs will increase the ROE. Similarly, increasing the asset turnover by increasing inventory turnover or reducing days receivables will increase the ROE. However, increasing financial leverage has dual, and possibly contradictory, effects. It increases not only the ROE through the Du Pont formula, but also the cost of equity.

A firm's cost of equity, or equivalently investors' expected return on equity, can be estimated using the Capital Asset Pricing Model (CAPM). According to the model, the expected return on equity is a function of a firm's equity beta (β_E) which, in turn, is a function of both leverage and asset risk (β_A):

$$K_E = R_F + \beta_E (R_M - R_F) \qquad (5)$$

where:

R_M = return on the market portfolio
R_F = risk-free rate of return

$$\beta_E = [\, \beta_A - \beta_D (D/V)\,] (V/E) \qquad (6)$$

because:

$$\beta_A = \beta_D (D/V) + \beta_E (E/V) \qquad (7)$$

and

$$\text{Firm Value (V)} = \text{Debt Value (D)} + \text{Equity Value (E)} \qquad (8)$$

Assuming riskless debt, meaning the beta of debt is zero, then equation 6 can be written as:

$$\beta_E = \beta_A\ (V/E) \tag{9}$$

As financial leverage (D/V) increases, the ratio of firm value to equity value (V/E) increases, the equity beta increases, and, according to equation 5, the expected return on equity increases. The expected return increases because equity cash flows are riskier: leverage increases debtholders fractional claim on the firm's cash flows. As a result, an increase in leverage can either increase or decrease the ratio in equation 4 depending on whether the return on equity (the numerator) or the cost of equity (the denominator) increases faster.

Advantage Horizon

Equation 4 presents a firm's market-to-book ratio as a stable perpetuity under the assumption that its profitability remains constant forever. An alternative, and more realistic assumption, is that firms generate positive abnormal returns—returns in excess of their cost of capital—for only a limited number of years. The period during which firms generate positive abnormal returns is known as the advantage horizon.

Using a variation of the simple valuation model in equation 4, Appendix 1 derives the market-to-book ratio as an annuity rather than a stable perpetuity. It assumes that a firm's equity returns can be divided into two parts: *normal* returns equal to the firm's cost of equity (K_E) and *abnormal* returns equal to the actual ROE less the cost of equity (ROE - K_E). Viewed in this fashion, one can think of abnormal returns and the advantage horizon in the same way Stewart (1991) defines economic value added (EVA) and the competitive advantage period (CAP). Equation A1.8 from the Appendix 1 is:[2]

$$\textbf{Market/Book} = \mathbf{1 + (ROE\text{-}K_E) * [(1/K_E) - (1/\ (K_E(1+K_E)^n)]} \tag{10}$$

where the advantage horizon is defined as n years. According to this formula, the greater the spread between a firm's return on equity and its cost of equity (ROE - K_E), the longer the advantage horizon (increasing n), and the sooner abnormal returns occur (positive abnormal returns in early years), the higher the market-to-book ratio. Firms that earn normal returns (K_E = ROE) in all periods should have market-to-book ratios equal to one; firms that generate negative abnormal returns during the advantage (disadvantage) period should have market-to-book ratios less than one.

Equation 10 is more realistic than equation 4 because most firms earn positive abnormal returns for only a limited number of years. The presence of positive abnormal returns encourages entry by new firms and increased competition by existing firms. Over time, competition reduces excess returns to the point where firms just earn the expected, or normal, rate of return. Although there is typically an inverse relation between the magnitude of positive abnormal profits and the length of the advantage horizon, this model implies that firms should seek to extend the advantage horizon as long as possible for a given level of profitability.

Ghemawat (1991) refers to this ability to preserve competitive advantage as sustainability and asserts it is a key determinant of value creation. Sustainability, he maintains, depends on a firm's ability to create scarcity value and for the firm's owners to capture or appropriate this value. Threats to scarcity value include imitation and substitution. A firm can defend against imitation by erecting barriers to entry or forestalling entry through aggressive positioning; a firm can defend against substitution by continually improving or augmenting its product. Threats to appropriability include

---

[2] This formula is a variation of the accounting-based valuation methods described in Bernard (1994); Palepu, Bernard, and Healy (1996), and Ohlson (1995).

slack and hold-up both of which result from misaligned incentives. Slack occurs when firms fail to create as much value as they are capable of creating; hold-up occurs when non-owners, instead of owners, capture value. Non-owners are often able to capture value when they provide complementary, and necessary, inputs.

Reinvestment

The third value driver, reinvestment, builds on the other two factors and incorporates the concept of growth. Firms that have attractive investment opportunities, meaning that investments are expected to generate positive abnormal earnings, can create equity value by reinvesting earnings or by investing additional equity. Appendix 2 derives a valuation model which allows for reinvestment of earnings at rate γ where γ equals the retention rate or the fraction of net income reinvested in the firm. The quantity γROE is a firm's sustainable growth rate, the rate at which it can grow its assets (or sales if they are proportional to assets) without changing its capital structure or raising external equity. With reinvestment, the valuation model becomes (equation A2.4):

$$\text{Market/Book} = [\text{ROE}(1 - \gamma)] / (K_E - \gamma\text{ROE}) \qquad (11)$$

When a firm pays out all of its earnings as dividends, then the retention rate is zero ($\gamma = 0$) and equation 11 reduces to the simple valuation model in equation 4. Assuming a firm has attractive investment opportunities in which it can generate positive abnormal returns (ROE>K_E), then it can increase value by retaining a larger fraction of earnings and investing them in the business. Thus reinvestment and growth creates value only when a firm can generate positive abnormal returns on future investment opportunities. Those firms with the greatest number and the most profitable investment opportunities should have the highest market-to-book ratios provided they are able to fund the projects.

In fact, it is often convenient to think of firm value as consisting of two parts: the present value of assets in place and the present value of future growth opportunities (Myers, 1977). The former require little in the way of additional investment, while the latter are investment opportunities which are expected to earn positive abnormal returns. These investment opportunities are called "real" options because they resemble financial options, particularly call options. They can be interpreted and managed using option pricing theory and valued using option pricing techniques (see Luehrman, 1995).

Numerical Example

Combining equations 10 and 11 produces a single valuation model that incorporates all three value drivers. Exhibit 1 shows this model as well as the relation between a hypothetical firm's market-to-book ratio and the value drivers. The exhibit presents three cases with differing levels of reinvestment (γ = 0%, 33%, and 66%). For each case, there is a sensitivity table showing how the market-to-book ratio depends on the advantage horizon and level of profitability (ROE).

Case #1 (no reinvestment) shows that more profitable firms have higher market-to-book ratios—the ratio increases as one reads across the rows. As stated earlier, the impact of the advantage horizon depends on whether a firm generates positive or negative abnormal earnings. The longer a firm can generate positive abnormal earnings, the greater its market-to-book ratio. However, because of discounting, abnormal earnings in later years have a smaller impact on the market-to-book ratio than abnormal earnings in early years. Alternatively, firms that generate negative abnormal earnings have market-to-book ratios less than one. Moreover, their market-to-book ratio falls as the advantage (disadvantage) horizon gets longer. Finally, the market-to-book ratio is equal to one and is independent of the advantage horizon for firms that generate normal earnings (the case where ROE=K_E).

Cases #2 and #3 (with reinvestment rates equal to 33% and 66%, respectively) illustrate the impact of reinvestment. Like the advantage horizon, reinvestment creates additional value only for firms that generate positive abnormal earnings. When firms are able to generate positive abnormal returns (ROE = 25%), have a long advantage horizon (30 years), and reinvest a large fraction of earnings (γ = 66%), they create significant value. The difference between the market-to-book ratio in the high return/long horizon with no reinvestment (case #1) and with reinvestment (case #3) is large: 1.66 vs. 4.27.

Empirical Evidence

This section presents empirical evidence on the relation between the value drivers and value creation. Despite the assumptions imbedded in the simple valuation models, they do, nonetheless, yield predictions which are consistent with what we observe in practice.

Profitability

The model predicts that there is a relation between a firm's market-to-book ratio and the ratio of its return on equity to its cost of equity. Given a set of firms in a single industry, the model implies that there should be a positive relation between ROE's and market-to-book ratios for these firms assuming their costs of capital are approximately equal. To a first approximation, it is reasonable to assume that firms in the same industry will have similar capital costs because they hold similar assets and, typically, have similar capital structures.

Exhibit 2 shows the relation between market-to-book ratios and firm profitability for two quite different industries: grocery stores and oil field service companies. Whereas the grocery industry is a retail business with high inventories and low margins, the oil-field services industry is a service business with industrial customers and higher margins. Yet in both cases, there is a very clear, positive relation between equity value and ROE's: higher ROE's are associated with higher market-to-book ratios. Fruhan (1996) presents similar evidence for a much wider range of industries including newspapers, telecommunications, and specialty chemicals.

There are at least two reasons why this relation does not hold perfectly. First, not all firms in the same industry have the same leverage or same asset risk. Thus, financial and operating differences cause the cost of equity to differ across firms. Second, accounting data is subject to manipulation by managers. On the one hand, managers provide valuable information through their choice of accounting disclosures and policies. On the other hand, they are biased which may lead them to distort reported numbers. Fortunately, however, most distortions occur through accruals which eventually get reversed. Because accounting data is subject to this kind of manipulation, it is critical to understand whether the reported numbers reflect economic reality. To the extent high ROE's reflect economic reality, and not unreasonable deferral of costs or a one-time aberrations, then the relation shown in exhibit 2 will be stronger. When accounting data does not reflect economic reality, one must undo the distortions before trying to make substantive conclusions about the business or its prospects.

Advantage Horizon

Several researchers have studied the length of the advantage horizon. For example, Fruhan (1995) examined a sample of 87 "high-performing" firms defined as those firms with sales of greater than $200 million and an average ROE of greater than 25% for five consecutive years between 1976-82. He calculated the median ROE for the firms from 1976-78 and from 1989-93, and then compared these medians against the average ROE for firms on the S&P 400 (see Exhibit 3). Whereas the median ROE for the high-performing subgroup was 21% above the average ROE for the S&P 400 in 1976-82, it was only 2% above in the later period. Thus the high-performing firms' abnormal earnings had largely dissipated over the fifteen year interval.

Palepu *et al* (1996, pp. 5.4-5.7) report similar findings: abnormally high or low ROE's tend to revert to normal levels, roughly between 10-14%, often within five years and usually within ten years.[3] The reversion in ROE's is largely due to reversion in profit margins rather than reversion in asset turnover or leverage which remain relatively constant over time. The fact that advantage horizon lasts for five or ten years provides some justification for using five or ten-year projections in discounted cash flow analysis.

In another study, Ghemawat (1991) examined the returns on investment (ROI) for 692 business units from 1971-1980. After sorting the business units by their ROI in 1971, he divided the sample into two equal subgroups and calculated the average ROI for each subgroup over the next ten years. Initially, the top group had an average ROI of 39% compared to 3% for the bottom group. The 36% spread between the two groups decreased to less than 3% by the end of ten years: the average ROI for the top group had decreased to 21.5% while the average ROI for the bottom group increased to 18.0%.

While the evidence consistently shows that the advantage horizon is finite, firms like Coca-Cola, Wal-Mart, and Microsoft have been able to extend their advantage horizons for many years. These firms have been able to create tremendous value for shareholders by sustaining their ability to generate positive abnormal profits.

Reinvestment

The key insight from the model regarding investment is that reinvestment of earnings is value enhancing only when investment opportunities generate expected returns in excess of the cost of equity (ROE>K_E). Because investment opportunities vary across firms and vary over time for the same firm, it is impossible to make conclusive statements on the value of reinvestment. Nevertheless, there is some evidence that reinvestment creates value. Recent studies have shown that firms which announce major capital expenditure or research and development (R&D) programs experience positive abnormal equity returns.[4] The market interprets these announcements as good news and their stock prices usually increase. While it may be the case that firms announce only their most positive NPV investments, Fruhan (1979, Table 1-6) provides evidence from a sample of almost 1500 firms that broadly supports the relation among high profitability, high reinvestment, and high equity valuations.

Acquisitions represent another form of investment for many firms. Jensen and Ruback (1983) review the many studies on acquirer returns surrounding merger announcements. They conclude that, on average, acquirer shareholders do not lose and target shareholders gain from merger

[3] See also Freeman, Ohlson, and Penman (1982).
[4] McConnell and Muscarella (1985) analyze capital expenditure announcements while Chan, Martin, and Kensinger (1990) analyze R&D expenditure announcements.

announcements. Thus, acquisitions create net gains for both firms combined even though they do not increase acquirer shareholder value.

Jensen (1986, 1993) presents an opposing view. He argues that managers often overinvest, i.e. invest in negative net present value projects, especially when their firms generate substantial free cash flow. Their incentive to overinvest results from their compensation being tied, indirectly, to firm size which, in turn, is a function of the amount investment. They are able to over invest because internal control systems such as board oversight are weak. In the absence of effective internal control systems, external forces such as the market for corporate control discipline investment activity. Jensen cites the oil industry in general and the Gulf Oil takeover in particular as examples where takeovers eliminated wasteful capital expenditures. Just as investing in positive NPV projects creates value, so, too, does eliminating negative NPV investments.

Warren Buffet, the prominent investor and chairman of Berkshire Hathaway, acknowledged the problem of overinvestment in his company's 1984 annual report:

> Many corporations that show consistently good returns have, indeed, employed a large portion of their retained earnings on an economically unattractive, even disastrous, basis. Their marvelous core businesses camouflage repeated failures in capital allocation elsewhere (usually involving high-priced acquisitions). The managers at fault periodically report on the lessons they have learned from the latest disappointment. They then usually seek out future lessons. (Failure seems to go to their heads.). . . In such cases, shareholders would be far better off if the earnings were retained to expand only the high-return business, with the balance being paid in dividends or used to repurchase stock...

Although stated in his characteristically droll way, Buffet's point is clear: reinvestment destroys value unless it generates an appropriate risk-adjusted rate of return.

References

Bernard, Victor L, 1994, Accounting-Based valuation methods, Determinants of market-to-Book Ratios, and Implications for Financial Statement Analysis," University of Michigan working paper.

Chan, Su Han, John D. Martin, and John W. Kensigner, 1990, Corporate research and development expenditures and share value, *Journal of Financial Economics*, 26: 255-276.

Freeman, Robert, James Ohlson, and Stephen Penman, 1982, Book Rate-of-Return and Prediction of Earnings Changes: An Empirical Investigation, *Journal of Accounting Research*, Autumn: 639-653.

Fruhan, William F., 1996, How do we link together profitability, cash flows, capital cost, growth, and stock value?, Harvard Business School presentation.

Fruhan, William F., 1979, Financial Strategy: Studies in the Creation, Transfer, and Destruction of Shareholder Value, (Irwin, Homewood, IL).

Ghemawat, Pankaj, 1991, Commitment, (The Free Press, Macmillan, Inc., New York, NY).

Jensen, Michael C., 1986, Agency Costs of Free Cash Flow, Corporate Finance, and Takeovers, *American Economic Review*, 2:323-329.

Jensen, Michael C., 1993, The Modern Industrial Revolution, Exit, and the Failure of Internal Control Systems, *Journal of Finance*, 3:831-880.

Jensen, Michael C., and Richard S. Ruback, 1983, The market for Corporate Control: The Scientific Evidence, *Journal of Financial Economics*, 11:5-50.

Luehrman, Timothy A., Capital Projects as Real Options: An introduction, Harvard Business School Note N9-295-074 (Rev. March 22, 1995)

McConnell, John J., and Chris J. Muscarella, 1985, Corporate Capital Expenditure Decisions and the Market Value of the Firm, *Journal of Financial Economics*, 14: 399-422.

Myers, Stewart C., 1977, Determinants of Corporate Borrowing, *Journal of Financial Economics*, 5: 147-175.

Ohlson, J. A., 1995, Earnings, Book Values, and Dividends in Equity Valuation, *Contemporary Accounting Research*, 661-687.

Palepu, Krishna G., Victor L. Bernard, and Paul M. Healy, 1996, Business Analysis & Valuation Using Financial Statements, South-Western College Publishing (Cincinnati, Ohio).

Stewart, G. Bennett, III, 1991, The Quest for Value: The EVA Management Guide, (Harper Business, New York, NY).

Exhibit 1: Numerical example of the relation between the value drivers and value creation

Combining equations 10 and 11 yields the following equation:

Market/Book = $[(1+\gamma ROE) / (1+K_E)]^n + [ROE(1-\gamma) / (K_E - \gamma ROE)] [1 - ((1+\gamma ROE) / (1+K_E))^n]$

This Exhibit shows the hypothetical market-to-book ratios as a function of the three value drivers: profitability, advantage horizon, and re-investment.; assuming the firm has a cost of equity equal to 15%. The three cases differ by the level of reinvestment which varies from 0% to 66%.

Case #1: Reinvestment rate (γ) = 0%

	Return on Equity (ROE)		
Advantage Horizon	**5%**	**15%**	**25%**
5 years	0.66	1.00	1.34
15 years	0.42	1.00	1.58
30 years	0.34	1.00	1.66

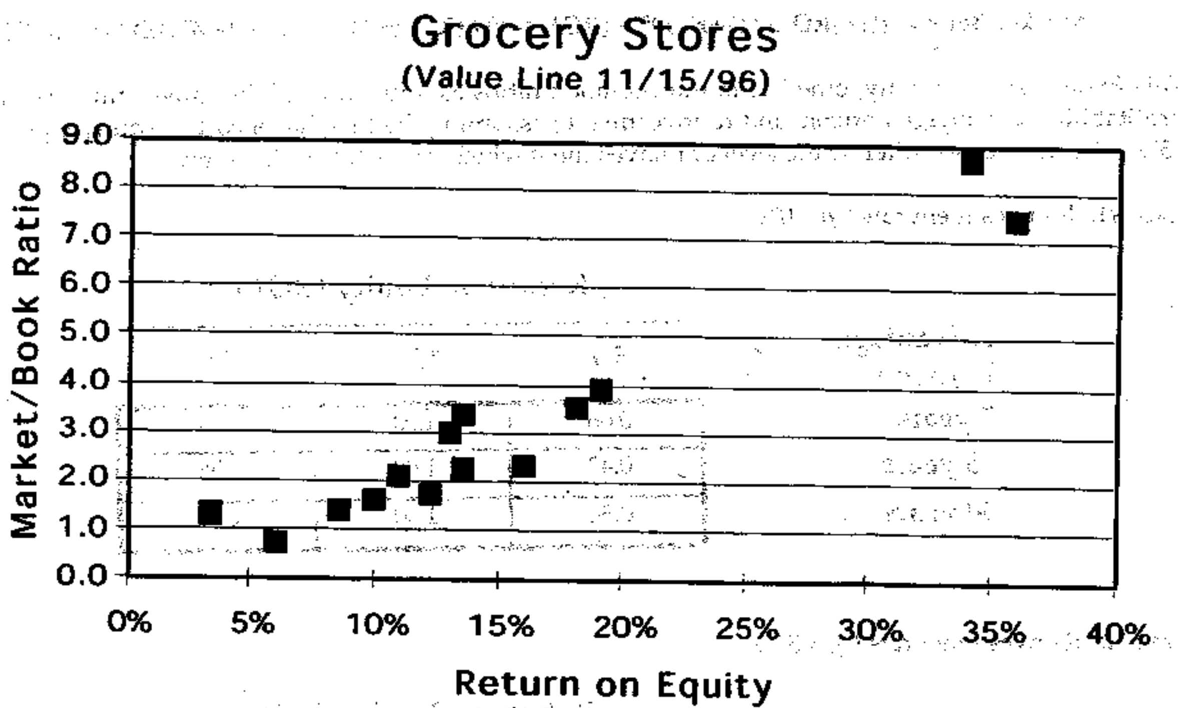
Case #2: Reinvestment rate (γ) = 33%

	Return on Equity (ROE)		
Advantage Horizon	**5%**	**15%**	**25%**
5 years	0.65	1.00	1.39
15 years	0.37	1.00	1.88
30 years	0.27	1.00	2.24

Case #3: Reinvestment rate (γ) = 66%

	Return on Equity (ROE)		
Advantage Horizon	**5%**	**15%**	**25%**
5 years	0.65	1.00	1.45
15 years	0.32	1.00	2.43
30 years	0.18	1.00	4.27

Exhibit 2: Relation between Return on Equity (ROE) and Market-to-Book Ratio



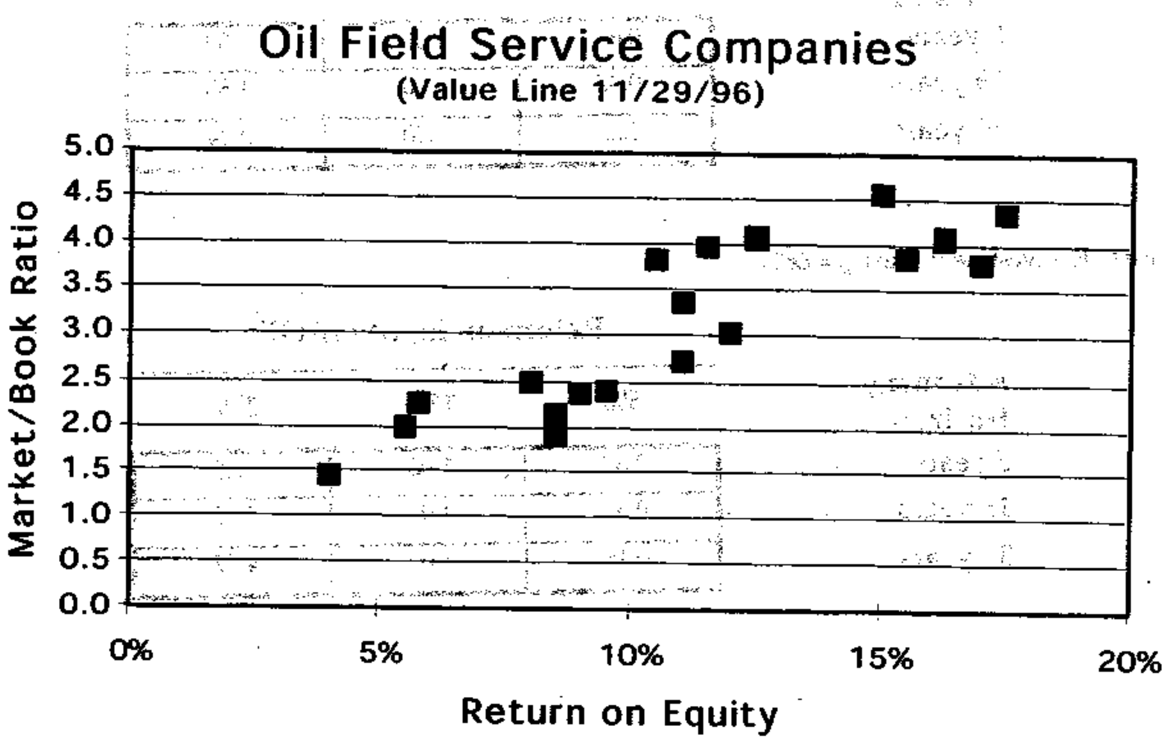


Exhibit 3: Advantage horizon

Fruhan (1995) analyzed the advantage horizon of a sample of 87 high-performing firms. To be included in the sample, firms had to have an average ROE of more than 25% for five consecutive years between 1976-82 and have sales greater than \$200 million. He found the following:

Top Performers:	**1976-78 Average ROE**
1. Petrie Stores	2.03
2. H&R Block	1.45
3. Standard Microsystems	1.43
4. Airborne Freight	0.77
5. Wendy's International	0.74
6. Commerce Clearing House	0.67
7. Avon Products	0.67
8. Southwest Airlines	0.63
9. Charming Shoppes	0.56
10. Loctite Corp.	0.56

For the period from 1976-78:

Median ROE for the top 87 firms	= 37%
S&P 400 Average ROE	= 15%
Spread	= 21%

For the period from 1989-93:

Median ROE for the top 87 firms	= 17%
S&P 400 Average ROE	= 15%
Spread	= 2%

Lesson: The advantage horizon is finite.

Appendix 1: Equity value and the advantage horizon

Equations 1 and 3 show that a firm's equity market value is a function of its return on equity (ROE) and cost of equity (K_E). Assuming no retention of earnings and constant returns, equity value is:

$$E_{MV} = ROE^*E_{BV} /(1+K_E) + ROE^*E_{BV} /(1+K_E)^2 + \ldots \quad \text{(A1.1)}$$

dividing through by the book value of equity (E_{BV}) yields

$$\text{Market/Book} = E_{MV} /E_{BV} = ROE/ (1+K_E) + ROE /(1+K_E)^2 + \ldots \quad \text{(A1.2)}$$

The ROE can be divided into two parts: ROE = (ROE - K_E) + K_E. The first term (ROE- KE) consists of "abnormal" earnings, returns to equity in excess of the cost of equity; the second term consists of "normal" earnings because that is the expected return on equity. Substituting back into equation A1.2 yields:

$$\text{Market/Book} = [ROE\text{-}K_E) + K_E]/(1+K_E) + [ROE\text{-}K_E) + K_E] /(1+K_E)^2 + \ldots \quad \text{(A1.3)}$$

$$\text{Market/Book} = (ROE\text{-}K_E)/(1+K_E) + ([ROE\text{-}K_E)/(1+K_E)^2 + \ldots$$

$$+ K_E/(1+K_E) + K_E/(1+K_E)^2 + \ldots \quad \text{(A1.4)}$$

Equation A1.4 is the sum of two geometric series, one of normal earnings and one of abnormal earnings. The present value of the normal earnings (using a perpetuity formula) is one:

$$1 = K_E/K_E = K_E/(1+K_E) + K_E/(1+K_E)^2 + \ldots \quad \text{(A1.5)}$$

The present value of the abnormal earnings depends on how long the firm expects to earn abnormal earnings. It can be thought of as an annuity: The firm receives a stream of abnormal earnings for a period of n years. The present value of an annuity can be written as:

$$\text{present value} = (ROE\text{-}K_E) * [(1/K_E) - (1/(K_E(1+K_E)^n)] \quad \text{(A1.6)}$$

Combining equations A1.5 and A1.6 yields:

$$\textbf{Market/Book} = \mathbf{1 + \{(ROE\text{-}K_E)^*[(1/K_E) - (1/(K_E(1+K_E)^n)]\}} \quad \text{(A1.7)}$$

as n approaches infinity, equation A1.7 reduces to equation 4 in the note.

Appendix 2: Equity value and reinvestment

This appendix derives a model of equity valuation as a growing perpetuity. Given a firm with a constant return on equity (ROE), it can either retain its earnings or pay them out to equityholders as dividends. Assuming the firms retains a fraction of earnings (γ) and pays out the remainder, then the market value of equity can be determined as follows.

Time	Total Earnings	Amount Paid Out (ECF)	Amount Retained	Book Value of Equity
t=0				E_0
t=1	ROE^*E_0	$(1-\gamma)^*\ ROE^*E_0$	$(\gamma)^*\ ROE^*E_0$	$E_1 = E_0 + (\gamma)^*\ ROE^*E_0$ $E_1 = E_0\,(1+\gamma ROE)$
t=2	ROE^*E_1 $ROE^*[E_0(1+\gamma ROE)]$	$(1-\gamma)^*\ ROE^*E_1$ $(1-\gamma)^*\ ROE^*\ E_0\,(1+\gamma ROE)$	$(\gamma)^*\ ROE^*E_1$ $(\gamma)^*\ ROE^*\ E_0\,(1+\gamma ROE)$	$E_2 = E_1 + (\gamma)^*\ ROE^*E_1$ $E_2 = E_1\,(1+\gamma ROE)$ $E_2 = E_0\,(1+\gamma ROE)^2$
t=3	ROE^*E_2 $ROE^*[E_0(1+\gamma ROE)^2]$	$(1-\gamma)^*\ ROE^*E_2$ $(1-\gamma)^*\ ROE^*\ E_0\,(1+\gamma ROE)^2$	$(\gamma)^*\ ROE^*E_2$ $(\gamma)^*\ ROE^*\ E_0\,(1+\gamma ROE)^2$	$E_3 = E_2 + (\gamma)^*\ ROE^*E_2$ $E_3 = E_2\,(1+\gamma ROE)$ $E_3 = E_0\,(1+\gamma ROE)^3$
t=4	(etc.)			
Growth Rate	γROE	γROE	γROE	γROE

Value = discounted present value of payouts (equity cash flows)

$$= \frac{((1-\gamma)^*\ ROE^*E_0)}{(1+K_E)} + \frac{((1-\gamma)^*\ ROE^*\ E_0\,(1+\gamma ROE)}{(1+K_E)^2} + \ldots \qquad \text{(A2.1)}$$

$$= \frac{((1-\gamma)^*\ ROE^*E_0)}{(1+K_E)} \quad \{1 + [(1+\gamma ROE)/(1+K_E)] + [(1+\gamma ROE)/(1+K_E)]^2 + \ldots\} \qquad \text{(A2.2)}$$

Equation A-2 is a growing perpetuity with growth rate equal to γROE. It can be rewritten as:

$$\text{Equity Value} = \frac{(1-\gamma)^*\ ROE^*E_0}{(K_E - \gamma ROE)} \qquad \text{(A2.3)}$$

After multiplying through by the book value of equity (E_0), one gets the ratio of equity at market value to equity at book value ($E_{MV}/E_{BV} = V/E_0$):

$$\textbf{Market/Book} = \frac{(1-\gamma)^*\ ROE}{(K_E - \gamma ROE)} \qquad \text{(A2.4)}$$

Netspar

Network for Studies on Pensions, Aging and Retirement

Netspar DISCUSSION PAPERS

Casper van Ewijk, Henri de Groot and Coos Santing

A Meta-Analysis of the Equity Premium

Discussion Paper 09/2010-050

A Meta-Analysis of the Equity Premium

Casper van Ewijk,[a,b] Henri L.F. de Groot[a,c,d] and Coos Santing[e]

[a] CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands
[b] University of Amsterdam, Amsterdam, The Netherlands
[c] Tinbergen Institute, Amsterdam-Rotterdam, The Netherlands
[d] Dept. of Spatial Economics, VU University Amsterdam, The Netherlands
[e] Ministry of Finance, The Hague, The Netherlands

Abstract

The equity premium is a key parameter in asset allocation policies. There is a vigorous debate in the literature regarding the actual measurement of the equity premium, its size and the determinants of its variation. This study aims to take stock of this literature by means of a meta-analysis. We identify how the size of the equity premium depends on the way it is measured, along with its evolution over time and its variation across regions in the world. We find that the equity premium is significantly lower if measured by ex ante methods rather than ex post, in more recent periods, and for more developed countries. In addition, looking at the underlying fundamentals, we find that larger volatility in GDP growth tends to raise the equity premium while a higher nominal interest rate has a negative impact on the equity premium.

Keywords: equity premium, meta-analysis

JEL codes: D53, E44, G12, N20

Useful comments by Clemens Kool, Peter Schotman, Bas ter Weel and Ed Westerhout are gratefully acknowledged. We are also grateful to Jan Luiten van Zanden for sharing historical interest rates and inflation rates (from International Institute of Social History) that underlie our analysis in Section 4.3. The usual disclaimer applies.

1. Introduction: The Equity Premium

The equity premium is a key parameter in asset allocation policies. It measures the excess return above the risk-free return and as such it can be seen as the price for risk. There has been a lively debate in the theoretical as well as the empirical literature on the measurement, size and sources of variation of the equity premium. In their seminal contribution, Mehra and Prescott (1985) identified the famous equity premium puzzle according to which there is a discrepancy between the equity premium as measured empirically and the premium that follows from standard theory. Mehra and Prescott calculated a historical equity premium of 6.2 percent in the United States for the period 1889–1978. Economic theory, based on the consumption capital asset pricing model (CCAPM), only justifies a premium up to a maximum of about 0.35 percent using conventional values for risk aversion. Their study initiated an intense debate in the scientific literature on the determination and size of the equity premium, both on the theoretical side (cf. Weil, 1989, Kocherlakota, 1996, Campbell and Cochrane, 1999, and many others) and on the empirical side of the puzzle. This paper focuses on the empirical aspects of the discussion, and aims to take stock of the existing literature by performing a meta-analysis of a wide selection of empirical studies on the equity premium, and to explain the sources of variation in this literature.

Meta-analysis provides us with a toolkit of statistical techniques enabling a quantitative review of the existing literature. As such, it complements narrative reviews.[1] Meta-analysis originated in the experimental sciences and was later on extended to fields such as the medical sciences where it has gained the status of a common practise instrument to merge results from different trials on the effectiveness of a specific drug or treatment. The research method has subsequently been introduced in psychology and education and is gradually gaining ground in economics (see, e.g., Florax et al., 2002, for an overview). Nowadays meta-analyses have been performed for a wide array of both microeconomic and macroeconomic issues. This study adds a new topic to the list which is at the heart of finance and also has close ties to macroeconomics.

Considering the empirics of the equity premium, four major issues stand out. First, the equity premium as measured from ex post stock returns proves to be quite sensitive to the observation period. This even holds for the long periods that are often used to identify the premium, which is obviously due to the large volatility of stock prices. This causes controversy on the 'true' value of the equity premium. For example, Siegel (1992) suggests that the high equity premium found by Mehra and Prescott (1985) was the result of the relatively low risk free rate in the period 1889–1978. Siegel found that the equity premium in this period is 4% higher than in the two decades just before and after this

---

[1] For good overviews of the literature, see Dimson et al. (2002) and Mehra (2008). See also Fernandez (2009a,b) for studies complementary to our meta-analysis which are based on a survey among professors and a review of information provided in 150 textbooks in finance.

period (viz. the periods 1880–1888 and 1979–1990, respectively). Including these adjacent periods would lower the equity premium by some 0.8% points.

A second, and related, controversy concerns the question whether the equity premium is constant over time. Several authors suggest that the equity premium is declining over time, especially since World War II (e.g., Blanchard, 1993, Siegel, 1999, Dimson et al., 2002), whereas others claim that the equity premium will continue to remain high (e.g., Mehra, 2003).

Third, the equity premium may vary across space. There is no strict need that the equity premium should be identical across countries and regions. Differences in stage of development leading to different aggregate risks, or differences in institutions leading to differences in leverage, could well explain different values of the equity premium. Moreover, as better time series tend to be available for the more successful stock markets, in particular the United States, this may have caused a bias in research as well. Jorion and Goetzmann (1999) conclude that the high equity premium obtained for U.S. equities could be the exception rather than the rule. Extending the data set to other markets – including the ones that did not survive – they find a lower estimate of the world rate of return on equity by 0.29 % points. Since that study the scope of research is broadened as more data become available for other countries. An important study in this respect is the *"Triumph of the Optimists"* by Dimson et al. (2002) who have calculated the equity premium for 17 countries over a period of 101 years.

A final issue is whether the equity premium should be measured ex post or ex ante. In ex post studies the equity premium is calculated as the difference in the mean return on stocks, either taken geometrically or arithmetically, and the risk free rate, mostly the short term interest rate (T-bills) or long term government bonds. This ex post approach is taken by Mehra and Prescott (1985) as well as many others (cf. Siegel, 1999, Dimson et al., 2002). Ex ante studies, in contrast, take the dividend yield or the earnings-price ratio as a starting point and derive the implied equity premium using an estimate for the capital gains. Seminal contributions here are Blanchard (1993), and Fama and French (1988, 2002) who found substantially lower estimates for the equity premium – ranging from 2.5% to 3% in the last study – than in most ex post studies.

After having addressed these issues, our analysis will be extended by looking at some fundamentals of the equity premium. First, we will have a closer look at the relationship between the equity premium and the interest rate and the rate of inflation. Next, we will investigate two underlying macroeconomic determinants. It is typically argued in the literature that the equity premium is higher in emerging markets than in mature markets (Shackman, 2006, and Erbas and Mirakhor, 2007). Investing in developing countries is generally perceived to be more risky, which has to be compensated in terms of a higher return. The stage of development of a country will be proxied by its Gross Domestic Product (GDP) per capita. Another macroeconomic factor that can influence the equity premium is the size of aggregate risk, here measured by the volatility of GDP growth. It is well known that higher volatility of consumption leads to higher required returns (Weil, 1989). In this vein Lettau et al. (2008) provide evidence that decreasing macroeconomic risk explains the boom of the stock markets in the 1990s. We will consider whether differences in the volatility of the economy indeed affect the equity premium. In this respect this study may contribute to the understanding of the impact of the credit crisis on the equity premium, even though the credit crisis itself is beyond the scope of this study (the most recent paper on the equity premium included in our meta-analysis being from 2008).

The remainder of this paper is structured as follows. Section 2 discusses several measurement issues, and identifies potential sources of variation in the equity premium. It thus paves the road for the selection of moderator variables to be employed in the meta-regression analysis. Section 3 describes the selection process of the primary studies of the meta-analysis and provides summary statistics of the explanatory variables. Section 4 discusses the results of the meta-regression, investigates the impact of structural underlying variables, and finally constructs benchmark values for the equity premium. Section 5 concludes.

2. How to measure the equity premium?

The literature on the equity premium provides no unanimity on how to measure the equity premium. In theory the equity premium represents the additional risk premium on equity relative to the return on safe assets. Or, more precisely the equity premium (*EP*) is defined as difference between the required return on equity ($\bar{r}_e$) and the risk free rate (r_f):

$$EP \equiv \bar{r}_e - r_f \,. \tag{1}$$

Assuming market efficiency, the required rate of return equals the expected rate of return (viz. $\bar{r}_e = E[r_e]$). There are a number of issues concerning the measurement of the equity premium. First and most fundamental, there is the difference between ex post and ex ante approaches to estimate the equity premium. Second, the choice of the market portfolio of stocks may matter for the height of the equity premium. In general, authors use a wide portfolio corresponding to well-established indices for official stock markets. Second, as purely safe assets do not exist in practice, one has to find a suitable proxy for the risk free rate. Third, there is a more technical issue of measuring returns as an arithmetic or geometric mean. Each of these issues is briefly discussed below.

Ex post or ex ante measurement of the equity premium

Mehra and Prescott (1985) measure the equity premium by calculating the historical return on stocks compared to the risk free rate. This 'ex post' approach is followed by many others (e.g., Dimson et al., 2002). It is not undisputed though. In particular, this method may be biased if the equity premium is not stationary during the observation period. Rising price earnings ratios over a prolonged period after World War II (up to the credit crisis) may point to a secular decline in the risk premium on equity. Indeed, building on Gordon's (1962) dividend discount model, Blanchard (1993) estimated that the equity premium in the United States had fallen to 2-3% in the early 1990s. Essentially, this 'ex ante' method takes the equity price as the present value of future dividends or earnings. Then, estimating future growth of earnings (dividends), one can calculate the equity premium implied in observed earnings to price ratio, or dividend to price ratio. Blanchard's finding of a declining premium was confirmed in other ex ante studies such as Jagannathan et al. (2000), and Fama and French (2002).[2]

The choice in method can thus have substantial consequences for the size of the equity premium. For the United States, Fama and French (2002) find that the ex-post equity premium for the period 1951–2001 is almost three times as high as the ex-ante estimate. In a stationary environment both methods, ex ante and ex post, are expected to converge in the long run. In a non-stationary environment, however, the outcome can differ for the two methods, even producing seemingly contradictory results (e.g., Lengwiler, 2004). This is because changes in the required rate of return produce just the opposite effect on the realised return through the revaluation of stocks. For this reason Dimson et al. (2002) warn not to extrapolate the high post-war returns into the future. As these high ex-post returns were caused by the revaluation of stocks due to a *fall* in the prospective rate of return, they rather point to low future returns.

Choice of market portfolio

Most authors measure the equity premium using the well-known stock market indices for a broad market portfolio, such as Standard and Poors for the United States and the MSCI for the developed countries. Usually midcaps are not included in the data. This may matter, as the equity premium depends on the risk profile of the companies, and also on the equity-debt composition in financing the firm. Higher risk and higher leverage imply higher returns on equity. As most authors use broad market portfolios, we will make no further distinction with regard to the portfolio in the meta-analysis. When using long time series one should furthermore be aware of the sensitivity of the results for survivorship of companies over time. If indexes are constructed by only including companies that are

[2] Early 'ex ante' studies focused on the equity premium *per se*. Others have extended this framework by allowing the projected growth of dividends and earnings to depend on other variables. This leads to the so-called conditional model of the equity premium, as distinct from the unconditional model employed by, for example, Fama and French (2002). Claus and Thomas (2001) use several accounting variables to do this. Earlier, Blanchard (1993) used the unconditional dividend model, but took account of expectations of the interest rate and inflation rate.

present today, a bias is created since companies that went bankrupt are excluded by construction (Brown et al., 1995). However, the general idea is that survivorship bias in stock market returns is small. In our meta-analysis we will therefore neglect the potential influence of 'survivorship bias'. However, Jorion and Goetzmann point out that there may exist a survival bias across stock markets as well, as existing long data series tend to focus on markets that have been successful up to date. Also, time series often break down during deep crises such as wars and revolutions. Indeed, the very focus in research on the most successful stock market, *viz.* the United States, may lead to a significant bias. Constructing data for other stock markets Jorion and Goetzmann show that U.S. equities have the highest return over the period 1921–1996, at 4.3%, versus a mean return for other countries in the sample of only 0.8%. Taking the average of all countries, including these other markets, lowers the world market return by 0.29% points relative to the U.S. return.

Risk free rate

The second important measurement issue concerns the choice of the risk free rate. In theory, a risk free asset should deliver an income flow in real terms that is independent of the state of the world (Lengwiler, 2004). Unfortunately such an asset does not exist. Government paper comes closest, as it has low default risk.[3] Therefore, most studies on the equity premium use the return on short term treasury bills or long term bonds as a proxy for the risk free rate. A disadvantage of such assets is that their real return depends on inflation. Inflation-indexed governments bonds do exist, but are only recently available. Economists therefore prefer treasury bills (T-bills) or notes with a short time to maturity, as they are less sensitive to inflation and interest rate risk. Others, however, prefer long term bonds, as this is more in line with the long-term character of equity.[4] The impact of the risk-free asset against which the equity premium is determined will be identified in the meta-analysis by using a dummy indicating whether the risk-free rate is proxied by T-bills (short-term) or long-term bonds.[5]

Arithmetic versus geometric measurement of mean returns

Using historical time series, the return on equity can be calculated as a geometric mean (*GR*) or an arithmetic mean (*AR*). The difference relates to the way in which series of returns are averaged over time. If returns are measured arithmetically, the average is taken as the sum of the returns per period divided by the number of periods. If returns are measured geometrically this is calculated as the

---

[3] In deep crises, such as wars and revolutions, also governments may default on their liabilities. For this reason Jorion and Goetzmann (1999) focus on real equity return, that is the return relative to commodities, rather than on the equity premium which measures the return relative to government debt.

[4] Recently, some work is being done on the term structure of the equity premium (cf. Lemke and Werner, 2009). In this meta-analysis we will take account of the term of the risk free rate, but ignore potential differences in the equity premium arising from a term structure as knowledge on this is still pre-mature.

[5] See Dimson et al. (2007) for an extensive discussion on the impact of maturity of the risk free rate on the equity premium.

compound rate of return (Derrig and Orr, 2003). Arithmetic returns tend to be higher than the geometric returns. With lognormal returns the expected geometric return (*GR*) converges to the expected arithmetic return minus half the variance, that is $GR = AR - ½\,\sigma^2$ (see, e.g., Welch, 2000, Dimson et al., 2002, and Ibbotson and Chen, 2002). The arithmetic mean is generally considered to produce the best estimate of the mean return; the geometric mean approximates the median return rather than the mean (Campbell et al., 1997, Jacquier et al., 2003, and Ten Cate, 2009). In the meta-regression model the difference between the arithmetic and geometric return is captured by a simple dummy variable.

3. Data Sources and Summary Statistics

This section describes the selection of the studies that are used in our meta-analysis, and provides a brief characterization of the database by some descriptive statistics. The formal meta-regression model and its results will be presented in the next section. The equity premium puzzle that was identified by Mehra and Prescott (1985) resulted in a flood of studies on the equity premium, both theoretical and empirical. We focus on the empirical studies. To construct the database for the meta-analysis, we started using the search engine Econlit covering published articles in English in academic journals.[6] The keywords used for our search were 'equity premium'. This resulted in 242 hits of which 15 studies measure the size of the equity premium. Using the technique of snowballing (see, for example, Cooper and Hedges, 1994), nine other studies were found which were added to the database. We are thus left with 24 studies that form the heart of our meta-analysis. Each study reports several equity premiums, covering different time periods, countries and methodologies.[7] The resulting database consists of 535 observations. Appendix A provides a list of all studies and their summary statistics. The studies are also clearly marked in the list of references.

Clearly, the database is not balanced across the spatial and time dimension. In the spatial distribution, there is a bias towards developed countries, in particular the United States. Over the past couple of years, however, the sample of countries for which equity risk premiums are available has increased substantially due to, for example, studies by Dimson et al. (2002), Shackman (2005), and Salomons and Grootveld (2003). In total, our database includes 44 countries. Almost half of the observations (256) refer to the United States. For many other countries, there is only a couple of observations available. We therefore combine these countries into relatively homogeneous regions, viz. Canada, Oceania (Australia, New Zealand and Japan), Canada, Western Europe, Advanced Emerging Countries (including amongst others Brazil, Mexico, Poland and South Africa), Secondary

---

[6] Econlit American Economic Association's electronic bibliography contains 750 journals since 1962 (see www.econlit.org).
[7] There are studies reporting premiums covering a broad time span as well as premiums for sub-periods within this broad time span. In these cases, the former is omitted from the analysis to avoid double counting.

Emerging Countries (including amongst others Argentina, China, India, Turkey), and the Asian Tigers.[8]

Across the temporal dimension there is a bias towards more recent periods. Some studies cover a long time span of almost two centuries (from 1830 to present), but most studies cover more recent periods. About 9% of the observations is characterized by a mid-year before 1900. About 13% has a mid-year that falls in the period 1900–1950. For the remaining 78%, the mid-year is 1950 or later.[9] Concerning the way of measurement, over 80% of the observations measure the equity premium on an ex post basis. Furthermore, the majority concerns equity premiums that are measured arithmetically (354 compared to 181 on a geometric basis).[10] Finally, of the 535 observations, 310 are calculated with T-bills or closely related substitutes. The other 225 equity premiums are calculated with bonds proxying for the risk free asset.

3.1 Descriptive analysis of the data

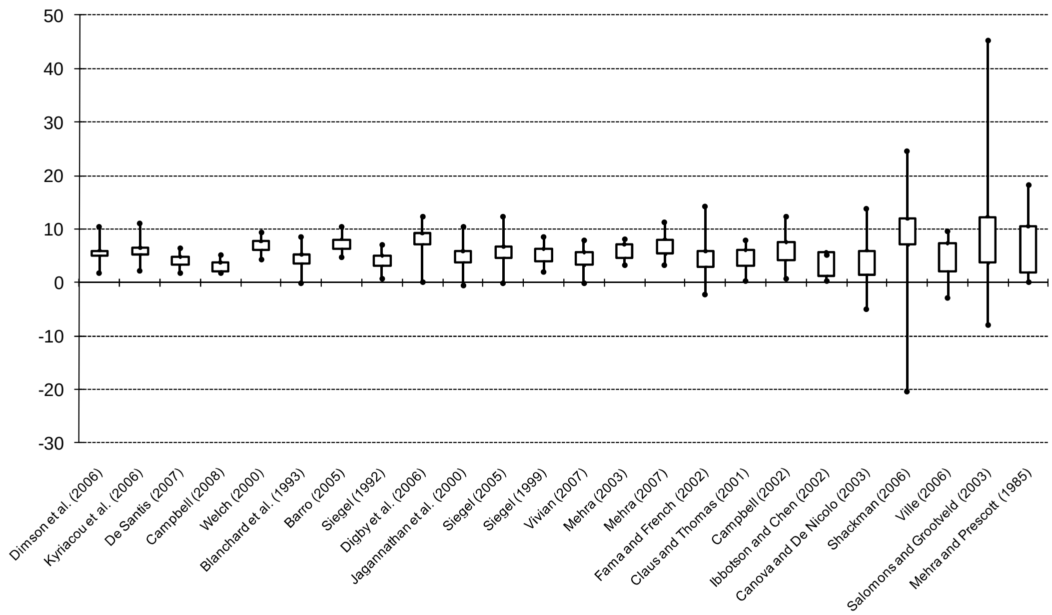
The within-study distribution of the observations is presented in Figure 1. For each individual study it gives the minimum and maximum value of the equity premium along with a 95% confidence interval.[11] The primary studies are ordered according to the within-study variation measured by the size of the 95% confidence interval.

According to Figure 1, some studies in the meta-analysis report negative equity premiums (viz. Blanchard et al., 1993, Canova and Nicolo, 2003, Digby et al., 2006, Fama and French, 2002, Jagannathan et al., 2000, Salomons and Grootveld, 2003, Shackman, 2006, Siegel, 2005, Ville, 2006, and Vivian, 2007). There are also very large equity premiums as is the case for the study by Salomons and Grootveld (2003). We see large differences for the within-study variation of the equity premium. For Dimson et al. (2006), the lower bound of the 95% confidence interval is 5.0% and the upper bound is 6.0%. In contrast, for Mehra and Prescott (1985) the lower bound is 1.9% and the upper bound is 10.5%.

---

[8] Further details on country groupings are available upon request.
[9] The mid-year is the average of the initial and final year of the period covered by the observation.
[10] If studies do not report the method to calculate returns the arithmetic one is assumed. We performed a robustness check to investigate the sensitivity for this assumption. Details are available upon request from the authors.
[11] The confidence interval of the mean is equal to the within study mean plus or minus two times the within study standard-deviation divided by the square root of the number of observations.

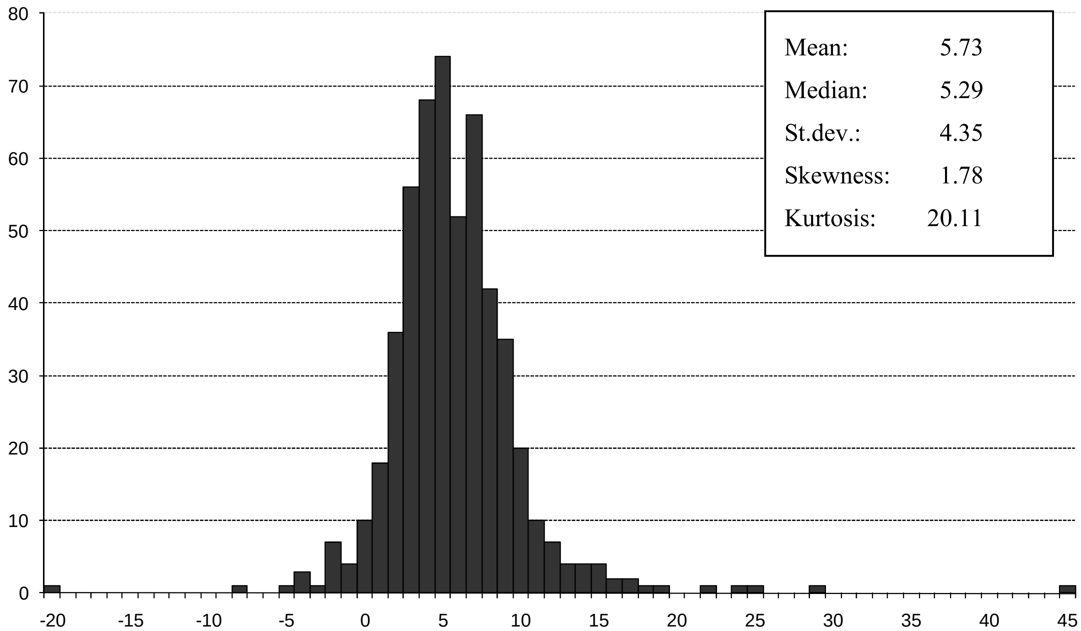
Figure 1. Within- and between-study variation of the Equity Premium



Note: lines indicate minimum and maximum EP's found in the respective studies. The boxes indicate a 95% confidence interval around the mean of the respective studies.

Figure 2 further describes the distribution of the equity premium for the entire sample of 535 observations. The mean is 5.73. The null-hypothesis of a normal distribution is clearly rejected (*p*-value <0.001). There are 24 observations with a negative equity premium, whereas 48 observations have equity premiums exceeding 10%.

Figure 2. Histogram the Equity Premium



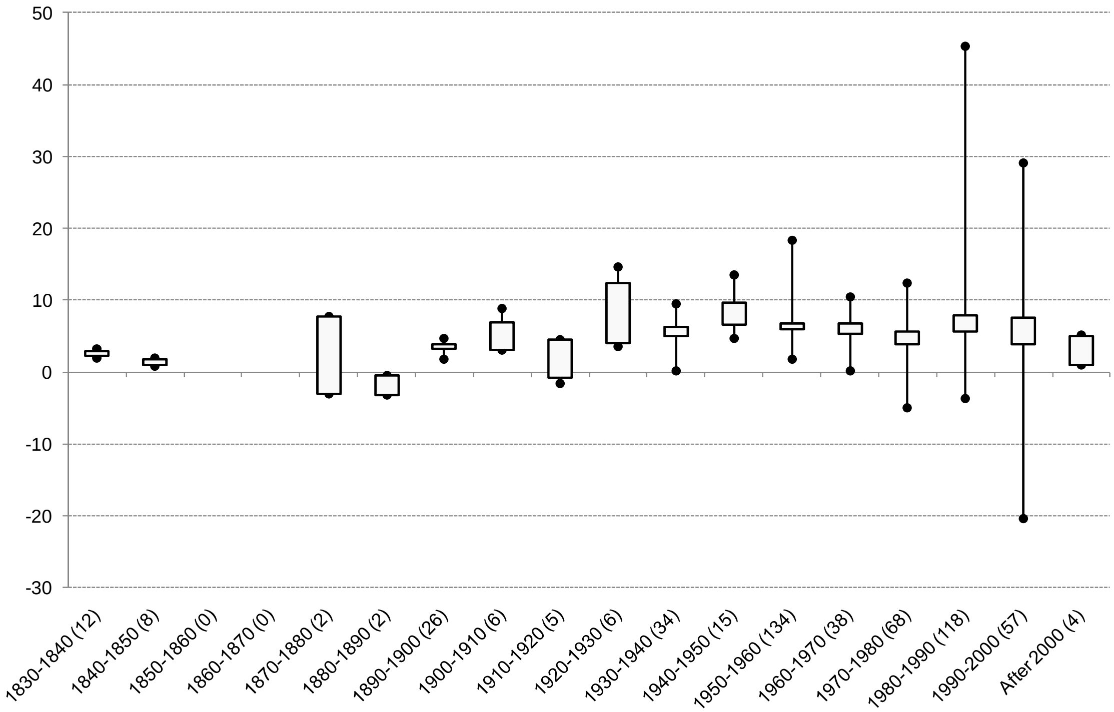
Time Variation

Figure 3 gives an impression of the temporal variation of the equity premium. More precisely, each observation is expressed for the mid-year of the period on which this observation is based. This figure confirms the overall picture that the equity premium was low until 1920, high in the 1920s and again high in the post war period. Short term deviations from this overall pattern are observed in the 1970s (with a dip and a recovery thereafter). The recent crisis on the financial markets falls beyond the scope of all studies included in the sample.[12]

---

[12] It should be noted that this is not a complete representation of the variation of the equity premium over time. As the data points refer to the mid year of observation periods with different lengths, the evolution of the equity premium is smoothed. Restricting the dataset to only observation periods of 10 years or less, shows a similar pattern but with greater volatility. Looking at the length of the period studied in somewhat greater detail, we can distinguish several categories, viz. 0–10 years (123 observations), 11–20 years (66 observations), 21–30 years (79 observations), 31–50 years (51 observations), 51–100 years (110 observations) and more than 100 years (106 observations). In our database, there are no observations based on periods shorter than 5 years or longer than 203 years. Further details on the impact of differences in the length of the observation period are available upon request from the authors.

Figure 3. Variation over time in the equity premium by mid year of the observation period

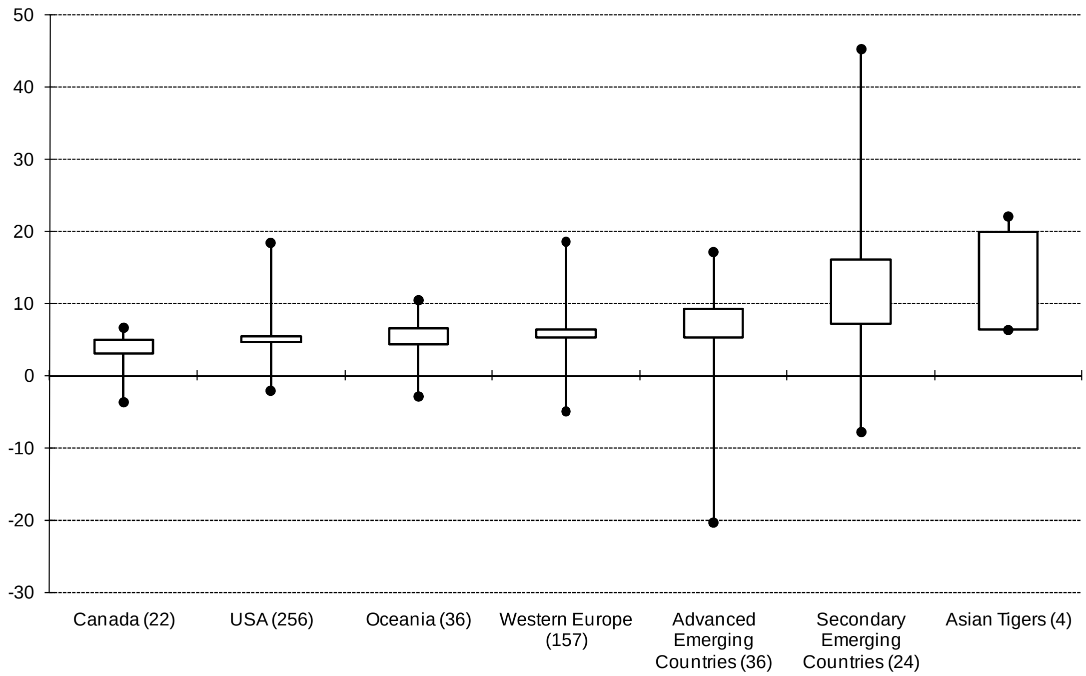


Note: lines indicate minimum and maximum EP's found in the respective periods. The boxes indicate a 95% confidence interval around the mean of the respective regions. The number of observations for each period is indicated in brackets.

Spatial Variation

The equity premium also varies considerably over space as is shown in Figure 4. To obtain a more balanced set, some countries are grouped into relatively homogeneous groups. We find that the equity premium is relatively high in emerging countries. The lowest average equity premium is found in Canada, and the highest is found for the Asian Tigers. The mean of the equity premium for these groups of countries varies from 3.95 percent in Canada to 13.14 in the Asian Tigers.

Figure 4. The Equity Premium by Country or Region



Note: lines indicate minimum and maximum EP's found in the respective regions. The boxes indicate a 95% confidence interval around the mean of the respective regions. The number of observations for each region is indicated in brackets

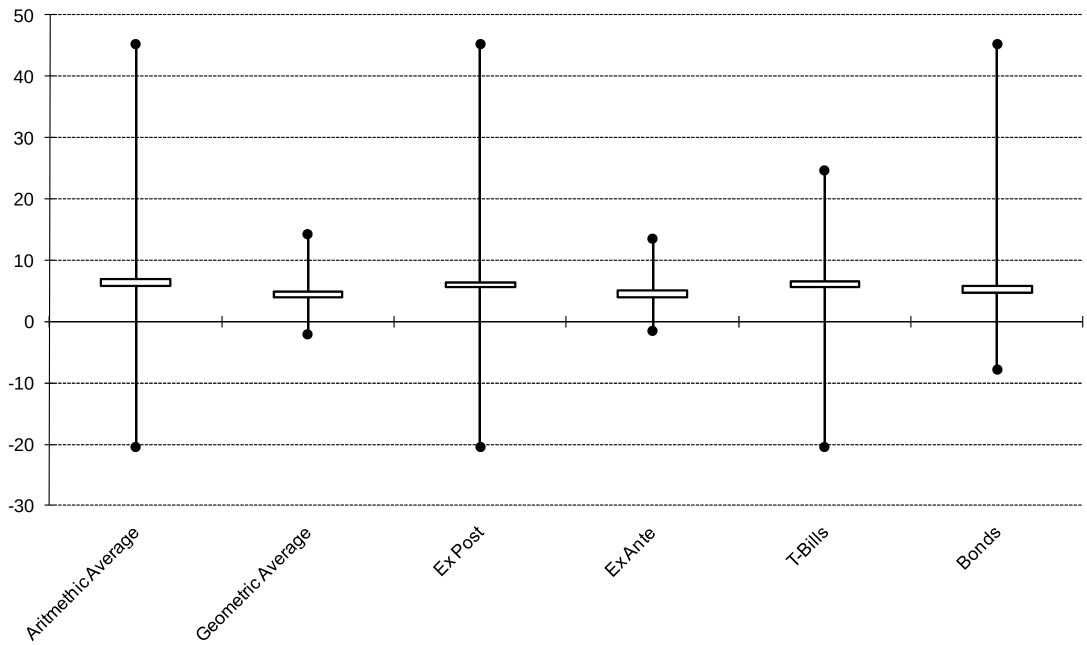
Variation in Method

Finally, Figure 5 illustrates the variation in the equity premium due to differences in definition of method of measurement. The mean of the observations calculating an arithmetic average is 6.37% whereas the mean of the observations calculating a geometric average is 4.46%. This is in line what might be expected on the basis of the variance in the series (see Section 2).[13] The second measurement issue is whether the equity premium is measured ex-ante or ex-post. As was explained in Section 2, the ex ante approach tends to produce lower estimates. This is confirmed by Figure 5. The average mean for the ex-post equity premium is 6.03%, whereas the mean of the ex-ante equity premium is 4.48%, a gap of 1.55% points which is in line with half the variance. Finally, the results for the equity premium depend on the proxy for the risk free rate. The mean of the equity premium calculated with T-bills as risk free rate is 6.07%, whereas the mean with bonds as risk free rate is 5.26%, a difference of 0.81% points.

---

[13] For a few observations it is unknown whether the mean is arithmetic or geometric. We have reckoned these to be arithmetic. Alternatively, if these observations with unknown method were assumed to be geometric the mean of the equity premiums with an arithmetic average is 6.59% and the mean of the equity premium with the geometric average is 4.98%. The difference in between measurement methods would then decrease from 1.8% to 1.6%.

Figure 5. Equity Premiums according to Method



Note: lines indicate minimum and maximum EP's found using the respective methods. The boxes indicate a 95% confidence interval around the mean for the respective methods.

To conclude this section, we present in Table 1 the simple correlations between the equity premium and the main explanatory variables. As to be expected, the equity premium tends to be higher in studies that use the arithmetic mean, the ex post method and the short term interest rate.

Table 1. Simple correlation matrix for equity premium and methods (N=535)

	Equity Premium	Arithmetic mean	Ex Post	T-Bill
Equity Premium	1.00	0.21	0.14	0.09
Arithmetic mean	0.21	1.00	0.07	–0.12
Ex Post	0.14	0.07	1.00	0.16
T-Bill	0.09	–0.12	0.16	1.00

4. The Meta-Regression Analysis

In this section, we turn to a meta-regression analysis to identify the (conditional) effects of the moderator variables on the equity premium. First, we present the basic meta-regression model and discuss its results. Then we extend the model including underlying fundamentals of the equity premium to get better insight into what explains the variation of the equity premium over time and across regions. Finally, we quantify benchmark values for the equity premium on the basis of the data set in this study.

4.1 The Meta Regression Model

The factors that may cause variation in the equity premium were identified in the previous sections. We will estimate meta-regression models that allow us to identify the contribution of these factors to the observed variation in the equity premium. For this purpose, we use the Huber-White estimator. This estimator simultaneously corrects for heteroskedasticity and cluster autocorrelation (see Williams, 2000, and Wooldridge, 2002, Section 13.8.2). The advantage of this estimator is that it accounts for the pooled data set-up by allowing for different variances and non-zero co-variances for clusters of observations taken from the same study.[14] More specifically, we postulate the following simple model:

$$EP_i = \alpha_0 + \sum_k \alpha_k Z_{ik} + \varepsilon_i \qquad (2)$$

where EP is the equity premium derived from the primary studies (indexed i= 1,2 ....., L) – as defined in equation (1) – and Z are the explanatory variables (indexed k= 1,....., K). The effect of the explanatory variables is measured by the regression coefficients α_k. The explanatory factors that we consider are (i) characteristics of the methodology used to derive the equity premium; (ii) temporal sources of variation; (iii) spatial sources of variation; and (iv) characteristics of the economy.

---

[14] Dependence may also occur for estimates from the same country or time period. Robust standard errors accounting for spatial or temporal dependence of the observations are presented in Appendix B.

The first three sets of factors will be central in the Section 4.2 in which we present the basic model. The three method variables (arithmetic versus geometric, ex post versus ex ante, and the use of treasury bills versus bonds) that we consider in our basic specification are easily captured by a dummy variable because each of them only has two categories. For the observation period, we include two dummy variables characterizing (i) the mid year to which the observation pertains and (ii) the length of the period covered by the observation. Regarding spatial variation, we include dummies for the countries and regions distinguished. Section 4.3 elaborates on this basic model by adding underlying fundamental determinants of the equity premium.[15]

4.2 Basic results

Table 2 describes the results of our base model in which we consider the impact of research method, and spatial and temporal factors. In the base specification (0) we only include the dummy variables capturing variation in methods. In specification (1), we also consider spatial variation, and we make a distinction between three different time periods.[16] All three methodological variables in specification (1) have a statistically significant impact on the equity premium. Equity premiums with an arithmetic average are on average 1.37% larger than equity premiums with a geometric average. This is fairly close to the 1.28% estimate reported as an average in Dimson et al. (2002).

The economic significance of the other methodology variables is somewhat smaller, but still substantial. Equity premiums that have been measured ex-post are on average 1.31% higher than equity premiums that are measured ex-ante. The size of this effect is comparable to other studies: Salomons (2008) estimates a difference between ex post and ex ante measurement of 1.08% for the United States in the period 1871–2003, and Madsen (2004) estimates a difference of 3% for the major industrialised countries in the period 1878–2002. The use of T-bills as risk free rate results on average in a 0.81% higher equity premium than the use of bonds as risk free rate. This is slightly higher than the 0.5% found by Dimson et al. (2002).

The country dummies capture differences in the equity premium relative to the United States which is taken as our benchmark country. The country effects for Canada, Secondary Emerging Countries and Asian Tigers are statistically significant. On average, an equity premium in Secondary Emerging Countries is 5.25% higher than in the United States and 6.60% in the Asian Tigers. In

---

[15] A distinctive feature of this meta-analysis is that the equity premium is often calculated rather than estimated. This implies that we cannot apply standard practice in most meta-analyses which is to weight observations with the standard error of the estimate in order to correct for variation in the precision or accuracy of observations. In our basic model we will not apply any weighting of observations. As it could be argued that the variance decreases with the number of observations, and thus with the length of the observation period, we have by means of robustness check also applied a weighting scheme based on the square root of the length of the observation time period (T). This hardly affects the results that we present. Further information is available upon request from the authors.

[16] The two specification tests indicate that the model is correctly specified. The White test and Breusch-Pagan test present evidence for heteroscedasticity of the error term of the equity premium, as has been expected.

contrast, Canada faces an equity premium that is 1.72% lower than in the United States. Equity premiums in Oceania, Western Europe, the Advanced Emerging Countries are not statistically different from those in the United States. Economically the magnitude of equity premiums which are calculated in emerging countries is very large, suggesting that the excess return for risky assets is substantially larger in those countries.

Table 2. Equity premium: base model

	Spec. 0	Spec. 1	Spec. 2	Spec. 3
Constant	2.94***	4.00***	4.10***	3.84***
	(0.44)	(0.62)	(0.59)	(0.66)
Arithmetic mean	1.96***	1.37***	1.42***	1.41***
	(0.45)	(0.29)	(0.33)	(0.33)
Ex Post	1.22***	1.31***	1.05***	1.17***
	(0.32)	(0.26)	(0.30)	(0.40)
T-bill used	0.89*	0.81***	0.92***	0.89***
	(0.50)	(0.29)	(0.26)	(0.25)
Region effects (relative to USA)				
Canada		–1.72***	–1.65***	–1.60***
		(0.50)	(0.48)	(0.51)
Oceania		–0.53	–0.64	–0.69
		(0.74)	(0.63)	(0.68)
Western Europe		–0.03	–0.22	–0.17
		(0.52)	(0.64)	(0.66)
Advanced emerging		1.17	1.31	1.39
		(0.85)	(0.86)	(0.88)
Secondary emerging		5.25***	5.95***	5.93***
		(0.43)	(0.74)	(0.75)
Asian Tigers		6.60***	7.11***	7.06***
		(2.23)	(2.01)	(2.02)
Period effects (relative to 1910–1950)				
Before 1910		–3.54***	–3.46***	–3.38***
		(0.58)	(0.57)	(0.51)
After 1950		–0.74	0.16	0.29
		(0.66)	(0.62)	(0.57)
Trend after 1950			–0.04**	–0.05*
			(0.02)	(0.02)
Length of period < 40 years				0.42
				(0.63)
# observations	535	535	535	535
R^2	0.07	0.21	0.22	0.22

Note: cluster robust standard errors corrected for within-study dependence are reported in parentheses. Statistical significance of the estimated coefficients is indicated by ***, ** and * referring, respectively, to the 1%, 5% and 10% significance level. Appendix B provides a more detailed cluster analysis taking account of dependence by country/region and time period.

Regarding variation over time, we find that the pre-war period (before 1910) was characterized by a substantially lower equity risk-premium than the period 1910–1950. A similar conclusion was drawn by Dimson et al. (2006) and Siegel (1992). The number of observations in the 19th century is, however, limited. In the second specification, we extend the basic specification (1) by allowing for a time trend in the equity premium in the post-war period. The results reveal that this trend is significantly negative, resulting in an annual decline of the equity premium by 0.038% points (cumulating to 0.94 % in 25 years). Apart from some variation in the size of the coefficients, the qualitative results described in specification (1) are unaffected by the inclusion of the time trend.

In specification (3) in Table 2, we look at the effect of the length of the observation period by including a dummy for shorter periods (0–40 years). Although positive, the effect is statistically insignificant. Inclusion of the effect hardly affects the other results. We will therefore take specification (2) as our basis model in the remaining.

4.3 Underlying fundamentals

Going one step beyond the standard meta-analysis we will also explore some underlying economic fundamentals of the equity premium. Therefore we extend the previous analysis by adding some underlying explanatory variables which may be relevant to the equity premium. This provides us with a more substantive way of identifying sources of variation and can enhance the understanding of the deeper determinants of observed variation over time and space. Specifically, we look at the impact of volatility of income, the stage of development of the country, the interest rate and inflation.

Both the stage of economic development and income volatility can influence the price of risk underlying the equity premium. The stage of development can be regarded as a proxy for the maturity of financial markets in the country or region at hand. In general, mature markets offer better opportunities for spreading risks, and could therefore lead to a lower equity premium (cf. Levine et al., 2006). Volatility is taken as an indicator for the size of risk in the economy. It is well established that equity returns tend to be higher in periods of high volatility in stock markets (cf. Lettau et al., 2008). Here we include the volatility in GDP as the underlying explanatory variable.

These additional variables are not directly available in the studies on the equity premium in our sample. We therefore have to revert to other sources. The stage of economic development can be proxied by Gross Domestic Product (GDP) per capita. The database of Maddison (2007) provides information on GDP per capita for many countries and over a long time period. The benchmark year of the database is 1990 and GDP is measured in Geary-Khamis dollars. These Geary-Khamis dollars convert local currencies into international dollars by using purchasing power parity rates. For each observation, GDP per capita is measured at the mid-year of the period for each observation of the equity premium. Information on GDP per capita could be obtained for 500 observations (the Maddison data are only available for periods after 1870). The lowest GDP per capita is observed in India, Pakistan, the Philippines and Indonesia. The United States has the highest GDP per capita. There is not only variation across countries but also over time. The GDP per capita in the United States was \$2,570 in 1876 and increased to \$28,347 in 2001. The degree of uncertainty in an economy is measured by the variance of the economic growth (GDP) for the period of observation. Doing this we are able to construct GDP variances for 494 of our observations. The largest variance is found for the 1940s for the United States. For the period of the 'great moderation' in the 1990s, the variance of economic growth is lowest, again in the United States. Table 3 describes the partial correlations between the variables. This shows a positive covariance of the equity premium and volatility, and negative covariance with GDP and inflation. Furthermore, the strong correlations between volatility and the interest rate, volatility and GDP, and the interest rate and inflation stand out.

Table 3. Simple correlation matrix equity premium and economic variables (N=460)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1) Equity Risk Premium	1.00	0.18	0.17	0.15	0.22	−0.11	−0.21	0.01
(2) Arithmetic mean		1.00	0.04	−0.14	−0.07	−0.07	0.07	0.11
(3) Ex Post			1.00	0.19	0.14	−0.20	0.03	0.10
(4) T-Bill				1.00	-0.04	−0.02	0.15	0.13
(5) Log(business cycle)					1.00	−0.59	−0.56	−0.13
(6) Log(GDP per capita)						1.00	0.16	−0.04
(7) Interest							1.00	0.58
(8) Inflation								1.00

The results of our regression analysis are presented in Table 4. For reference, specification (0) reiterates our basic model in the previous analysis, viz. specification (2) in Table 2, here taken for the comprehensive data set including GDP as well as interest rates and inflation. Specification (1) includes volatility measured as the variance of economic growth and GDP per capita. The number of observations decreases slightly as compared to the basic specification presented in Table 2 due to missing data for periods before 1870. The effect of the variance of economic growth is statistically significant and has the expected positive effect. The impact is substantial: an increase in volatility by 1 standard deviation leads to a 1.7%-point higher equity premium. The effect of GDP per capita is positive, but statistically only marginally significant. This is largely caused by the fact that region-dummies have been included. These pick up a large part of the impact of GDP per capita. Omitting the region-dummies results in a statistically significant negative effect of GDP per capita (see also the partial correlations in Table 3). The coefficients of the other explanatory variables are comparable to those in the basic specification in Table 2.

Table 4. Equity premium, model including economic variables

	Spec. 0	Spec. 1	Spec. 2	Spec. 3
Constant	4.02***	–23.78*	5.09***	–6.99
	(0.71)	(11.76)	(0.77)	(6.11)
Arithmetic mean	1.22***	1.35***	1.26***	1.20***
	(0.29)	(0.30)	(0.33)	(0.31)
Ex Post	1.35***	1.00***	1.33***	1.37***
	(0.31)	(0.32)	(0.31)	(0.34)
T-bill used	0.82**	0.97***	1.13***	1.05***
	(0.36)	(0.32)	(0.30)	(0.30)
Canada	–1.75***	–1.32***	–1.11**	–0.90*
	(0.49)	(0.43)	(0.45)	(0.44)
Oceania	–0.45	0.90	–0.85	–0.09
	(0.73)	(0.77)	(0.66)	(0.51)
Western Europe	–0.31	1.22	–0.001	0.73
	(0.45)	(0.97)	(0.60)	(0.89)
Advanced emerging	1.51	4.44***	3.46***	6.42***
	(0.97)	(1.51)	(1.14)	(1.75)
Secondary emerging		8.28***		
		(1.39)		
Asian Tigers		7.25***		
		(2.12)		
Before 1910	–2.46***	–0.29	–1.73***	–0.68
	(0.70)	(1.00)	(0.58)	(0.51)
After 1950	–0.68	–0.34	0.88	0.80
	(0.71)	(0.47)	(0.52)	(0.53)
Volatility (log var GDP)		1.49***		0.60**
		(0.43)		(0.25)
GDP per capita (log)		2.51**		1.14*
		(1.15)		(0.62)
Nominal interest rate			–0.53***	–0.52***
			(0.13)	(0.14)
Inflation rate			0.03	–0.02
			(0.15)	(0.17)
# observations	438	493	460	438
R^2	0.13	0.25	0.26	0.28

Note: cluster robust standard errors corrected for within-study dependence are reported in parentheses. Statistical significance of the estimated coefficients is indicated by ***, ** and * referring, respectively, to the 1%, 5% and 10% significance level. The dummy for Secondary Emerging Countries and the Asian Tigers is omitted in specifications (3) and (4) because of lacking data. For comparison, specification (0) uses the specification in Table 2 using a sample of observations that is equal to the sample underling specification (3).

Specification (2) considers the impact of the nominal interest rates and inflation.[17] Since interest rates are not available for the Secondary Emerging Countries and the Asian Tigers, these had to be omitted from the sample. Nominal interest rates are clearly negatively associated with the equity premium. A one percent increase in the interest rate leads to a half percent decline in the rate of return on equity. The result for inflation reported in specification (2) is statistically and economically insignificant.

Finally, specification (3) includes all economic indicators in one equation. The previous results stand upright. Also here we find a positive impact of GDP per capita which captures the variation of GDP per capita within the groups of countries that are distinguished by the dummies. Again, omitting all country and region dummies would alter this result and produce a negative association.

These results have been tested for their robustness. Instead of the volatility of GDP we also considered an alternative measure of macroeconomic uncertainty, viz. the fraction of economic downturns during the observation period. This variable is not statistically significant, and as the number of observations drops also the significance of other variable deteriorates as well. Also for the stage of economic development we looked at other – more direct – indicators, such as market capitalization and credit to the private sector. Market capitalisation ratios are available in the databases of Levine et al. (2006) and the World Development Indicators (World Bank, 2006). The data are available for almost every country but the time period is limited. For WDI, the period is restricted to 1988–2006 and for Levine to 1976–2006. The sample of observations for which this information can be used is thus relatively small. Credit to the private sector is available in the database by Levine et al. (2006) for the period 1960–2005. Using these data we are left with 285 observations. The lowest amount of credit to the private sector relative to GDP is measured for Venezuela, Argentina and Mexico. In these countries the ratio is only 0.1. The highest one is measured in Japan where in the 1990s the ratio of credit to the private sector to GDP was 1.8. In most countries the ratio of credit to the private sector to GDP is about 0.5. This variable is statistically significant at the 5% significance level when country dummies are dropped. With country dummies included the effect is statistically insignificant at the 10% significance level.

4.4 Benchmark values for the equity premium

The equity premium is a crucial parameter in today's financial decision making. This applies to households who have to decide on their investment portfolio, to pension funds determining the financial strategy, and governments who have estimate future tax revenues. This meta-analysis can

[17] Data were kindly made available by Jan Luiten van Zanden and are derived from (i) Mitchell, B.R. (1998), International historical statistics: Africa, Asia and Oceania, 1750-1993, London: Macmillan; (ii) Mitchell, B.R. (1998), International historical statistics: Europe, 1750-1993, London: Macmillan; (iii) Mitchell, B.R. (1998): International historical statistics: The Americas 1750-1993, London: Macmillan. Further information was derived from Dimson et al., Morningstar Encorr, and IMF (2009), International Financial Statistics.

help to narrow down the uncertainty about the equity premium and provide benchmark values that are useful for economists, policymakers and investors. The meta-analysis also allows us to construct confidence intervals for these benchmarks, although these should be treated with caution as we are not certain what is the best specification to use. In the remainder, we use specification (2) in Table 2, thus including a trend term for the post war period.[18] This model includes a time trend for the post war period. Furthermore, we focus on the results for the United States – as this provides the best benchmark with most of the literature – and on the results using the ex ante method, as this method can take account of possible non-stationarity in the data.

As there is no general consensus on the way to define the equity premium, Table 5 provides two benchmarks, and their confidence intervals, depending on whether the equity premium is measured relative to the T-bill rate or the bond rate. These benchmarks refer to the year 2000. The 90% confidence intervals are given between parentheses.

Table 5. Benchmark values for the equity premium in the year 2000

	Mean	Confidence interval
T-bill	4.7	3.6 – 5.9
Bonds	3.8	2.8 – 4.8

The bench-marks are taken for the ex ante method. This is to be preferred because this method is better able to take account of the time variation in the equity premium. Furthermore, we use arithmetic returns as these correspond to the mean of the underlying (asymmetric) distribution of the equity premium. We thus find a bench-mark for the equity premium of 4.7% relative to T-bills, and 3.8% relative to government bonds. Alternatively, using the geometric method the results would have been lower, namely a premium of 3.3% relative to T-bill rates (confidence interval 2.4 - 4.2) and 2.4% relative to bond rates (confidence interval 1.5 - 3.3). This, however, corresponds to the median rather than to the mean of the equity premium.

A few qualifications are in order. First, these bench-marks refer to the United States and cannot automatically be taken to be representative for the world. For European countries and Canada often lower equity premiums are found, while for emerging countries they tend to be higher. In addition, it has to be remembered that focussing on the United States may lead to a survival bias in the results. As mentioned earlier, Jorion and Goetzmann conclude that taking account of this bias will lead to lower world returns on equity by some 0.29% points.

---

[18] If one would neglect this downward trend, and base the benchmarks on the first regression in Table 4.1, the results would have been higher by about 0.9%-points.

A next and obvious limitation is that these benchmarks are constructed for the relatively steady period up to the year 2000. These results should therefore be regarded as a benchmark for the equity premium in a hypothetical steady situation. It is clear that the economy today is far from its normal state. Unfortunately, it is too early to assess the impact of the credit crisis on the equity premium. Using the extended model including the economic fundamentals (Table 4) one could argue that the higher volatility in GDP and lower interest rates would lead to a higher equity premium at present. This is particularly so, if – with hindsight – the volatility experienced in the period up to 2000 was low by historical standards (see also Lettau et al., 2006). On the other hand, the credit crisis may also have deteriorated other fundamentals underlying the equity price, namely expected profits. Therefore, it is impossible at this stage to establish the impact of the credit crisis on the equity premium with any reliability.

And there is a further issue in this regard. Even if the recent fall in equity prices has been triggered by higher volatility in the economy, and is thus associated with a higher prospective equity premium, that does not mean that this can be usefully exploited in terms of an investment strategy (see also Broer et al., 2010). As these high expected returns coincide with high volatility, they do not yield better investment opportunities but rather a shift along the risk-return frontier.

5. Conclusion

This meta-analysis provides an accurate measure of the factors that cause variation in the equity premium. Thereby it explains, to a considerable extent, the heterogeneity of the equity premium in the economic literature. We determine the effects of several factors on the equity premium. The first factor is the applied methodology to measure the equity premium. Variation in the equity premium is the result of calculating equity premiums ex-post or ex-ante, average returns arithmetically or geometrically and using T-bills or bonds as the risk free rate. This variation can easily add up to 3.5% points between the extremes of ex ante/geometric/bond rate on the one hand and ex post/arithmetic/T-bill rate on the other hand. This again indicates how important it is to be clear about the method of measurement.

The second factor is the variation over time. Several authors have pointed to a possible downward trend in the equity premium over time, which can be explained by the development of financial markets allowing for better diversification of risks. The meta-analysis confirms such a pattern. The precise results should be interpreted with care, however. One difficulty in the meta-analysis is that the underlying studies use different periods of observation, both in length and in precise dates. This makes it difficult to accurately pin down an observation of the equity premium to a certain period. At the same time the meta-analysis is of special value here, as it charts the – apparently discretionary – choices made by the different authors in a consistent manner. In the current study, we break down the time dimension into three periods: before 1910, the period after 1950, and the intermediate period characterized by the two World Wars. We also allow for the possibility of a trend in the post-war period.

The third factor concerns the spatial dimension. We find significant differences in equity premiums between the United States, Canada, Secondary Emerging Countries and the Asian Tigers. Emerging countries have a larger equity premium than the United States, whereas Canada has a lower equity premium. For Oceania (including Japan) and Western Europe the differences in comparison with the United States are small and statistically insignificant.

Finally, we have looked into some underlying determinants of the equity premium. The equity premium tends to be higher in periods and countries with larger economic volatility. There is also a clear negative effect of the interest rate, indicating that the return on equity does not vary one-for-one with changes in the interest rate. This also implies that the return on equity cannot be determined by adding a constant equity risk premium to a time varying short or long interest rate. The rate of return on equity has its own dynamics which is only partly associated with the dynamics of the interest rate.

The aim of this meta-analysis was to shed light on the ongoing debate on the height of the equity premium, which tends to be hampered by differences in definition, method of measurement and observation periods. We believe that charting this complex field from a different angle using meta-analysis provides a useful contribution to this literature. The analysis is not meant to replace other (econometric) techniques as being a superior one. Similarly, the value of the equity premium suggested by our analysis as a bench-mark is conditional on the model used in this paper, and should by not be interpreted as a consensus estimate of the equity premium. But exactly because of the uncertainty about the right method and model, meta-analysis is helpful for surveying this literature in a structured manner and enhancing our understanding of sources of variation in estimated equity premiums.

References[19]

Barro, R.J. (2005). Rare Events and the Equity Premium. *National Bureau of Economic Research Working Paper*, no. 11310, Cambridge, MA. ■

Brown, S.J., W.N. Goetzmann and S.A. Ross (1995). Survival. *Journal of Finance*, 50(3), 853–873.

Blanchard, O.J. (1993). Movements in the Equity Premium. *Brookings Papers on Economic Activity*, 1993(2), 75-138. ■

Broer, D.P., T. Knaap and E. Westerhout (2010). Risk Factors in Pension Returns. *Panel Paper*, Netspar, Tilburg (forthcoming).

Campbell, J.Y. and J.H. Cochrane (1999). By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy*, 107(2), 205–251.

Campbell, J.Y. (2008). Viewpoint: Estimating the Equity Premium. *Canadian Journal of Economics*, 41(1), 1–21. ■

[19] Studies included in the database are marked with ■

Campbell, J., A. Lo and A.C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton: Princeton University Press.

Canova, F. and Nicoló de G. (2003). The Properties of the Equity Premium and the Risk-Free Rate: An Investigation Across Time and Countries. *IMF Staff Papers*, 50(2), 222–248. ■

Cate, A. ten (2009). *Arithmetic and Geometric Mean Rates of Return in Discrete Time*. CPB Memorandum 223, CPB Netherlands Bureau for Economic Policy Analysis, The Hague.

Chen, M.H. and P.V. Bidarkota (2004). Consumption Equilibrium Asset Pricing in two Asian Emerging Markets. *Journal of Asian Economics*, 15, 305–319.

Claus, J. and J. Thomas (2001). Equity Premiums as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets. *Journal of Finance*, 56(5), 1629–1666. ■

Cooper, H. and L.V. Hedges (1994). *Handbook of Research Synthesis*. New York: Russell Sage Foundation.

Derrig, R.A. and E.D. Orr (2003). Equity Risk Premium: Expectations Great and Small. *North American Actuarial Journal*, 8(1), 45–69.

Dimson, E., P. Marsh and M. Staunton (2002). *Triumph of the Optimists: 101 Years of Global Investment Returns.* Princeton: Princeton University Press.

Dimson E., P. Marsh and M. Staunton (2006). The Worldwide Equity Premium: A Smaller Puzzle. *London Business School*, mimeo. ■

Erbas, S.N. and A. Mirakhor (2007). The Equity Premium Puzzle, Ambiguity Aversion, and Institutional Quality. *International Monetary Fund Working Paper,* no. 07-230, Washington.

Fama, E. and K. French (1988). Dividend Yields and Expected Returns on Stocks and Bonds. *Journal of Financial Economics*, 25, 23–49. ■

Fama, E. and K. French (2002). The Equity Premium. *Journal of Finance*, 57(2), 637–659. ■

Fernandez, P. (2009a). The Equity Premium in 150 Textbooks. *Working Paper IESE Business School*, no. WP-829, Barcelona.

Fernandez, P. (2009b). Market Risk Premium Used in 2008 by Professors: A Survey with 1,400 Answers. *Working Paper IESE Business School*, no. WP-796, Barcelona.

Florax, R.J.G.M., H.L.F. de Groot and R.A. de Mooij (2002). Meta-analysis: A Tool for Upgrading Inputs of Macroeconomic Policy Models. *Tinbergen Institute Discussion Paper*, no. 041(3), Amsterdam-Rotterdam.

Gordon, M.J. (1962). *The Investment, Financing, and Valuation of the Corporation*. Homewood, Illinois: Richard D. Irwin.

Ibbotson R.G. and P. Chen (2003). Stock Market Returns in the Long Run: Participating in the Real Economy. *Financial Analysts Journal*, 59(1), 88–98. ■

Jacquier, E., A. Kane and A. Marcus (2003). Geometric or Arithmetic Mean: A Reconsideration. *Financial Analysts Journal*, 59, 46–53.

Jagannathan, R., E.R. McGrattan and A. Scherbina (2000). The Declining U.S. Equity Premium. *Federal Reserve Bank of Minneapolis Quarterly Review*, 24(4), 3–19. ■

Jorion, Ph. and W.N. Goetzmann (1999). Global Stock Markets in the Twentieth Century. *Journal of Finance*, 54, 953–980

Kocherlakota, N.R. (1996). The Equity Premium: It's Still a Puzzle. *Journal of Economic Literature*, 34(1), 42–71.

Kyriacou, K., J.B. Madsen and B. Mase (2006). Does Inflation Exaggerate the Equity Premium? *Journal of Economic Studies*, 33(5), 344–356. ■

La Porta, R.F. Lopez-de-Silanes, A. Shleifer and R.W. Vishny (1998). Law and Finance. *Journal of Political Economy*, 106(6), 1113–1155.

Lemke, W. and T. Werner (2009). The Term Structure of Equity Premiums in an Affine Arbitrage-Free Model of Bond Market and Stock Market Dynamics. *ECB Working Paper*, no. 1045, European Central Bank.

Lengwiler, Y. (2004). *Microfoundations of Financial Economics: An Introduction to General Equilibrium Asset Pricing*. Princeton: Princeton University Press.

Lettau, M., S.C. Ludvigson and J.A. Wachter (2008). The Declining Equity Premium: What Role Does Macroeconomic Risk Play. *Review of Financial Studies*, 21(4), 1653–1687.
Levine, R. (1997). Financial Development and Economic Growth: Views and Agenda. *Journal of Economic Literature*, 35(2), 688–726.
Maddison, A. (2007). *Contours of the World Economy, 1-2030 AD*. Oxford: Oxford University Press.
Madsen, J.B. (2004). The Equity Premium Puzzle and the Ex Post Bias. *FRU Working Papers*, no. 01, University of Copenhagen, Department of Economics Finance Research Unit.
Mehra, R. (2003). The Equity Premium: Why Is It a Puzzle? *Financial Analysts Journal*, January / February, 54–69. ■
Mehra, R. (2007). The Equity Premium in India. In: K. Basu (ed.), *The Oxford Companion to Economics in India*. Oxford: Oxford University Press. ■
Mehra, R. (ed.) (2008). *Handbook of the Equity Risk Premium*. Amsterdam: Elsevier.
Mehra, R. and E.C. Prescott (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics*, 15(2), 145–161. ■
Salomons, R. and H. Grootveld (2003). The Equity Risk Premium: Emerging versus Developed Markets. *SOM Theme E: Financial Markets and Institutions*, University of Groningen. ■
Salomons, R. (2008). A Theoretical and Practical Perspective on the Equity Risk Premium. *Journal of Economic Surveys*, 22(2), 299–329.
Santis, M. de (2007). Movements in the Equity Premium: Evidence from a Time-Varying VAR. *Studies in Nonlinear Dynamics and Econometrics*, 11(4), 1–39. ■
Shackman, J.D. (2006). The Equity Premium and Market Integration: Evidence from International Data. *Journal of International Financial Markets, Institutions and Money*, 16(2), 155–179. ■
Siegel, J.J. (1992). The Real Rate of Interest from 1800-1990: A Study of the U.S. and the U.K. *Journal of Monetary Economics*, 29, 227–252. ■
Siegel, J.J. (1999). The Shrinking Equity Premium: Historical Facts and Future Forecasts. *Journal of Portfolio Management*, 26(1), 10–17. ■
Siegel, J.J. (2005). Perspectives on the Equity Risk Premium. *Financial Analyst Journal CFA Institute,* 61–73. ■
Ville, S. (2006). The Equity Premium Puzzle: Australia and the United States in Comparative Perspective. *University of Wollongong Economics Working Paper Series*, no. 06–25, Wollongong, Australia. ■
Vivian, A. (2007). The UK Equity Premium: 1901-2004. *Journal of Business and Accounting*, 34(9/10), 1496–1527. ■
Weil, Ph. (1989). The Equity Premium Puzzle and the Risk-Free Rate Puzzle. *Journal of Monetary Economics,* 24 (3), 401–421.
Welch, I. (2000). Views of Financial Economists of the Equity Premium and on Professional Controversies. *Journal of Business*, 73(4), 501–537. ■
Williams, R.L. (2000). A Note on Robust Variance Estimation for Cluster-Correlated Data. *Biometrics*, 56, 645–646.
Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.
World Bank (2006). *World Development Indicators*. Washington:World Bank.

Appendix A. Summary statistics per study

Study	# obs	Minimum ep	Average ep	Maximum ep	Mid year	Initial year	Final year
Barro (2005)	13	4.70	7.16	10.40	1968.00	1880	2004
Blanchard et al. (1993)	32	–0.20	4.37	8.50	1941.63	1802	1992
Campbell (2002)	15	0.80	5.93	12.35	1978.53	1891	1999
Campbell (2008)	8	1.80	2.95	5.10	1994.25	1982	2006
Canova and De Nicolo (2003)	21	–4.91	3.70	13.84	1985.67	1971	1999
Claus and Thomas (2001)	12	0.21	4.56	7.91	1993.17	1985	1999
De Santis (2007)	14	1.70	4.04	6.40	1966.39	1928	2004
Digby et al. (2006)	23	–0.02	8.14	12.30	1971.20	1910	2004
Dimson et al. (2006)	68	1.80	5.50	10.46	1952.50	1900	2005
Fama and French (2002)	33	–2.15	4.44	14.27	1942.06	1872	2000
Ibbotson and Chen (2002)	4	0.24	3.42	5.24	1963.00	1926	2000
Jagannathan et al. (2000)	38	–0.65	4.84	10.35	1967.13	1930	1999
Kyriacou et al. (2006)	50	2.18	5.95	11.02	1942.00	1871	2002
Mehra (2003)	8	3.30	5.95	8.00	1963.94	1802	2000
Mehra (2007)	12	3.30	6.73	11.30	1968.71	1802	2004
Mehra and Prescott (1985)	9	0.18	6.18	18.30	1933.50	1889	1978
Salomons and Grootveld (2003)	25	–7.86	7.99	45.26	1992.20	1976	2002
Shackman (2006)	39	–20.37	9.50	24.64	1986.00	1970	2002
Siegel (1992)	24	0.79	4.15	7.04	1920.67	1800	1990
Siegel (1999)	16	1.90	5.12	8.60	1917.00	1802	1998
Siegel (2005)	36	–0.21	5.68	12.34	1947.11	1802	2004
Ville (2006)	9	–2.91	4.73	9.53	1933.50	1889	1978
Vivian (2007)	14	–0.09	4.43	7.94	1974.36	1901	2004
Welch (2000)	12	4.30	6.90	9.40	1961.00	1870	1998
Grand Total	535	–20.37	5.73	45.26	1958.56	1800	2006

Appendix B. Accounting for dependence

Dependence among observations in meta-analysis studies may occur between estimates from the same study, country, region or time period and results in standard errors that are wrong. In the main text, we have accounted for within-study dependence by reporting Huber-White cluster robust standard errors. This Appendix shows results with standard errors that have been corrected for dependence across regions (Western Europe, Developing countries, Canada, Australia, South Africa, Japan and the United States) and time periods (pre-1910, 1910–1950 and post 1950). We take the specification (2) in Table 2 as the base specification. Comparable results for other specifications are available upon request.

Table B.1. Accounting for different types of dependence

	Base	Spatial	Temporal
Constant	4.10***	4.10***	4.10**
	(0.59)	(0.45)	(0.55)
Arithmetic mean	1.42***	1.42***	1.42***
	(0.33)	(0.22)	(0.13)
Ex Post	1.05***	1.05***	1.05
	(0.30)	(0.25)	(0.75)
T-bill used	0.92***	0.92***	0.92*
	(0.26)	(0.21)	(0.24)
Region effects (relative to USA)			
Canada	–1.65***	–1.65***	–1.65***
	(0.48)	(0.11)	(0.08)
Oceania	–0.64	–0.64***	–0.64
	(0.63)	(0.08)	(0.38)
Western Europe	–0.22	–0.22*	–0.22
	(0.64)	(0.10)	(0.11)
Advanced emerging	1.31	1.31***	1.31***
	(0.86)	(0.27)	(0.10)
Secondary emerging	5.95***	5.95***	5.95***
	(0.74)	(0.77)	(0.23)
Asian Tigers	7.11***	7.11***	7.11***
	(2.01)	(0.66)	(0.28)
Period effects			
Before 1910	–3.46***	–3.46***	–3.46***
	(0.57)	(0.36)	(0.19)
After 1950	0.16	0.16	0.16
	(0.62)	(0.70)	(0.16)
Trend after 1950	–0.04**	–0.04	–0.04**
	(0.02)	(0.04)	(0.004)
# observations	535	535	535
R^2	0.22	0.22	0.22

Note: Statistical significance of the estimated coefficients is indicated by ***, ** and * referring, respectively, to the 1%, 5% and 10% significance level.

Journal of Economic Perspectives—Volume 18, Number 3—Summer 2004—Pages 25–46

The Capital Asset Pricing Model: Theory and Evidence

Eugene F. Fama and Kenneth R. French

The capital asset pricing model (CAPM) of William Sharpe (1964) and John Lintner (1965) marks the birth of asset pricing theory (resulting in a Nobel Prize for Sharpe in 1990). Four decades later, the CAPM is still widely used in applications, such as estimating the cost of capital for firms and evaluating the performance of managed portfolios. It is the centerpiece of MBA investment courses. Indeed, it is often the only asset pricing model taught in these courses.[1]

The attraction of the CAPM is that it offers powerful and intuitively pleasing predictions about how to measure risk and the relation between expected return and risk. Unfortunately, the empirical record of the model is poor—poor enough to invalidate the way it is used in applications. The CAPM's empirical problems may reflect theoretical failings, the result of many simplifying assumptions. But they may also be caused by difficulties in implementing valid tests of the model. For example, the CAPM says that the risk of a stock should be measured relative to a comprehensive "market portfolio" that in principle can include not just traded financial assets, but also consumer durables, real estate and human capital. Even if we take a narrow view of the model and limit its purview to traded financial assets, is it

[1] Although every asset pricing model is a capital asset pricing model, the finance profession reserves the acronym CAPM for the specific model of Sharpe (1964), Lintner (1965) and Black (1972) discussed here. Thus, throughout the paper we refer to the Sharpe-Lintner-Black model as the CAPM.

■ *Eugene F. Fama is Robert R. McCormick Distinguished Service Professor of Finance, Graduate School of Business, University of Chicago, Chicago, Illinois. Kenneth R. French is Carl E. and Catherine M. Heidt Professor of Finance, Tuck School of Business, Dartmouth College, Hanover, New Hampshire. Their e-mail addresses are ⟨eugene.fama@gsb.uchicago.edu⟩ and ⟨kfrench@dartmouth.edu⟩, respectively.*

legitimate to limit further the market portfolio to U.S. common stocks (a typical choice), or should the market be expanded to include bonds, and other financial assets, perhaps around the world? In the end, we argue that whether the model's problems reflect weaknesses in the theory or in its empirical implementation, the failure of the CAPM in empirical tests implies that most applications of the model are invalid.

We begin by outlining the logic of the CAPM, focusing on its predictions about risk and expected return. We then review the history of empirical work and what it says about shortcomings of the CAPM that pose challenges to be explained by alternative models.

The Logic of the CAPM

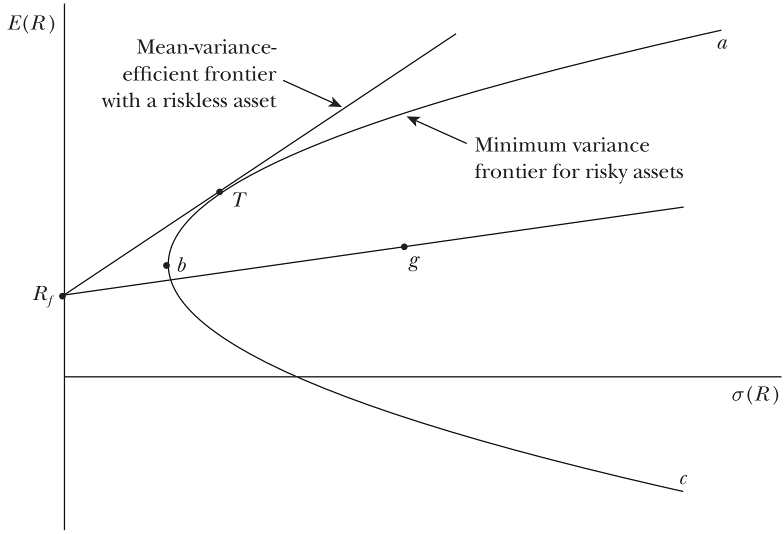
The CAPM builds on the model of portfolio choice developed by Harry Markowitz (1959). In Markowitz's model, an investor selects a portfolio at time $t - 1$ that produces a stochastic return at t. The model assumes investors are risk averse and, when choosing among portfolios, they care only about the mean and variance of their one-period investment return. As a result, investors choose "mean-variance-efficient" portfolios, in the sense that the portfolios 1) minimize the variance of portfolio return, given expected return, and 2) maximize expected return, given variance. Thus, the Markowitz approach is often called a "mean-variance model."

The portfolio model provides an algebraic condition on asset weights in mean-variance-efficient portfolios. The CAPM turns this algebraic statement into a testable prediction about the relation between risk and expected return by identifying a portfolio that must be efficient if asset prices are to clear the market of all assets.

Sharpe (1964) and Lintner (1965) add two key assumptions to the Markowitz model to identify a portfolio that must be mean-variance-efficient. The first assumption is *complete agreement*: given market clearing asset prices at $t - 1$, investors agree on the joint distribution of asset returns from $t - 1$ to t. And this distribution is the true one—that is, it is the distribution from which the returns we use to test the model are drawn. The second assumption is that there is *borrowing and lending at a risk-free rate*, which is the same for all investors and does not depend on the amount borrowed or lent.

Figure 1 describes portfolio opportunities and tells the CAPM story. The horizontal axis shows portfolio risk, measured by the standard deviation of portfolio return; the vertical axis shows expected return. The curve *abc*, which is called the minimum variance frontier, traces combinations of expected return and risk for portfolios of risky assets that minimize return variance at different levels of expected return. (These portfolios do not include risk-free borrowing and lending.) The tradeoff between risk and expected return for minimum variance portfolios is apparent. For example, an investor who wants a high expected return, perhaps at point *a*, must accept high volatility. At point *T*, the investor can have an interme-

Figure 1
Investment Opportunities



diate expected return with lower volatility. If there is no risk-free borrowing or lending, only portfolios above *b* along *abc* are mean-variance-efficient, since these portfolios also maximize expected return, given their return variances.

Adding risk-free borrowing and lending turns the efficient set into a straight line. Consider a portfolio that invests the proportion x of portfolio funds in a risk-free security and $1 - x$ in some portfolio g. If all funds are invested in the risk-free security—that is, they are loaned at the risk-free rate of interest—the result is the point R_f in Figure 1, a portfolio with zero variance and a risk-free rate of return. Combinations of risk-free lending and positive investment in g plot on the straight line between R_f and g. Points to the right of g on the line represent borrowing at the risk-free rate, with the proceeds from the borrowing used to increase investment in portfolio g. In short, portfolios that combine risk-free lending or borrowing with some risky portfolio g plot along a straight line from R_f through g in Figure 1.[2]

[2] Formally, the return, expected return and standard deviation of return on portfolios of the risk-free asset f and a risky portfolio g vary with x, the proportion of portfolio funds invested in f, as

$$R_p = xR_f + (1 - x)R_g,$$

$$E(R_p) = xR_f + (1 - x)E(R_g),$$

$$\sigma(R_p) = (1 - x)\sigma(R_g), \; x \leq 1.0,$$

which together imply that the portfolios plot along the line from R_f through g in Figure 1.

To obtain the mean-variance-efficient portfolios available with risk-free borrowing and lending, one swings a line from R_f in Figure 1 up and to the left as far as possible, to the tangency portfolio T. We can then see that all efficient portfolios are combinations of the risk-free asset (either risk-free borrowing or lending) and a single risky tangency portfolio, T. This key result is Tobin's (1958) "separation theorem."

The punch line of the CAPM is now straightforward. With complete agreement about distributions of returns, all investors see the same opportunity set (Figure 1), and they combine the same risky tangency portfolio T with risk-free lending or borrowing. Since all investors hold the same portfolio T of risky assets, it must be the value-weight market portfolio of risky assets. Specifically, each risky asset's weight in the tangency portfolio, which we now call M (for the "market"), must be the total market value of all outstanding units of the asset divided by the total market value of all risky assets. In addition, the risk-free rate must be set (along with the prices of risky assets) to clear the market for risk-free borrowing and lending.

In short, the CAPM assumptions imply that the market portfolio M must be on the minimum variance frontier if the asset market is to clear. This means that the algebraic relation that holds for any minimum variance portfolio must hold for the market portfolio. Specifically, if there are N risky assets,

$$\text{(Minimum Variance Condition for } M) \quad E(R_i) = E(R_{ZM}) + [E(R_M) - E(R_{ZM})]\beta_{iM}, \quad i = 1, \ldots, N.$$

In this equation, $E(R_i)$ is the expected return on asset i, and β_{iM}, the market beta of asset i, is the covariance of its return with the market return divided by the variance of the market return,

$$\text{(Market Beta)} \quad \beta_{iM} = \frac{\text{cov}(R_i, R_M)}{\sigma^2(R_M)}.$$

The first term on the right-hand side of the minimum variance condition, $E(R_{ZM})$, is the expected return on assets that have market betas equal to zero, which means their returns are uncorrelated with the market return. The second term is a risk premium—the market beta of asset i, β_{iM}, times the premium per unit of beta, which is the expected market return, $E(R_M)$, minus $E(R_{ZM})$.

Since the market beta of asset i is also the slope in the regression of its return on the market return, a common (and correct) interpretation of beta is that it measures the sensitivity of the asset's return to variation in the market return. But there is another interpretation of beta more in line with the spirit of the portfolio model that underlies the CAPM. The risk of the market portfolio, as measured by the variance of its return (the denominator of β_{iM}), is a weighted average of the covariance risks of the assets in M (the numerators of β_{iM} for different assets).

Thus, β_{iM} is the covariance risk of asset i in M measured relative to the average covariance risk of assets, which is just the variance of the market return.[3] In economic terms, β_{iM} is proportional to the risk each dollar invested in asset i contributes to the market portfolio.

The last step in the development of the Sharpe-Lintner model is to use the assumption of risk-free borrowing and lending to nail down $E(R_{ZM})$, the expected return on zero-beta assets. A risky asset's return is uncorrelated with the market return—its beta is zero—when the average of the asset's covariances with the returns on other assets just offsets the variance of the asset's return. Such a risky asset is riskless in the market portfolio in the sense that it contributes nothing to the variance of the market return.

When there is risk-free borrowing and lending, the expected return on assets that are uncorrelated with the market return, $E(R_{ZM})$, must equal the risk-free rate, R_f. The relation between expected return and beta then becomes the familiar Sharpe-Lintner CAPM equation,

$$\text{(Sharpe-Lintner CAPM)} \quad E(R_i) = R_f + [E(R_M) - R_f)]\beta_{iM}, \; i = 1, \ldots, N.$$

In words, the expected return on any asset i is the risk-free interest rate, R_f, plus a risk premium, which is the asset's market beta, β_{iM}, times the premium per unit of beta risk, $E(R_M) - R_f$.

Unrestricted risk-free borrowing and lending is an unrealistic assumption. Fischer Black (1972) develops a version of the CAPM without risk-free borrowing or lending. He shows that the CAPM's key result—that the market portfolio is mean-variance-efficient—can be obtained by instead allowing unrestricted short sales of risky assets. In brief, back in Figure 1, if there is no risk-free asset, investors select portfolios from along the mean-variance-efficient frontier from a to b. Market clearing prices imply that when one weights the efficient portfolios chosen by investors by their (positive) shares of aggregate invested wealth, the resulting portfolio is the market portfolio. The market portfolio is thus a portfolio of the efficient portfolios chosen by investors. With unrestricted short selling of risky assets, portfolios made up of efficient portfolios are themselves efficient. Thus, the market portfolio is efficient, which means that the minimum variance condition for M given above holds, and it is the expected return-risk relation of the Black CAPM.

The relations between expected return and market beta of the Black and Sharpe-Lintner versions of the CAPM differ only in terms of what each says about $E(R_{ZM})$, the expected return on assets uncorrelated with the market. The Black version says only that $E(R_{ZM})$ must be less than the expected market return, so the

[3] Formally, if x_{iM} is the weight of asset i in the market portfolio, then the variance of the portfolio's return is

$$\sigma^2(R_M) = Cov(R_M, R_M) = Cov\left(\sum_{i=1}^{N} x_{iM}R_i, \; R_M\right) = \sum_{i=1}^{N} x_{iM}Cov(R_i, R_M).$$

premium for beta is positive. In contrast, in the Sharpe-Lintner version of the model, $E(R_{ZM})$ must be the risk-free interest rate, R_f, and the premium per unit of beta risk is $E(R_M) - R_f$.

The assumption that short selling is unrestricted is as unrealistic as unrestricted risk-free borrowing and lending. If there is no risk-free asset and short sales of risky assets are not allowed, mean-variance investors still choose efficient portfolios—points above *b* on the *abc* curve in Figure 1. But when there is no short selling of risky assets and no risk-free asset, the algebra of portfolio efficiency says that portfolios made up of efficient portfolios are not typically efficient. This means that the market portfolio, which is a portfolio of the efficient portfolios chosen by investors, is not typically efficient. And the CAPM relation between expected return and market beta is lost. This does not rule out predictions about expected return and betas with respect to other efficient portfolios—if theory can specify portfolios that must be efficient if the market is to clear. But so far this has proven impossible.

In short, the familiar CAPM equation relating expected asset returns to their market betas is just an application to the market portfolio of the relation between expected return and portfolio beta that holds in any mean-variance-efficient portfolio. The efficiency of the market portfolio is based on many unrealistic assumptions, including complete agreement and either unrestricted risk-free borrowing and lending or unrestricted short selling of risky assets. But all interesting models involve unrealistic simplifications, which is why they must be tested against data.

Early Empirical Tests

Tests of the CAPM are based on three implications of the relation between expected return and market beta implied by the model. First, expected returns on all assets are linearly related to their betas, and no other variable has marginal explanatory power. Second, the beta premium is positive, meaning that the expected return on the market portfolio exceeds the expected return on assets whose returns are uncorrelated with the market return. Third, in the Sharpe-Lintner version of the model, assets uncorrelated with the market have expected returns equal to the risk-free interest rate, and the beta premium is the expected market return minus the risk-free rate. Most tests of these predictions use either cross-section or time-series regressions. Both approaches date to early tests of the model.

Tests on Risk Premiums

The early cross-section regression tests focus on the Sharpe-Lintner model's predictions about the intercept and slope in the relation between expected return and market beta. The approach is to regress a cross-section of average asset returns on estimates of asset betas. The model predicts that the intercept in these regressions is the risk-free interest rate, R_f, and the coefficient on beta is the expected return on the market in excess of the risk-free rate, $E(R_M) - R_f$.

Two problems in these tests quickly became apparent. First, estimates of beta for individual assets are imprecise, creating a measurement error problem when they are used to explain average returns. Second, the regression residuals have common sources of variation, such as industry effects in average returns. Positive correlation in the residuals produces downward bias in the usual ordinary least squares estimates of the standard errors of the cross-section regression slopes.

To improve the precision of estimated betas, researchers such as Blume (1970), Friend and Blume (1970) and Black, Jensen and Scholes (1972) work with portfolios, rather than individual securities. Since expected returns and market betas combine in the same way in portfolios, if the CAPM explains security returns it also explains portfolio returns.[4] Estimates of beta for diversified portfolios are more precise than estimates for individual securities. Thus, using portfolios in cross-section regressions of average returns on betas reduces the critical errors in variables problem. Grouping, however, shrinks the range of betas and reduces statistical power. To mitigate this problem, researchers sort securities on beta when forming portfolios; the first portfolio contains securities with the lowest betas, and so on, up to the last portfolio with the highest beta assets. This sorting procedure is now standard in empirical tests.

Fama and MacBeth (1973) propose a method for addressing the inference problem caused by correlation of the residuals in cross-section regressions. Instead of estimating a single cross-section regression of average monthly returns on betas, they estimate month-by-month cross-section regressions of monthly returns on betas. The times-series means of the monthly slopes and intercepts, along with the standard errors of the means, are then used to test whether the average premium for beta is positive and whether the average return on assets uncorrelated with the market is equal to the average risk-free interest rate. In this approach, the standard errors of the average intercept and slope are determined by the month-to-month variation in the regression coefficients, which fully captures the effects of residual correlation on variation in the regression coefficients, but sidesteps the problem of actually estimating the correlations. The residual correlations are, in effect, captured via repeated sampling of the regression coefficients. This approach also becomes standard in the literature.

Jensen (1968) was the first to note that the Sharpe-Lintner version of the

[4] Formally, if x_{ip}, $i = 1, \ldots, N$, are the weights for assets in some portfolio p, the expected return and market beta for the portfolio are related to the expected returns and betas of assets as

$$E(R_p) = \sum_{i=1}^{N} x_{ip}E(R_i), \text{ and } \beta_{pM} = \sum_{i=1}^{N} x_{ip}\beta_{pM}.$$

Thus, the CAPM relation between expected return and beta,

$$E(R_i) = E(R_f) + [E(R_M) - E(R_f)]\beta_{iM},$$

holds when asset i is a portfolio, as well as when i is an individual security.

relation between expected return and market beta also implies a time-series regression test. The Sharpe-Lintner CAPM says that the expected value of an asset's excess return (the asset's return minus the risk-free interest rate, $R_{it} - R_{ft}$) is completely explained by its expected CAPM risk premium (its beta times the expected value of $R_{Mt} - R_{ft}$). This implies that "Jensen's alpha," the intercept term in the time-series regression,

$$\text{(Time-Series Regression)} \quad R_{it} - R_{ft} = \alpha_i + \beta_{iM}(R_{Mt} - R_{ft}) + \varepsilon_{it},$$

is zero for each asset.

The early tests firmly reject the Sharpe-Lintner version of the CAPM. There is a positive relation between beta and average return, but it is too "flat." Recall that, in cross-section regressions, the Sharpe-Lintner model predicts that the intercept is the risk-free rate and the coefficient on beta is the expected market return in excess of the risk-free rate, $E(R_M) - R_f$. The regressions consistently find that the intercept is greater than the average risk-free rate (typically proxied as the return on a one-month Treasury bill), and the coefficient on beta is less than the average excess market return (proxied as the average return on a portfolio of U.S. common stocks minus the Treasury bill rate). This is true in the early tests, such as Douglas (1968), Black, Jensen and Scholes (1972), Miller and Scholes (1972), Blume and Friend (1973) and Fama and MacBeth (1973), as well as in more recent cross-section regression tests, like Fama and French (1992).

The evidence that the relation between beta and average return is too flat is confirmed in time-series tests, such as Friend and Blume (1970), Black, Jensen and Scholes (1972) and Stambaugh (1982). The intercepts in time-series regressions of excess asset returns on the excess market return are positive for assets with low betas and negative for assets with high betas.

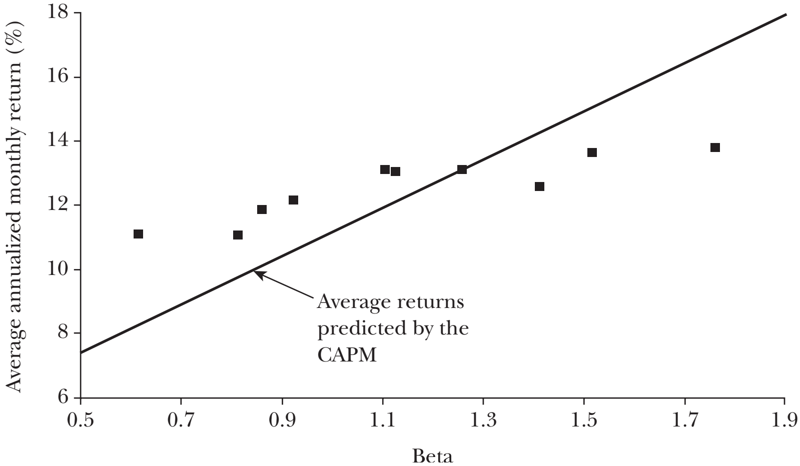
Figure 2 provides an updated example of the evidence. In December of each year, we estimate a preranking beta for every NYSE (1928–2003), AMEX (1963–2003) and NASDAQ (1972–2003) stock in the CRSP (Center for Research in Security Prices of the University of Chicago) database, using two to five years (as available) of prior monthly returns.[5] We then form ten value-weight portfolios based on these preranking betas and compute their returns for the next twelve months. We repeat this process for each year from 1928 to 2003. The result is 912 monthly returns on ten beta-sorted portfolios. Figure 2 plots each portfolio's average return against its postranking beta, estimated by regressing its monthly returns for 1928–2003 on the return on the CRSP value-weight portfolio of U.S. common stocks.

The Sharpe-Lintner CAPM predicts that the portfolios plot along a straight

[5] To be included in the sample for year t, a security must have market equity data (price times shares outstanding) for December of $t - 1$, and CRSP must classify it as ordinary common equity. Thus, we exclude securities such as American Depository Receipts (ADRs) and Real Estate Investment Trusts (REITs).

Figure 2

Average Annualized Monthly Return versus Beta for Value Weight Portfolios Formed on Prior Beta, 1928–2003



line, with an intercept equal to the risk-free rate, R_f, and a slope equal to the expected excess return on the market, $E(R_M) - R_f$. We use the average one-month Treasury bill rate and the average excess CRSP market return for 1928–2003 to estimate the predicted line in Figure 2. Confirming earlier evidence, the relation between beta and average return for the ten portfolios is much flatter than the Sharpe-Lintner CAPM predicts. The returns on the low beta portfolios are too high, and the returns on the high beta portfolios are too low. For example, the predicted return on the portfolio with the lowest beta is 8.3 percent per year; the actual return is 11.1 percent. The predicted return on the portfolio with the highest beta is 16.8 percent per year; the actual is 13.7 percent.

Although the observed premium per unit of beta is lower than the Sharpe-Lintner model predicts, the relation between average return and beta in Figure 2 is roughly linear. This is consistent with the Black version of the CAPM, which predicts only that the beta premium is positive. Even this less restrictive model, however, eventually succumbs to the data.

Testing Whether Market Betas Explain Expected Returns

The Sharpe-Lintner and Black versions of the CAPM share the prediction that the market portfolio is mean-variance-efficient. This implies that differences in expected return across securities and portfolios are entirely explained by differences in market beta; other variables should add nothing to the explanation of expected return. This prediction plays a prominent role in tests of the CAPM. In the early work, the weapon of choice is cross-section regressions.

In the framework of Fama and MacBeth (1973), one simply adds predetermined explanatory variables to the month-by-month cross-section regressions of returns on beta. If all differences in expected return are explained by beta, the average slopes on the additional variables should not be reliably different from zero. Clearly, the trick in the cross-section regression approach is to choose specific additional variables likely to expose any problems of the CAPM prediction that, because the market portfolio is efficient, market betas suffice to explain expected asset returns.

For example, in Fama and MacBeth (1973) the additional variables are squared market betas (to test the prediction that the relation between expected return and beta is linear) and residual variances from regressions of returns on the market return (to test the prediction that market beta is the only measure of risk needed to explain expected returns). These variables do not add to the explanation of average returns provided by beta. Thus, the results of Fama and MacBeth (1973) are consistent with the hypothesis that their market proxy—an equal-weight portfolio of NYSE stocks—is on the minimum variance frontier.

The hypothesis that market betas completely explain expected returns can also be tested using time-series regressions. In the time-series regression described above (the excess return on asset i regressed on the excess market return), the intercept is the difference between the asset's average excess return and the excess return predicted by the Sharpe-Lintner model, that is, beta times the average excess market return. If the model holds, there is no way to group assets into portfolios whose intercepts are reliably different from zero. For example, the intercepts for a portfolio of stocks with high ratios of earnings to price and a portfolio of stocks with low earning-price ratios should both be zero. Thus, to test the hypothesis that market betas suffice to explain expected returns, one estimates the time-series regression for a set of assets (or portfolios) and then jointly tests the vector of regression intercepts against zero. The trick in this approach is to choose the left-hand-side assets (or portfolios) in a way likely to expose any shortcoming of the CAPM prediction that market betas suffice to explain expected asset returns.

In early applications, researchers use a variety of tests to determine whether the intercepts in a set of time-series regressions are all zero. The tests have the same asymptotic properties, but there is controversy about which has the best small sample properties. Gibbons, Ross and Shanken (1989) settle the debate by providing an F-test on the intercepts that has exact small-sample properties. They also show that the test has a simple economic interpretation. In effect, the test constructs a candidate for the tangency portfolio T in Figure 1 by optimally combining the market proxy and the left-hand-side assets of the time-series regressions. The estimator then tests whether the efficient set provided by the combination of this tangency portfolio and the risk-free asset is reliably superior to the one obtained by combining the risk-free asset with the market proxy alone. In other words, the Gibbons, Ross and Shanken statistic tests whether the market proxy is the tangency portfolio in the set of portfolios that can be constructed by combining the market portfolio with the specific assets used as dependent variables in the time-series regressions.

Enlightened by this insight of Gibbons, Ross and Shanken (1989), one can see a similar interpretation of the cross-section regression test of whether market betas suffice to explain expected returns. In this case, the test is whether the additional explanatory variables in a cross-section regression identify patterns in the returns on the left-hand-side assets that are not explained by the assets' market betas. This amounts to testing whether the market proxy is on the minimum variance frontier that can be constructed using the market proxy and the left-hand-side assets included in the tests.

An important lesson from this discussion is that time-series and cross-section regressions do not, strictly speaking, test the CAPM. What is literally tested is whether a specific proxy for the market portfolio (typically a portfolio of U.S. common stocks) is efficient in the set of portfolios that can be constructed from it and the left-hand-side assets used in the test. One might conclude from this that the CAPM has never been tested, and prospects for testing it are not good because 1) the set of left-hand-side assets does not include all marketable assets, and 2) data for the true market portfolio of all assets are likely beyond reach (Roll, 1977; more on this later). But this criticism can be leveled at tests of any economic model when the tests are less than exhaustive or when they use proxies for the variables called for by the model.

The bottom line from the early cross-section regression tests of the CAPM, such as Fama and MacBeth (1973), and the early time-series regression tests, like Gibbons (1982) and Stambaugh (1982), is that standard market proxies seem to be on the minimum variance frontier. That is, the central predictions of the Black version of the CAPM, that market betas suffice to explain expected returns and that the risk premium for beta is positive, seem to hold. But the more specific prediction of the Sharpe-Lintner CAPM that the premium per unit of beta is the expected market return minus the risk-free interest rate is consistently rejected.

The success of the Black version of the CAPM in early tests produced a consensus that the model is a good description of expected returns. These early results, coupled with the model's simplicity and intuitive appeal, pushed the CAPM to the forefront of finance.

Recent Tests

Starting in the late 1970s, empirical work appears that challenges even the Black version of the CAPM. Specifically, evidence mounts that much of the variation in expected return is unrelated to market beta.

The first blow is Basu's (1977) evidence that when common stocks are sorted on earnings-price ratios, future returns on high E/P stocks are higher than predicted by the CAPM. Banz (1981) documents a size effect: when stocks are sorted on market capitalization (price times shares outstanding), average returns on small stocks are higher than predicted by the CAPM. Bhandari (1988) finds that high debt-equity ratios (book value of debt over the market value of equity, a measure of leverage) are associated with returns that are too high relative to their market betas.

Finally, Statman (1980) and Rosenberg, Reid and Lanstein (1985) document that stocks with high book-to-market equity ratios (B/M, the ratio of the book value of a common stock to its market value) have high average returns that are not captured by their betas.

There is a theme in the contradictions of the CAPM summarized above. Ratios involving stock prices have information about expected returns missed by market betas. On reflection, this is not surprising. A stock's price depends not only on the expected cash flows it will provide, but also on the expected returns that discount expected cash flows back to the present. Thus, in principle, the cross-section of prices has information about the cross-section of expected returns. (A high expected return implies a high discount rate and a low price.) The cross-section of stock prices is, however, arbitrarily affected by differences in scale (or units). But with a judicious choice of scaling variable X, the ratio X/P can reveal differences in the cross-section of expected stock returns. Such ratios are thus prime candidates to expose shortcomings of asset pricing models—in the case of the CAPM, shortcomings of the prediction that market betas suffice to explain expected returns (Ball, 1978). The contradictions of the CAPM summarized above suggest that earnings-price, debt-equity and book-to-market ratios indeed play this role.

Fama and French (1992) update and synthesize the evidence on the empirical failures of the CAPM. Using the cross-section regression approach, they confirm that size, earnings-price, debt-equity and book-to-market ratios add to the explanation of expected stock returns provided by market beta. Fama and French (1996) reach the same conclusion using the time-series regression approach applied to portfolios of stocks sorted on price ratios. They also find that different price ratios have much the same information about expected returns. This is not surprising given that price is the common driving force in the price ratios, and the numerators are just scaling variables used to extract the information in price about expected returns.

Fama and French (1992) also confirm the evidence (Reinganum, 1981; Stambaugh, 1982; Lakonishok and Shapiro, 1986) that the relation between average return and beta for common stocks is even flatter after the sample periods used in the early empirical work on the CAPM. The estimate of the beta premium is, however, clouded by statistical uncertainty (a large standard error). Kothari, Shanken and Sloan (1995) try to resuscitate the Sharpe-Lintner CAPM by arguing that the weak relation between average return and beta is just a chance result. But the strong evidence that other variables capture variation in expected return missed by beta makes this argument irrelevant. If betas do not suffice to explain expected returns, the market portfolio is not efficient, and the CAPM is dead in its tracks. Evidence on the size of the market premium can neither save the model nor further doom it.

The synthesis of the evidence on the empirical problems of the CAPM provided by Fama and French (1992) serves as a catalyst, marking the point when it is generally acknowledged that the CAPM has potentially fatal problems. Research then turns to explanations.

One possibility is that the CAPM's problems are spurious, the result of data dredging—publication-hungry researchers scouring the data and unearthing contradictions that occur in specific samples as a result of chance. A standard response to this concern is to test for similar findings in other samples. Chan, Hamao and Lakonishok (1991) find a strong relation between book-to-market equity (B/M) and average return for Japanese stocks. Capaul, Rowley and Sharpe (1993) observe a similar B/M effect in four European stock markets and in Japan. Fama and French (1998) find that the price ratios that produce problems for the CAPM in U.S. data show up in the same way in the stock returns of twelve non-U.S. major markets, and they are present in emerging market returns. This evidence suggests that the contradictions of the CAPM associated with price ratios are not sample specific.

Explanations: Irrational Pricing or Risk

Among those who conclude that the empirical failures of the CAPM are fatal, two stories emerge. On one side are the behavioralists. Their view is based on evidence that stocks with high ratios of book value to market price are typically firms that have fallen on bad times, while low B/M is associated with growth firms (Lakonishok, Shleifer and Vishny, 1994; Fama and French, 1995). The behavioralists argue that sorting firms on book-to-market ratios exposes investor overreaction to good and bad times. Investors overextrapolate past performance, resulting in stock prices that are too high for growth (low B/M) firms and too low for distressed (high B/M, so-called value) firms. When the overreaction is eventually corrected, the result is high returns for value stocks and low returns for growth stocks. Proponents of this view include DeBondt and Thaler (1987), Lakonishok, Shleifer and Vishny (1994) and Haugen (1995).

The second story for explaining the empirical contradictions of the CAPM is that they point to the need for a more complicated asset pricing model. The CAPM is based on many unrealistic assumptions. For example, the assumption that investors care only about the mean and variance of one-period portfolio returns is extreme. It is reasonable that investors also care about how their portfolio return covaries with labor income and future investment opportunities, so a portfolio's return variance misses important dimensions of risk. If so, market beta is not a complete description of an asset's risk, and we should not be surprised to find that differences in expected return are not completely explained by differences in beta. In this view, the search should turn to asset pricing models that do a better job explaining average returns.

Merton's (1973) intertemporal capital asset pricing model (ICAPM) is a natural extension of the CAPM. The ICAPM begins with a different assumption about investor objectives. In the CAPM, investors care only about the wealth their portfolio produces at the end of the current period. In the ICAPM, investors are concerned not only with their end-of-period payoff, but also with the opportunities they will have to consume or invest the payoff. Thus, when choosing a portfolio at time $t - 1$, ICAPM investors consider how their wealth at t might vary with future *state variables*, including labor income, the prices of consumption goods and the nature of portfolio opportunities at t, and expectations about the labor income, consumption and investment opportunities to be available after t.

Like CAPM investors, ICAPM investors prefer high expected return and low return variance. But ICAPM investors are also concerned with the covariances of portfolio returns with state variables. As a result, optimal portfolios are "multifactor efficient," which means they have the largest possible expected returns, given their return variances and the covariances of their returns with the relevant state variables.

Fama (1996) shows that the ICAPM generalizes the logic of the CAPM. That is, if there is risk-free borrowing and lending or if short sales of risky assets are allowed, market clearing prices imply that the market portfolio is multifactor efficient. Moreover, multifactor efficiency implies a relation between expected return and beta risks, but it requires additional betas, along with a market beta, to explain expected returns.

An ideal implementation of the ICAPM would specify the state variables that affect expected returns. Fama and French (1993) take a more indirect approach, perhaps more in the spirit of Ross's (1976) arbitrage pricing theory. They argue that though size and book-to-market equity are not themselves state variables, the higher average returns on small stocks and high book-to-market stocks reflect unidentified state variables that produce undiversifiable risks (covariances) in returns that are not captured by the market return and are priced separately from market betas. In support of this claim, they show that the returns on the stocks of small firms covary more with one another than with returns on the stocks of large firms, and returns on high book-to-market (value) stocks covary more with one another than with returns on low book-to-market (growth) stocks. Fama and French (1995) show that there are similar size and book-to-market patterns in the covariation of fundamentals like earnings and sales.

Based on this evidence, Fama and French (1993, 1996) propose a three-factor model for expected returns,

$$\text{(Three-Factor Model)} \quad E(R_{it}) - R_{ft} = \beta_{iM}[E(R_{Mt}) - R_{ft}] + \beta_{is}E(SMB_t) + \beta_{ih}E(HML_t).$$

In this equation, SMB_t (small minus big) is the difference between the returns on diversified portfolios of small and big stocks, HML_t (high minus low) is the difference between the returns on diversified portfolios of high and low B/M stocks, and the betas are slopes in the multiple regression of $R_{it} - R_{ft}$ on $R_{Mt} - R_{ft}$, SMB_t and HML_t.

For perspective, the average value of the market premium $R_{Mt} - R_{ft}$ for 1927–2003 is 8.3 percent per year, which is 3.5 standard errors from zero. The average values of SMB_t, and HML_t are 3.6 percent and 5.0 percent per year, and they are 2.1 and 3.1 standard errors from zero. All three premiums are volatile, with annual standard deviations of 21.0 percent ($R_{Mt} - R_{ft}$), 14.6 percent (SMB_t) and 14.2 percent (HML_t) per year. Although the average values of the premiums are large, high volatility implies substantial uncertainty about the true expected premiums.

One implication of the expected return equation of the three-factor model is that the intercept α_i in the time-series regression,

$$R_{it} - R_{ft} = \alpha_i + \beta_{iM}(R_{Mt} - R_{ft}) + \beta_{is}SMB_t + \beta_{ih}HML_t + \varepsilon_{it},$$

is zero for all assets i. Using this criterion, Fama and French (1993, 1996) find that the model captures much of the variation in average return for portfolios formed on size, book-to-market equity and other price ratios that cause problems for the CAPM. Fama and French (1998) show that an international version of the model performs better than an international CAPM in describing average returns on portfolios formed on scaled price variables for stocks in 13 major markets.

The three-factor model is now widely used in empirical research that requires a model of expected returns. Estimates of α_i from the time-series regression above are used to calibrate how rapidly stock prices respond to new information (for example, Loughran and Ritter, 1995; Mitchell and Stafford, 2000). They are also used to measure the special information of portfolio managers, for example, in Carhart's (1997) study of mutual fund performance. Among practitioners like Ibbotson Associates, the model is offered as an alternative to the CAPM for estimating the cost of equity capital.

From a theoretical perspective, the main shortcoming of the three-factor model is its empirical motivation. The small-minus-big (SMB) and high-minus-low (HML) explanatory returns are not motivated by predictions about state variables of concern to investors. Instead they are brute force constructs meant to capture the patterns uncovered by previous work on how average stock returns vary with size and the book-to-market equity ratio.

But this concern is not fatal. The ICAPM does not require that the additional portfolios used along with the market portfolio to explain expected returns "mimic" the relevant state variables. In both the ICAPM and the arbitrage pricing theory, it suffices that the additional portfolios are well diversified (in the terminology of Fama, 1996, they are multifactor minimum variance) and that they are sufficiently different from the market portfolio to capture covariation in returns and variation in expected returns missed by the market portfolio. Thus, adding diversified portfolios that capture covariation in returns and variation in average returns left unexplained by the market is in the spirit of both the ICAPM and the Ross's arbitrage pricing theory.

The behavioralists are not impressed by the evidence for a risk-based explanation of the failures of the CAPM. They typically concede that the three-factor model captures covariation in returns missed by the market return and that it picks up much of the size and value effects in average returns left unexplained by the CAPM. But their view is that the average return premium associated with the model's book-to-market factor—which does the heavy lifting in the improvements to the CAPM—is itself the result of investor overreaction that happens to be correlated across firms in a way that just looks like a risk story. In short, in the behavioral view, the market tries to set CAPM prices, and violations of the CAPM are due to mispricing.

The conflict between the behavioral irrational pricing story and the rational risk story for the empirical failures of the CAPM leaves us at a timeworn impasse. Fama (1970) emphasizes that the hypothesis that prices properly reflect available information must be tested in the context of a model of expected returns, like the CAPM. Intuitively, to test whether prices are rational, one must take a stand on what the market is trying to do in setting prices—that is, what is risk and what is the relation between expected return and risk? When tests reject the CAPM, one cannot say whether the problem is its assumption that prices are rational (the behavioral view) or violations of other assumptions that are also necessary to produce the CAPM (our position).

Fortunately, for some applications, the way one uses the three-factor model does not depend on one's view about whether its average return premiums are the rational result of underlying state variable risks, the result of irrational investor behavior or sample specific results of chance. For example, when measuring the response of stock prices to new information or when evaluating the performance of managed portfolios, one wants to account for known patterns in returns and average returns for the period examined, whatever their source. Similarly, when estimating the cost of equity capital, one might be unconcerned with whether expected return premiums are rational or irrational since they are in either case part of the opportunity cost of equity capital (Stein, 1996). But the cost of capital is forward looking, so if the premiums are sample specific they are irrelevant.

The three-factor model is hardly a panacea. Its most serious problem is the momentum effect of Jegadeesh and Titman (1993). Stocks that do well relative to the market over the last three to twelve months tend to continue to do well for the next few months, and stocks that do poorly continue to do poorly. This momentum effect is distinct from the value effect captured by book-to-market equity and other price ratios. Moreover, the momentum effect is left unexplained by the three-factor model, as well as by the CAPM. Following Carhart (1997), one response is to add a momentum factor (the difference between the returns on diversified portfolios of short-term winners and losers) to the three-factor model. This step is again legitimate in applications where the goal is to abstract from known patterns in average returns to uncover information-specific or manager-specific effects. But since the momentum effect is short-lived, it is largely irrelevant for estimates of the cost of equity capital.

Another strand of research points to problems in both the three-factor model and the CAPM. Frankel and Lee (1998), Dechow, Hutton and Sloan (1999), Piotroski (2000) and others show that in portfolios formed on price ratios like book-to-market equity, stocks with higher expected cash flows have higher average returns that are not captured by the three-factor model or the CAPM. The authors interpret their results as evidence that stock prices are irrational, in the sense that they do not reflect available information about expected profitability.

In truth, however, one can't tell whether the problem is bad pricing or a bad asset pricing model. A stock's price can always be expressed as the present value of expected future cash flows discounted at the expected return on the stock (Campbell and Shiller, 1989; Vuolteenaho, 2002). It follows that if two stocks have the same price, the one with higher expected cash flows must have a higher expected return. This holds true whether pricing is rational or irrational. Thus, when one observes a positive relation between expected cash flows and expected returns that is left unexplained by the CAPM or the three-factor model, one can't tell whether it is the result of irrational pricing or a misspecified asset pricing model.

The Market Proxy Problem

Roll (1977) argues that the CAPM has never been tested and probably never will be. The problem is that the market portfolio at the heart of the model is theoretically and empirically elusive. It is not theoretically clear which assets (for example, human capital) can legitimately be excluded from the market portfolio, and data availability substantially limits the assets that are included. As a result, tests of the CAPM are forced to use proxies for the market portfolio, in effect testing whether the proxies are on the minimum variance frontier. Roll argues that because the tests use proxies, not the true market portfolio, we learn nothing about the CAPM.

We are more pragmatic. The relation between expected return and market beta of the CAPM is just the minimum variance condition that holds in any efficient portfolio, applied to the market portfolio. Thus, if we can find a market proxy that is on the minimum variance frontier, it can be used to describe differences in expected returns, and we would be happy to use it for this purpose. The strong rejections of the CAPM described above, however, say that researchers have not uncovered a reasonable market proxy that is close to the minimum variance frontier. If researchers are constrained to reasonable proxies, we doubt they ever will.

Our pessimism is fueled by several empirical results. Stambaugh (1982) tests the CAPM using a range of market portfolios that include, in addition to U.S. common stocks, corporate and government bonds, preferred stocks, real estate and other consumer durables. He finds that tests of the CAPM are not sensitive to expanding the market proxy beyond common stocks, basically because the volatility of expanded market returns is dominated by the volatility of stock returns.

One need not be convinced by Stambaugh's (1982) results since his market proxies are limited to U.S. assets. If international capital markets are open and asset prices conform to an international version of the CAPM, the market portfolio should include international assets. Fama and French (1998) find, however, that betas for a global stock market portfolio cannot explain the high average returns observed around the world on stocks with high book-to-market or high earnings-price ratios.

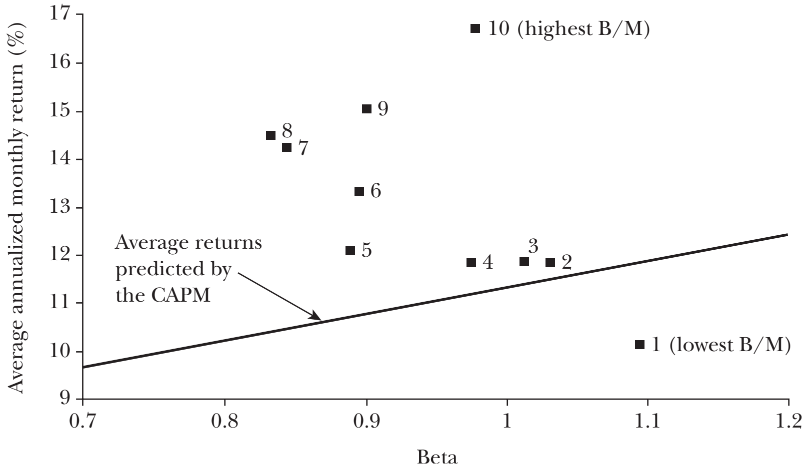
A major problem for the CAPM is that portfolios formed by sorting stocks on price ratios produce a wide range of average returns, but the average returns are not positively related to market betas (Lakonishok, Shleifer and Vishny, 1994; Fama and French, 1996, 1998). The problem is illustrated in Figure 3, which shows average returns and betas (calculated with respect to the CRSP value-weight portfolio of NYSE, AMEX and NASDAQ stocks) for July 1963 to December 2003 for ten portfolios of U.S. stocks formed annually on sorted values of the book-to-market equity ratio (B/M).[6]

Average returns on the B/M portfolios increase almost monotonically, from 10.1 percent per year for the lowest B/M group (portfolio 1) to an impressive 16.7 percent for the highest (portfolio 10). But the positive relation between beta and average return predicted by the CAPM is notably absent. For example, the portfolio with the lowest book-to-market ratio has the highest beta but the lowest average return. The estimated beta for the portfolio with the highest book-to-market ratio and the highest average return is only 0.98. With an average annualized value of the riskfree interest rate, R_f, of 5.8 percent and an average annualized market premium, $R_M - R_f$, of 11.3 percent, the Sharpe-Lintner CAPM predicts an average return of 11.8 percent for the lowest B/M portfolio and 11.2 percent for the highest, far from the observed values, 10.1 and 16.7 percent. For the Sharpe-Lintner model to "work" on these portfolios, their market betas must change dramatically, from 1.09 to 0.78 for the lowest B/M portfolio and from 0.98 to 1.98 for the highest. We judge it unlikely that alternative proxies for the market portfolio will produce betas and a market premium that can explain the average returns on these portfolios.

It is always possible that researchers will redeem the CAPM by finding a reasonable proxy for the market portfolio that is on the minimum variance frontier. We emphasize, however, that this possibility cannot be used to justify the way the CAPM is currently applied. The problem is that applications typically use the same

[6] Stock return data are from CRSP, and book equity data are from Compustat and the Moody's Industrials, Transportation, Utilities and Financials manuals. Stocks are allocated to ten portfolios at the end of June of each year t (1963 to 2003) using the ratio of book equity for the fiscal year ending in calendar year $t - 1$, divided by market equity at the end of December of $t - 1$. Book equity is the book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit (if available), minus the book value of preferred stock. Depending on availability, we use the redemption, liquidation or par value (in that order) to estimate the book value of preferred stock. Stockholders' equity is the value reported by Moody's or Compustat, if it is available. If not, we measure stockholders' equity as the book value of common equity plus the par value of preferred stock or the book value of assets minus total liabilities (in that order). The portfolios for year t include NYSE (1963–2003), AMEX (1963–2003) and NASDAQ (1972–2003) stocks with positive book equity in $t - 1$ and market equity (from CRSP) for December of $t - 1$ and June of t. The portfolios exclude securities CRSP does not classify as ordinary common equity. The breakpoints for year t use only securities that are on the NYSE in June of year t.

Figure 3
Average Annualized Monthly Return versus Beta for Value Weight Portfolios Formed on B/M, 1963–2003



market proxies, like the value-weight portfolio of U.S. stocks, that lead to rejections of the model in empirical tests. The contradictions of the CAPM observed when such proxies are used in tests of the model show up as bad estimates of expected returns in applications; for example, estimates of the cost of equity capital that are too low (relative to historical average returns) for small stocks and for stocks with high book-to-market equity ratios. In short, if a market proxy does not work in tests of the CAPM, it does not work in applications.

Conclusions

The version of the CAPM developed by Sharpe (1964) and Lintner (1965) has never been an empirical success. In the early empirical work, the Black (1972) version of the model, which can accommodate a flatter tradeoff of average return for market beta, has some success. But in the late 1970s, research begins to uncover variables like size, various price ratios and momentum that add to the explanation of average returns provided by beta. The problems are serious enough to invalidate most applications of the CAPM.

For example, finance textbooks often recommend using the Sharpe-Lintner CAPM risk-return relation to estimate the cost of equity capital. The prescription is to estimate a stock's market beta and combine it with the risk-free interest rate and the average market risk premium to produce an estimate of the cost of equity. The typical market portfolio in these exercises includes just U.S. common stocks. But empirical work, old and new, tells us that the relation between beta and average return is flatter than predicted by the Sharpe-Lintner version of the CAPM. As a result, CAPM estimates of the cost of equity for high beta stocks are too high (relative to historical average returns) and estimates for low beta stocks are too low (Friend and Blume, 1970). Similarly, if the high average returns on value stocks (with high book-to-market ratios) imply high expected returns, CAPM cost of equity estimates for such stocks are too low.[7]

The CAPM is also often used to measure the performance of mutual funds and other managed portfolios. The approach, dating to Jensen (1968), is to estimate the CAPM time-series regression for a portfolio and use the intercept (Jensen's alpha) to measure abnormal performance. The problem is that, because of the empirical failings of the CAPM, even passively managed stock portfolios produce abnormal returns if their investment strategies involve tilts toward CAPM problems (Elton, Gruber, Das and Hlavka, 1993). For example, funds that concentrate on low beta stocks, small stocks or value stocks will tend to produce positive abnormal returns relative to the predictions of the Sharpe-Lintner CAPM, even when the fund managers have no special talent for picking winners.

The CAPM, like Markowitz's (1952, 1959) portfolio model on which it is built, is nevertheless a theoretical tour de force. We continue to teach the CAPM as an introduction to the fundamental concepts of portfolio theory and asset pricing, to be built on by more complicated models like Merton's (1973) ICAPM. But we also warn students that despite its seductive simplicity, the CAPM's empirical problems probably invalidate its use in applications.

■ *We gratefully acknowledge the comments of John Cochrane, George Constantinides, Richard Leftwich, Andrei Shleifer, René Stulz and Timothy Taylor.*

[7] The problems are compounded by the large standard errors of estimates of the market premium and of betas for individual stocks, which probably suffice to make CAPM estimates of the cost of equity rather meaningless, even if the CAPM holds (Fama and French, 1997; Pastor and Stambaugh, 1999). For example, using the U.S. Treasury bill rate as the risk-free interest rate and the CRSP value-weight portfolio of publicly traded U.S. common stocks, the average value of the equity premium $R_{Mt} - R_{ft}$ for 1927–2003 is 8.3 percent per year, with a standard error of 2.4 percent. The two standard error range thus runs from 3.5 percent to 13.1 percent, which is sufficient to make most projects appear either profitable or unprofitable. This problem is, however, hardly special to the CAPM. For example, expected returns in all versions of Merton's (1973) ICAPM include a market beta and the expected market premium. Also, as noted earlier the expected values of the size and book-to-market premiums in the Fama-French three-factor model are also estimated with substantial error.

References

Ball, Ray. 1978. "Anomalies in Relationships Between Securities' Yields and Yield-Surrogates." *Journal of Financial Economics.* 6:2, pp. 103–26.

Banz, Rolf W. 1981. "The Relationship Between Return and Market Value of Common Stocks." *Journal of Financial Economics.* 9:1, pp. 3–18.

Basu, Sanjay. 1977. "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis." *Journal of Finance.* 12:3, pp. 129–56.

Bhandari, Laxmi Chand. 1988. "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence." *Journal of Finance.* 43:2, pp. 507–28.

Black, Fischer. 1972. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business.* 45:3, pp. 444–54.

Black, Fischer, Michael C. Jensen and Myron Scholes. 1972. "The Capital Asset Pricing Model: Some Empirical Tests," in *Studies in the Theory of Capital Markets.* Michael C. Jensen, ed. New York: Praeger, pp. 79–121.

Blume, Marshall. 1970. "Portfolio Theory: A Step Towards its Practical Application." *Journal of Business.* 43:2, pp. 152–74.

Blume, Marshall and Irwin Friend. 1973. "A New Look at the Capital Asset Pricing Model." *Journal of Finance.* 28:1, pp. 19–33.

Campbell, John Y. and Robert J. Shiller. 1989. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies.* 1:3, pp. 195–228.

Capaul, Carlo, Ian Rowley and William F. Sharpe. 1993. "International Value and Growth Stock Returns." *Financial Analysts Journal.* January/February, 49, pp. 27–36.

Carhart, Mark M. 1997. "On Persistence in Mutual Fund Performance." *Journal of Finance.* 52:1, pp. 57–82.

Chan, Louis K.C., Yasushi Hamao and Josef Lakonishok. 1991. "Fundamentals and Stock Returns in Japan." *Journal of Finance.* 46:5, pp. 1739–789.

DeBondt, Werner F. M. and Richard H. Thaler. 1987. "Further Evidence on Investor Overreaction and Stock Market Seasonality." *Journal of Finance.* 42:3, pp. 557–81.

Dechow, Patricia M., Amy P. Hutton and Richard G. Sloan. 1999. "An Empirical Assessment of the Residual Income Valuation Model." *Journal of Accounting and Economics.* 26:1, pp. 1–34.

Douglas, George W. 1968. *Risk in the Equity Markets: An Empirical Appraisal of Market Efficiency.* Ann Arbor, Michigan: University Microfilms, Inc.

Elton, Edwin J., Martin J. Gruber, Sanjiv Das and Matt Hlavka. 1993. "Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios." *Review of Financial Studies.* 6:1, pp. 1–22.

Fama, Eugene F. 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance.* 25:2, pp. 383–417.

Fama, Eugene F. 1996. "Multifactor Portfolio Efficiency and Multifactor Asset Pricing." *Journal of Financial and Quantitative Analysis.* 31:4, pp. 441–65.

Fama, Eugene F. and Kenneth R. French. 1992. "The Cross-Section of Expected Stock Returns." *Journal of Finance.* 47:2, pp. 427–65.

Fama, Eugene F. and Kenneth R. French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics.* 33:1, pp. 3–56.

Fama, Eugene F. and Kenneth R. French. 1995. "Size and Book-to-Market Factors in Earnings and Returns." *Journal of Finance.* 50:1, pp. 131–55.

Fama, Eugene F. and Kenneth R. French. 1996. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance.* 51:1, pp. 55–84.

Fama, Eugene F. and Kenneth R. French. 1997. "Industry Costs of Equity." *Journal of Financial Economics.* 43:2 pp. 153–93.

Fama, Eugene F. and Kenneth R. French. 1998. "Value Versus Growth: The International Evidence." *Journal of Finance.* 53:6, pp. 1975–999.

Fama, Eugene F. and James D. MacBeth. 1973. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy.* 81:3, pp. 607–36.

Frankel, Richard and Charles M.C. Lee. 1998. "Accounting Valuation, Market Expectation, and Cross-Sectional Stock Returns." *Journal of Accounting and Economics.* 25:3 pp. 283–319.

Friend, Irwin and Marshall Blume. 1970. "Measurement of Portfolio Performance under Uncertainty." *American Economic Review.* 60:4, pp. 607–36.

Gibbons, Michael R. 1982. "Multivariate Tests of Financial Models: A New Approach." *Journal of Financial Economics.* 10:1, pp. 3–27.

Gibbons, Michael R., Stephen A. Ross and Jay Shanken. 1989. "A Test of the Efficiency of a Given Portfolio." *Econometrica.* 57:5, pp. 1121–152.

Haugen, Robert. 1995. *The New Finance: The*

Case against Efficient Markets. Englewood Cliffs, N.J.: Prentice Hall.

Jegadeesh, Narasimhan and Sheridan Titman. 1993. "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *Journal of Finance.* 48:1, pp. 65–91.

Jensen, Michael C. 1968. "The Performance of Mutual Funds in the Period 1945–1964." *Journal of Finance.* 23:2, pp. 389–416.

Kothari, S. P., Jay Shanken and Richard G. Sloan. 1995. "Another Look at the Cross-Section of Expected Stock Returns." *Journal of Finance.* 50:1, pp. 185–224.

Lakonishok, Josef and Alan C. Shapiro. 1986. Systemaitc Risk, Total Risk, and Size as Determinants of Stock Market Returns." *Journal of Banking and Finance.* 10:1, pp. 115–32.

Lakonishok, Josef, Andrei Shleifer and Robert W. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance.* 49:5, pp. 1541–578.

Lintner, John. 1965. "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *Review of Economics and Statistics.* 47:1, pp. 13–37.

Loughran, Tim and Jay. R. Ritter. 1995. "The New Issues Puzzle." *Journal of Finance.* 50:1, pp. 23–51.

Markowitz, Harry. 1952. "Portfolio Selection." *Journal of Finance.* 7:1, pp. 77–99.

Markowitz, Harry. 1959. *Portfolio Selection: Efficient Diversification of Investments.* Cowles Foundation Monograph No. 16. New York: John Wiley & Sons, Inc.

Merton, Robert C. 1973. "An Intertemporal Capital Asset Pricing Model." *Econometrica.* 41:5, pp. 867–87.

Miller, Merton and Myron Scholes. 1972. "Rates of Return in Relation to Risk: A Reexamination of Some Recent Findings," in *Studies in the Theory of Capital Markets.* Michael C. Jensen, ed. New York: Praeger, pp. 47–78.

Mitchell, Mark L. and Erik Stafford. 2000. "Managerial Decisions and Long-Term Stock Price Performance." *Journal of Business.* 73:3, pp. 287–329.

Pastor, Lubos and Robert F. Stambaugh. 1999. "Costs of Equity Capital and Model Mispricing." *Journal of Finance.* 54:1, pp. 67–121.

Piotroski, Joseph D. 2000. "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers." *Journal of Accounting Research.* 38:Supplement, pp. 1–51.

Reinganum, Marc R. 1981. "A New Empirical Perspective on the CAPM." *Journal of Financial and Quantitative Analysis.* 16:4, pp. 439–62.

Roll, Richard. 1977. "A Critique of the Asset Pricing Theory's Tests' Part I: On Past and Potential Testability of the Theory." *Journal of Financial Economics.* 4:2, pp. 129–76.

Rosenberg, Barr, Kenneth Reid and Ronald Lanstein. 1985. "Persuasive Evidence of Market Inefficiency." *Journal of Portfolio Management.* Spring, 11, pp. 9–17.

Ross, Stephen A. 1976. "The Arbitrage Theory of Capital Asset Pricing." *Journal of Economic Theory.* 13:3, pp. 341–60.

Sharpe, William F. 1964. "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk." *Journal of Finance.* 19:3, pp. 425–42.

Stambaugh, Robert F. 1982. "On The Exclusion of Assets from Tests of the Two-Parameter Model: A Sensitivity Analysis." *Journal of Financial Economics.* 10:3, pp. 237–68.

Stattman, Dennis. 1980. "Book Values and Stock Returns." *The Chicago MBA: A Journal of Selected Papers.* 4, pp. 25–45.

Stein, Jeremy. 1996. "Rational Capital Budgeting in an Irrational World." *Journal of Business.* 69:4, pp. 429–55.

Tobin, James. 1958. "Liquidity Preference as Behavior Toward Risk." *Review of Economic Studies.* 25:2, pp. 65–86.

Vuolteenaho, Tuomo. 2002. "What Drives Firm Level Stock Returns?" *Journal of Finance.* 57:1, pp. 233–64.

THE JOURNAL OF FINANCE • VOL. LVII, NO. 2 • APRIL 2002

The Equity Premium

EUGENE F. FAMA and KENNETH R. FRENCH*

ABSTRACT

We estimate the equity premium using dividend and earnings growth rates to measure the expected rate of capital gain. Our estimates for 1951 to 2000, 2.55 percent and 4.32 percent, are much lower than the equity premium produced by the average stock return, 7.43 percent. Our evidence suggests that the high average return for 1951 to 2000 is due to a decline in discount rates that produces a large unexpected capital gain. Our main conclusion is that the average stock return of the last half-century is a lot higher than expected.

THE EQUITY PREMIUM—the difference between the expected return on the market portfolio of common stocks and the risk-free interest rate—is important in portfolio allocation decisions, estimates of the cost of capital, the debate about the advantages of investing Social Security funds in stocks, and many other applications. The average return on a broad portfolio of stocks is typically used to estimate the expected market return. The average real return for 1872 to 2000 on the S&P index (a common proxy for the market portfolio, also used here) is 8.81 percent per year. The average real return on six-month commercial paper (a proxy for the risk-free interest rate) is 3.24 percent. This large spread (5.57 percent) between the average stock return and the interest rate is the source of the so-called equity premium puzzle: Stock returns seem too high given the observed volatility of consumption (Mehra and Prescott (1985)).

We use fundamentals (dividends and earnings) to estimate the expected stock return. Along with other evidence, the expected return estimates from fundamentals help us judge whether the realized average return is high or low relative to the expected value.

The logic of our approach is straightforward. The average stock return is the average dividend yield plus the average rate of capital gain:

$$A(R_t) = A(D_t/P_{t-1}) + A(GP_t), \tag{1}$$

* Fama is from the University of Chicago and French is from Dartmouth College. The comments of John Campbell, John Cochrane, Kent Daniel, John Heaton, Jay Ritter, Andrei Shleifer, Rex Sinquefield, Tuomo Vuolteenaho, Paul Zarowin, and seminar participants at Boston College, Dartmouth College, the NBER, Purdue University, the University of Chicago, and Washington University have been helpful. Richard Green (the editor) and the two referees get special thanks.

where D_t is the dividend for year t, P_{t-1} is the price at the end of year $t-1$, $GP_t = (P_t - P_{t-1})/P_{t-1}$ is the rate of capital gain, and $A(\ )$ indicates an average value. (Throughout the paper, we refer to D_t/P_{t-1} as the dividend yield and D_t/P_t is the dividend–price ratio. Similarly, Y_t/P_{t-1}, the ratio of earnings for year t to price at the end of year $t-1$, is the earnings yield and Y_t/P_t is the earnings–price ratio.)

Suppose the dividend–price ratio, D_t/P_t, is stationary (mean reverting). Stationarity implies that if the sample period is long, the compound rate of dividend growth approaches the compound rate of capital gain. Thus, an alternative estimate of the expected stock return is

$$A(RD_t) = A(D_t/P_{t-1}) + A(GD_t), \tag{2}$$

where $GD_t = (D_t - D_{t-1})/D_{t-1}$ is the growth rate of dividends. We call (2) the dividend growth model.

The logic that leads to (2) applies to any variable that is cointegrated with the stock price. For example, the dividend–price ratio may be nonstationary because firms move away from dividends toward share repurchases as a way of returning earnings to stockholders. But if the earnings–price ratio, Y_t/P_t, is stationary, the average growth rate of earnings, $A(GY_t) = A((Y_t - Y_{t-1})/Y_{t-1})$, is an alternative estimate of the expected rate of capital gain. And $A(GY_t)$ can be combined with the average dividend yield to produce another estimate of the expected stock return:

$$A(RY_t) = A(D_t/P_{t-1}) + A(GY_t). \tag{3}$$

We call (3) the earnings growth model.[1]

We should be clear about the expected return concept targeted by (1), (2), and (3). D_t/P_t and Y_t/P_t vary through time because of variation in the conditional (point-in-time) expected stock return and the conditional expected growth rates of dividends and earnings (see, e.g., Campbell and Shiller (1989)). But if the stock return and the growth rates are stationary (they have constant unconditional means), D_t/P_t and Y_t/P_t are stationary. Then, like the average return (1), the dividend and earnings growth models (2) and (3) provide estimates of the unconditional expected stock return. In short, the focus of the paper is estimates of the unconditional expected stock return.

The estimate of the expected real equity premium for 1872 to 2000 from the dividend growth model (2) is 3.54 percent per year. The estimate from the average stock return, 5.57 percent, is almost 60 percent higher. The difference between the two is largely due to the last 50 years. The equity premium for 1872 to 1950 from the dividend growth model, 4.17 percent per year, is close to the estimate from the average return, 4.40 percent. In con-

[1] Motivated by the model in Lettau and Ludvigson (2001), one can argue that if the ratio of consumption to stock market wealth is stationary, the average growth rate of consumption is another estimate of the expected rate of capital gain. We leave this path to future work.

trast, the equity premium for 1951 to 2000 produced by the average return, 7.43 percent per year, is almost three times the estimate, 2.55 percent, from (2). The estimate of the expected real equity premium for 1951 to 2000 from the earnings growth model (3), 4.32 percent per year, is larger than the estimate from the dividend growth model (2). But the earnings growth estimate is still less than 60 percent of the estimate from the average return.

Three types of evidence suggest that the lower equity premium estimates for 1951 to 2000 from fundamentals are closer to the expected premium. (a) The estimates from fundamentals are more precise. For example, the standard error of the estimate from the dividend growth model is less than half the standard error of the estimate from the average return. (b) The Sharpe ratio for the equity premium from the average stock return for 1951 to 2000 is just about double that for 1872 to 1950. In contrast, the equity premium from the dividend growth model has a similar Sharpe ratio for 1872 to 1950 and 1951 to 2000. (c) Most important, valuation theory specifies relations among the book-to-market ratio, the return on investment, and the cost of equity capital (the expected stock return). The estimates of the expected stock return for 1951 to 2000 from the dividend and earnings growth models line up with other fundamentals in the way valuation theory predicts. But the book-to-market ratio and the return on investment suggest that the expected return estimate from the average stock return is too high.

Our motivation for the dividend growth model (2) is simpler and more general, but (2) can be viewed as the expected stock return estimate of the Gordon (1962) model. Our work is thus in the spirit of a growing literature that uses valuation models to estimate expected returns (e.g., Blanchard (1993), Claus and Thomas (2001), and Gebhardt, Lee, and Swaminathan (2001)). Claus and Thomas and Gebhardt, Lee, and Swaminathan use forecasts by security analysts to estimate expected cash flows. Their analyst forecasts cover short periods (1985 to 1998 and 1979 to 1995). We use realized dividends and earnings from 1872 to 2000. This 129-year period provides a long perspective, which is important for judging the competing expected return estimates from fundamentals and realized stock returns. Moreover, though the issue is controversial (Keane and Runkle (1998)), Claus and Thomas find that analyst forecasts are biased; they tend to be substantially above observed growth rates. The average growth rates of dividends and earnings we use are unbiased estimates of expected growth rates.

Like us, Blanchard (1993) uses dividend growth rates to estimate the expected rate of capital gain, which he combines with an expected dividend yield to estimate the expected stock return. But his focus is different and his approach is more complicated than ours. He is interested in the path of the conditional expected stock return. His conditional expected return is the sum of the fitted values from time-series regressions of the realized dividend yield and a weighted average of 20 years of future dividend growth rates on four predetermined variables (the dividend yield, the real rate of capital gain, and the levels of interest rates and inflation). He focuses on describing the path of the conditional expected return in terms of his four explanatory variables.

In contrast, our prime interest is the unconditional expected return, which we estimate more simply as the sum of the average dividend yield and the average growth rate of dividends or earnings. This approach is valid if the dividend–price and earnings–price ratios are stationary. And we argue below that it continues to produce estimates of the average expected stock return when the price ratios are subject to reasonable forms of nonstationarity. Given its simplicity and generality, our approach is an attractive addition to the research toolbox for estimating the expected stock return.

Moreover, our focus is comparing alternative estimates of the unconditional expected stock return over the long 1872 to 2000 period, and explaining why the expected return estimates for 1951 to 2000 from fundamentals are much lower than the average return. Our evidence suggests that much of the high return for 1951 to 2000 is unexpected capital gain, the result of a decline in discount rates.

Specifically, the dividend–price and earnings–price ratios fall from 1950 to 2000; the cumulative percent capital gain for the period is more than three times the percent growth in dividends or earnings. All valuation models agree that the two price ratios are driven by expectations about future returns (discount rates) and expectations about dividend and earnings growth. Confirming Campbell (1991), Cochrane (1994), and Campbell and Shiller (1998), we find that dividend and earnings growth rates for 1950 to 2000 are largely unpredictable. Like Campbell and Shiller (1998), we thus infer that the decline in the price ratios is mostly due to a decline in expected returns. Some of this decline is probably expected, the result of reversion of a high 1950 conditional expected return to the unconditional mean. But most of the decline in the price ratios seems to be due to the unexpected decline of expected returns to ending values far below the mean.

The paper proceeds as follows. The main task, addressed in Sections I and II, is to compare and evaluate the estimates of the unconditional annual expected stock return provided by the average stock return and the dividend and earnings growth models. Section III then considers the issues that arise if the goal is to estimate the long-term expected growth of wealth, rather than the unconditional expected annual (simple) return. Section IV concludes.

I. The Unconditional Annual Expected Stock Return

Table I shows estimates of the annual expected real equity premium for 1872 to 2000. The market portfolio is the S&P 500 and its antecedents. The deflator is the Producer Price Index until 1925 (from Shiller (1989)) and the Consumer Price Index thereafter (from Ibbotson Associates). The risk-free interest rate is the annual real return on six-month commercial paper, rolled over at midyear. The risk-free rate and S&P earnings data are from Shiller, updated by Vuolteenaho (2000) and us. Beginning in 1925, we construct S&P book equity data from the book equity data in Davis, Fama, and French (2000), expanded to include all NYSE firms. The data on dividends, prices, and returns for 1872 to 1925 are from Shiller. Shiller's annual data on the

Table I

Real Equity Premium and Related Statistics for the S&P Portfolio

The inflation rate for year t is $Inf_t = L_t/L_{t-1} - 1$, where L_t is the price level at the end of year t. The real return for year t on six-month (three-month for the year 2000) commercial paper (rolled over at midyear) is F_t. The nominal values of book equity and price for the S&P index at the end of year t are b_t and p_t. Nominal S&P dividends and earnings for year t are d_t and y_t. Real rates of growth of dividends, earnings, and the stock price are $GD_t = (d_t/d_{t-1})*(L_{t-1}/L_t) - 1$, $GY_t = (y_t/y_{t-1})*(L_{t-1}/L_t) - 1$, and $GP_t = (p_t/p_{t-1})*(L_{t-1}/L_t) - 1$. The real dividend yield is $D_t/P_{t-1} = (d_t/p_{t-1})*(L_{t-1}/L_t)$. The real income return on investment is $Y_t/B_{t-1} = (1 + y_t/b_{t-1})*(L_{t-1}/L_t) - 1$. The dividend growth estimate of the real S&P return for t is $RD_t = D_t/P_{t-1} + GD_t$, the earnings growth estimate is $RY_t = D_t/P_{t-1} + GY_t$, and R_t is the realized real S&P return. The dividend and earnings growth estimates of the real equity premium for year t are $RXD_t = RD_t - F_t$ and $RXY_t = RY_t - F_t$, and $RX_t = R_t - F_t$ is the real equity premium from the realized real return. The Sharpe ratio for $RD_t - F_t$ (the mean of $RD_t - F_t$ divided by the standard deviation of R_t) is SD, SY is the Sharpe ratio for $RY_t - F_t$ (the mean of $RY_t - F_t$ divided by the standard deviation of R_t), and SR is the Sharpe ratio for $R_t - F_t$ (the mean of $R_t - F_t$ divided by the standard deviation of R_t). Except for the Sharpe ratios, all variables are expressed as percents, that is, they are multiplied by 100.

	Inf_t	F_t	D_t/P_{t-1}	GD_t	GY_t	GP_t	RD_t	RY_t	R_t	RXD_t	RXY_t	RX_t	SD	SY	SR
Means of annual values of variables															
1872–2000	2.16	3.24	4.70	2.08	NA	4.11	6.78	NA	8.81	3.54	NA	5.57	0.20	NA	0.31
1872–1950	0.99	3.90	5.34	2.74	NA	2.96	8.07	NA	8.30	4.17	NA	4.40	0.22	NA	0.23
1951–2000	4.00	2.19	3.70	1.05	2.82	5.92	4.74	6.51	9.62	2.55	4.32	7.43	0.15	0.25	0.44
Standard deviations of annual values of variables															
1872–2000	7.51	8.48	1.39	12.37	NA	17.83	12.56	NA	18.03	13.00	NA	18.51			
1872–1950	9.11	10.63	1.12	15.28	NA	18.48	15.41	NA	18.72	16.02	NA	19.57			
1951–2000	3.11	2.46	1.17	5.09	13.79	16.77	5.21	13.51	17.03	5.62	14.02	16.73			
Means of annual continuously compounded returns and growth rates															
1872–2000	1.86	2.87		1.34	NA	2.48			7.00						
1872–1950	0.59	3.33		1.60	NA	1.22			6.41						
1951–2000	3.88	2.14		0.92	1.89	4.46			7.94						

	b_t/p_t	RD_t	RY_t	R_t	Y_t/B_{t-1}
Means of annual values of variables					
1951–2000	0.66	4.74	6.51	9.62	7.60

level of the S&P (used to compute returns and other variables involving price) are averages of daily January values. The S&P dividend, price, and return data for 1926 to 2000 are from Ibbotson Associates, and the returns for 1926 to 2000 are true annual returns.

Without showing the details, we can report that the CRSP value-weight portfolio of NYSE, AMEX, and Nasdaq stocks produces average returns and dividend growth estimates of the expected return close to the S&P estimates for periods after 1925 when both indices are available. What one takes to be the risk-free rate has a bigger effect. For example, substituting the one-month Treasury bill rate for the six-month commercial paper rate causes estimates of the annual equity premium for 1951 to 2000 to rise by about one percent. But for our main task—comparing equity premium estimates from (1), (2), and (3)—differences in the risk-free rate are an additive constant that does not affect inferences.

One can estimate expected returns in real or nominal terms. Since portfolio theory says the goal of investment is consumption, real returns seem more relevant, and only results for real returns are shown. Because of suspicions about the quality of the price deflator during the early years of 1872 to 2000, we have replicated the results for nominal returns. They support all the inferences from real returns.

The dividend and earnings growth models (2) and (3) assume that the market dividend–price and earnings–price ratios are stationary. The first three annual autocorrelations of D_t/P_t for 1872 to 2000 are 0.73, 0.51, and 0.47. For the 1951 to 2000 period that occupies much of our attention, the autocorrelations are 0.83, 0.72, and 0.69. The autocorrelations are large, but their decay is roughly like that of a stationary first-order autoregression (AR1). This is in line with formal evidence (Fama and French (1988), Cochrane (1994), and Lamont (1998)) that the market dividend–price ratio is highly autocorrelated but slowly mean-reverting. S&P earnings data for the early years of 1872 to 2000 are of dubious quality (Shiller (1989)), so we estimate expected returns with the earnings growth model (3) only for 1951 to 2000. The first three autocorrelations of Y_t/P_t for 1951 to 2000, 0.80, 0.70, and 0.61, are again roughly like those of a stationary AR1.

We emphasize, however, that our tests are robust to reasonable nonstationarity of D_t/P_t and Y_t/P_t. It is not reasonable that the expected stock return and the expected growth rates of dividends and earnings that drive D_t/P_t and Y_t/P_t are nonstationary processes that can wander off to infinity. But nonstationarity of D_t/P_t and Y_t/P_t due to structural shifts in productivity or preferences that permanently change the expected return or the expected growth rates is reasonable. Such regime shifts are not a problem for the expected return estimates from (2) and (3), as long as D_t/P_t and Y_t/P_t mean-revert within regimes. If the regime shift is limited to expected dividend and earnings growth rates, the permanent change in expected growth rates is offset by a permanent change in the expected dividend yield, and (2) and (3) continue to estimate the (stationary) expected stock return. (An Appendix, available on request, provides an example.) If there is a permanent shift in the expected stock return, it is nonstationary, but like the average return in (1), the dividend and earnings growth models in (2) and (3) estimate the average expected return during the sample period.

Indeed, an advantage of the expected return estimates from fundamentals is that they are likely to be less sensitive than the average return to long-lived shocks to dividend and earnings growth rates or the expected stock return. For example, a permanent shift in the expected return affects the average dividend yield, which is common to the three expected return estimates, but it produces a shock to the capital gain term in the average return in (1) that is not shared by the estimates in (2) and (3). In short, the estimates of the expected stock return from fundamentals are likely to be more precise than the average stock return.

A. The Equity Premium

For much of the period from 1872 to 2000—up to about 1950—the dividend growth model and the average stock return produce similar estimates of the expected return. Thereafter, the two estimates diverge. To illustrate, Table I shows results for 1872 to 1950 (79 years) and 1951 to 2000 (50 years). The year 1950 is a big year, with a high real stock return (23.40 percent), and high dividend and earnings growth estimates of the return (29.96 percent and 24.00 percent). But because the three estimates of the 1950 return are similarly high, the ordering of expected return estimates, and the inferences we draw from them, are unaffected by whether 1950 is allocated to the earlier or the later period. Indeed, pushing the 1950 break-year backward or forward several years does not affect our inferences.

For the earlier 1872 to 1950 period, there is not much reason to favor the dividend growth estimate of the expected stock return over the average return. Precision is not an issue; the standard errors of the two estimates are similar (1.74 percent and 2.12 percent), the result of similar standard deviations of the annual dividend growth rate and the rate of capital gain, 15.28 percent and 18.48 percent. Moreover, the dividend growth model and the average return provide similar estimates of the expected annual real return for 1872 to 1950, 8.07 percent and 8.30 percent. Given similar estimates of the expected return, the two approaches produce similar real equity premiums for 1872 to 1950, 4.17 percent (dividend growth model) and 4.40 percent (stock returns).

The competition between the dividend growth model and the average stock return is more interesting for 1951 to 2000. The dividend growth estimate of the 1951 to 2000 expected return, 4.74 percent, is less than half the average return, 9.62 percent. The dividend growth estimate of the equity premium, 2.55 percent, is 34 percent of the estimate from returns, 7.43 percent. The 1951 to 2000 estimates of the expected stock return and the equity premium from the earnings growth model, 6.51 percent and 4.32 percent, are higher than for the dividend growth model. But they are well below the estimates from the average return, 9.62 percent and 7.43 percent.

B. Evaluating the Expected Return Estimates for 1951 to 2000

We judge that the estimates of the expected stock return for 1951 to 2000 from fundamentals are closer to the true expected value, for three reasons.

(a) The expected return estimates from the dividend and earnings growth models are more precise than the average return. The standard error of the dividend growth estimate of the expected return for 1951 to 2000 is 0.74 percent, versus 2.43 percent for the average stock return. Since earnings growth is more volatile than dividend growth, the standard error of the expected return from the earnings growth model, 1.93 percent, is higher than the estimate from the dividend growth model, but it is smaller than the 2.43 percent standard error of the average stock return. Claus and Thomas (2001) also argue that expected return estimates from fundamentals are more precise than average returns, but they provide no direct evidence.

(b) Table I shows Sharpe ratios for the three equity premium estimates. Only the average premium in the numerator of the Sharpe ratio differs for the three estimates. The denominator for all three is the standard deviation of the annual stock return. The Sharpe ratio for the dividend growth estimate of the equity premium for 1872 to 1950, 0.22, is close to that produced by the average stock return, 0.23. More interesting, the Sharpe ratio for the equity premium for 1951 to 2000 from the dividend growth model, 0.15, is lower than but similar to that for 1872 to 1950. The Sharpe ratio for the 1951 to 2000 equity premium from the earnings growth model, 0.25, is somewhat higher than the dividend growth estimate, 0.15, but it is similar to the estimates for 1872 to 1950 from the dividend growth model, 0.22, and the average return, 0.23.

In asset pricing theory, the Sharpe ratio is related to aggregate risk aversion. The Sharpe ratios for the 1872 to 1950 and 1951 to 2000 equity premiums from the dividend growth model and the earnings growth model suggest that aggregate risk aversion is roughly similar in the two periods. In contrast, though return volatility falls a bit, the equity premium estimate from the average stock return increases from 4.40 percent for 1872 to 1950 to 7.43 percent for 1951 to 2000, and its Sharpe ratio about doubles, from 0.23 to 0.44. It seems implausible that risk aversion increases so much from the earlier to the later period.

(c) Most important, the behavior of other fundamentals favors the dividend and earnings growth models. The average ratio of the book value of equity to the market value of equity for 1951 to 2000 is 0.66, the book-to-market ratio B_t/P_t is never greater than 1.12, and it is greater than 1.0 for only 6 years of the 50-year period. Since, on average, the market value of equity is substantially higher than its book value, it seems safe to conclude that, on average, the expected return on investment exceeds the cost of capital.

Suppose investment at time $t-1$ generates a stream of equity earnings for $t, t+1, \ldots, t+N$ with a constant expected value. The average income return on book equity, $A(Y_t/B_{t-1})$, is then an estimate of the expected return on equity's share of assets. It is an unbiased estimate when N is infinite and it is upward biased when N is finite. In either case, if the expected return on investment exceeds the cost of capital, we should find that (except for sampling error) the average income return on book equity is greater than estimates of the cost of equity capital (the expected stock return):

$$A(Y_t/B_{t-1}) > E(R). \tag{4}$$

Table I shows that (4) is confirmed when we use the dividend and earnings growth models to estimate the expected real stock return for 1951 to 2000. The estimates of $E(R)$, 4.74 percent (dividend growth model) and 6.51 percent (earnings growth model), are below 7.60 percent, the average real income return on book equity, $A(Y_t/B_{t-1})$. In contrast, the average real stock return for 1951 to 2000, 9.62 percent, exceeds the average income return by more than 2 percent. An expected stock return that exceeds the expected income return on book equity implies that the typical corporate investment has a negative net present value. This is difficult to reconcile with an average book-to-market ratio substantially less than one.

To what extent are our results new? Using analyst forecasts of expected cash flows and a more complicated valuation model, Claus and Thomas (2001) produce estimates of the expected stock return for 1985 to 1998 far below the average return. Like us, they argue that the estimates from fundamentals are closer to the true expected return. We buttress this conclusion with new results on three fronts. (a) The long-term perspective provided by the evidence that, for much of the 1872 to 2000 period, average returns and fundamentals produce similar estimates of the expected return. (b) Direct evidence that the expected return estimates for 1951 to 2000 from fundamentals are more precise. (c) Sharpe ratios and evidence on how the alternative expected return estimates line up with the income return on investment. These new results provide support for the expected return estimates from fundamentals, and for the more specific inference that the average stock return for 1951 to 2000 is above the expected return.

II. Unexpected Capital Gains

Valuation theory suggests three potential explanations for why the 1951 to 2000 average stock return is larger than the expected return. (a) Dividend and earnings growth for 1951 to 2000 is unexpectedly high. (b) The expected (post-2000) growth rates of dividends and earnings are unexpectedly high. (c) The expected stock return (the equity discount rate) is unexpectedly low at the end of the sample period.

A. Is Dividend Growth for 1951 to 2000 Unexpectedly High?

If the prosperity of the United States over the last 50 years was not fully anticipated, dividend and earnings growth for 1951 to 2000 exceed 1950 expectations. Such unexpected in-sample growth produces unexpected capital gains. But it does not explain why the average return for 1951 to 2000 (the average dividend yield plus the average rate of capital gain) is so much higher than the expected return estimates from fundamentals (the average dividend yield plus the average growth rate of dividends or earnings). To see the point, note that unexpected in-sample dividend and earnings growth do not affect either the 1950 or the 2000 dividend–price and earnings–price ratios. (The 2000 ratios depend on post-2000 expected returns and growth rates.) Suppose D_t/P_t and E_t/P_t were the same in 1950 and 2000. Then the total percent growth in dividends and earnings during the period would be the same as the percent growth in the stock price. And (1), (2), and (3) would provide similar estimates of the expected stock return.

It is worth dwelling on this point. There is probably survivor bias in the U.S. average stock return for 1872 to 1950, as well as for 1951 to 2000. During the 1872 to 2000 period, it was not a foregone conclusion that the U.S. equity market would survive several financial panics, the Great Depression, two world wars, and the cold war. The average return for a market that survives many potentially cataclysmic challenges is likely to be higher than the expected return (Brown, Goetzmann, and Ross (1995)). But if the positive bias shows up only as higher than expected dividend and earnings growth during the sample period, there is similar survivor bias in the expected return estimates from fundamentals—a problem we do not solve. Our more limited goal is to explain why the average stock return for 1951 to 2000 is so high relative to the expected return estimates from the dividend and earnings growth models.

Since unexpected growth for 1951 to 2000 has a similar effect on the three expected return estimates, the task of explaining why the estimates are so different falls to the end-of-sample values of future expected returns and expected dividend and earnings growth. We approach the problem by first looking for evidence that expected dividend or earnings growth is high at the end of the sample period. We find none. We then argue that the large spread of capital gains over dividend and earnings growth for 1951 to 2000, or equivalently, the low end-of-sample dividend–price and earnings–price ratios, are due to an unexpected decline in expected stock returns to unusually low end-of-sample values.

B. Are Post-2000 Expected Dividend and Earnings Growth Rates Unusually High?

The behavior of dividends and earnings provides little evidence that rationally assessed (i.e., true) long-term expected growth is high at the end of the sample period. If anything, the growth rate of real dividends declines during the 1951 to 2000 period (Table II). The average growth rate for the first two decades, 1.60 percent, is higher than the average growth rates for the last three, 0.68 percent. The regressions in Table III are more formal evidence on the best forecast of post-2000 real dividend growth rates. Re-

Table II

Means of Simple Real Equity Premium and Related Statistics for the S&P Portfolio for 10-year Periods

The inflation rate for year t is $Inf_t = L_t/L_{t-1} - 1$, where L_t is the price level at the end of year t. The real return for year t on six-month (three-month for the year 2000) commercial paper (rolled over at midyear) is F_t. The nominal price of the S&P index at the end of year t is p_t. Nominal S&P dividends and earnings for year t are d_t and y_t. Real rates of growth of dividends, earnings, and the stock price are $GD_t = (d_t/d_{t-1})*(L_{t-1}/L_t) - 1$, $GY_t = (y_t/y_{t-1})*(L_{t-1}/L_t) - 1$, and $GP_t = (p_t/p_{t-1})*(L_{t-1}/L_t) - 1$. The real dividend yield is $D_t/P_{t-1} = (d_t/p_{t-1})*(L_{t-1}/L_t)$. The dividend growth estimate of the real S&P return for t is $RD_t = D_t/P_{t-1} + GD_t$, the earnings growth estimate is $RY_t = D_t/P_{t-1} + GY_t$, and R_t is the realized real S&P return. The dividend and earnings growth estimates of the real equity premium for year t are $RXD_t = RD_t - F_t$ and $RXY_t = RY_t - F_t$, and $RX_t = R_t - F_t$ is the real equity premium from the realized real return. All variables are expressed as percents, that is, they are multiplied by 100.

	Inf_t	F_t	D_t/P_{t-1}	GD_t	GY_t	GP_t	RD_t	RY_t	R_t	RXD_t	RXY_t	RX_t
1872–1880	−2.77	9.86	6.29	4.62	NA	7.13	10.91	NA	13.42	1.06	NA	3.56
1881–1890	−1.72	7.23	5.04	0.69	NA	0.04	5.73	NA	5.08	−1.51	NA	−2.15
1891–1900	0.18	5.08	4.40	4.49	NA	4.75	8.89	NA	9.15	3.81	NA	4.08
1901–1910	1.95	3.18	4.45	3.25	NA	2.33	7.70	NA	6.78	4.52	NA	3.60
1911–1920	6.82	0.82	5.70	−3.43	NA	−6.52	2.27	NA	−0.83	1.45	NA	−1.64
1921–1930	−1.70	7.41	5.72	9.07	NA	11.83	14.78	NA	17.54	7.37	NA	10.13
1931–1940	−1.23	2.80	5.31	0.36	NA	2.21	5.67	NA	7.52	2.87	NA	4.72
1941–1950	6.04	−4.57	5.90	3.02	NA	2.33	8.91	NA	8.22	13.48	NA	12.79
1951–1960	1.79	1.05	4.68	1.22	0.61	10.64	5.90	5.30	15.32	4.85	4.24	14.27
1961–1970	2.94	2.27	3.21	1.98	2.07	2.69	5.19	5.27	5.90	2.92	3.01	3.63
1971–1980	8.11	−0.30	4.04	−0.86	3.47	−1.92	3.18	7.50	2.12	3.48	7.80	2.42
1981–1990	4.51	5.32	4.19	2.32	0.37	5.40	6.51	4.56	9.59	1.19	−0.75	4.28
1991–2000	2.68	2.61	2.36	0.58	7.58	12.80	2.94	9.94	15.16	0.32	7.32	12.54

Table III

Regressions to Forecast Real Dividend and Earnings Growth Rates, GD_t and GY_t

The price level at the end of year t is L_t. The nominal values of book equity and price for the S&P index at the end of year t are b_t and p_t. Nominal S&P dividends and earnings for year t are d_t and y_t. The real dividend and earnings growth rates for year t are $GD_t = (d_t/d_{t-1})*(L_{t-1}/L_t) - 1$ and $GY_t = (y_t/y_{t-1})*(L_{t-1}/L_t) - 1$, and R_t is the realized real return on the S&P portfolio for year t. The regression intercept is *Int*, and t-Stat is the regression coefficient (*Coef*) divided by its standard error. The regression R^2 is adjusted for degrees of freedom. Except for the dividend payout ratio, d_t/y_t, all variables are expressed as percents, that is, they are multiplied by 100.

Panel A: One Year: The Regressions Forecast Real Dividend Growth, GD_t, with Variables Known at $t-1$

	Int	d_{t-1}/y_{t-1}	d_{t-1}/p_{t-1}	GD_{t-1}	GD_{t-2}	GD_{t-3}	R_{t-1}	R_{t-2}	R_{t-3}	R^2
1875–1950, $N = 76$ years										
Coef	29.56	−23.12	−2.63	−0.12	−0.07	−0.03	0.22	0.13	0.09	0.38
t-Stat	3.22	−3.17	−1.77	−1.08	−0.64	−0.29	2.24	1.37	1.01	
1951–2000, $N = 50$ years										
Coef	−2.16	2.97	0.11	−0.07	−0.20	−0.06	0.11	0.07	0.01	0.01
t-Stat	−0.40	0.33	0.16	−0.45	−1.57	−0.45	2.17	1.33	0.22	

Panel B: Two Years: The Regressions Forecast Real Dividend Growth, GD_t, with Variables Known at $t-2$

	Int	d_{t-2}/y_{t-2}	d_{t-2}/p_{t-2}	GD_{t-2}	GD_{t-3}	R_{t-2}	R_{t-3}	R^2
1875–1950, $N = 76$ years								
Coef	6.61	−11.60	0.31	−0.26	0.05	0.24	0.11	0.07
t-Stat	0.64	−1.28	0.18	−2.02	0.39	2.03	1.00	
1951–2000, $N = 50$ years								
Coef	−4.11	7.62	0.32	−0.14	−0.03	0.05	−0.01	−0.05
t-Stat	−0.73	0.81	0.46	−1.13	−0.28	0.99	−0.16	

Panel C: One Year: The Regressions Forecast Real Earnings Growth, GY_t, with Variables Known at $t-1$

	Int	Y_{t-1}/B_{t-2}	d_{t-1}/y_{t-1}	y_{t-1}/p_{t-1}	GY_{t-1}	GY_{t-2}	GY_{t-3}	R_{t-1}	R_{t-2}	R_{t-3}	R^2
1951–2000, $N = 50$ years											
Coef	5.48	0.11	13.06	−1.36	0.21	−0.13	−0.31	0.28	−0.25	0.03	0.40
t-Stat	0.33	0.11	0.52	−1.91	1.17	−0.89	−2.64	2.39	−2.18	0.26	

Panel D: Two Years: The Regressions Forecast Real Earnings Growth, GY_t, with Variables Known at $t-2$

	Int	Y_{t-2}/B_{t-3}	d_{t-2}/y_{t-2}	y_{t-2}/p_{t-2}	GY_{t-2}	GY_{t-3}	R_{t-2}	R_{t-3}	R^2
1951–2000, $N = 50$ years									
Coef	−7.60	0.46	2.05	−0.74	−0.16	−0.39	−0.31	−0.12	0.23
t-Stat	−0.43	1.66	0.76	−1.02	−0.92	−2.54	−2.59	−0.97	

gressions are shown for forecasts one year ahead (the explanatory variables for year t dividend growth are known at the end of year $t-1$) and two years ahead (the explanatory variables are known at the end of year $t-2$).

The regression for 1875 to 1950 suggests strong forecast power one year ahead. The slopes on the lagged payout ratio, the dividend–price ratio, and the stock return are close to or more than two standard errors from zero, and the regression captures 38 percent of the variance of dividend growth. Even in the 1875 to 1950 period, however, power to forecast dividend growth does not extend much beyond a year. When dividend growth for year t is explained with variables known at the end of year $t-2$, the regression R^2 falls from 0.38 to 0.07. Without showing the details, we can report that extending the forecast horizon from two to three years causes all hint of forecast power to disappear. Thus, for 1875 to 1950, the best forecast of dividend growth more than a year or two ahead is the historical average growth rate.

We are interested in post-2000 expected dividend growth, and even the short-term forecast power of the dividend regressions for 1872 to 1950 evaporates in the 1951 to 2000 period. The lagged stock return has some information ($t = 2.17$) about dividend growth one year ahead. But the 1951 to 2000 regression picks up only one percent of the variance of dividend growth. And forecast power does not improve for longer forecast horizons. Our evidence that dividend growth is essentially unpredictable during the last 50 years confirms the results in Campbell (1991), Cochrane (1991, 1994), and Campbell and Shiller (1998). If dividend growth is unpredictable, the historical average growth rate is the best forecast of future growth.

Long-term expected earnings growth also is not unusually high in 2000. There is no clear trend in real earnings growth during the 1951 to 2000 period. The most recent decade, 1991 to 2000, produces the highest average growth rate, 7.58 percent per year (Table II). But earnings growth is volatile. The standard errors of 10-year average growth rates vary around 5 percent. It is thus not surprising that 1981 to 1990, the decade immediately preceding 1991 to 2000, produces the lowest average real earnings growth rate, 0.37 percent per year.

The regressions in Table III are formal evidence on the predictability of earnings growth during the 1951 to 2000 period. There is some predictability of near-term growth, but it is largely due to transitory variation in earnings that is irrelevant for forecasting long-term earnings. In the 1951 to 2000 regression to forecast earnings growth one year ahead, the slope on the first lag of the stock return is positive (0.28, $t = 2.39$), but the slope on the second lag is negative (-0.25, $t = -2.18$) and about the same magnitude. Thus, the prediction of next year's earnings growth from this year's return is reversed the following year. In the one-year forecast regression for 1951 to 2000, the only variable other than lagged returns with power to forecast earnings growth ($t = -2.64$) is the third lag of earnings growth. But the slope is negative, so it predicts that the strong earnings growth of recent years is soon to be reversed.

In the 1951 to 2000 regression to forecast earnings one year ahead, there is a hint ($t = -1.91$) that the low earnings–price ratio at the end of the period implies higher than average expected growth one year ahead. But the effect peters out quickly; the slope on the lagged earnings–price ratio in the regression to forecast earnings growth two years ahead is -1.02 standard errors from zero. The only variables with forecast power two years ahead are the second lag of the stock return and the third lag of earnings growth. But the slopes on these variables are negative, so again the 2000 prediction is that the strong earnings growth of recent years is soon to be reversed. And again, regressions (not shown) confirm that forecast power for 1951 to 2000 does not extend beyond two years. Thus, beyond two years, the best forecast of earnings growth is the historical average growth rate.

In sum, the behavior of dividends for 1951 to 2000 suggests that future growth is largely unpredictable, so the historical mean growth rate is a near optimal forecast of future growth. Earnings growth for 1951 to 2000 is somewhat predictable one and two years ahead, but the end-of-sample message is that the recent high growth rates are likely to revert quickly to the historical mean. It is also worth noting that the market survivor bias argument of Brown, Goetzmann, and Ross (1995) suggests that past average growth rates are, if anything, upward biased estimates of future growth. In short, we find no evidence to support a forecast of strong future dividend or earnings growth at the end of our sample period.

C. Do Expected Stock Returns Fall during the 1951 to 2000 Period?

The S&P dividend–price ratio, D_t/P_t, falls from 7.18 percent at the end of 1950 to a historically low 1.22 percent at the end of 2000 (Figure 1). The growth in the stock price, P_{2000}/P_{1950}, is thus 5.89 times the growth in dividends, D_{2000}/D_{1950}. The S&P earnings–price ratio, Y_t/P_t, falls from 13.39 percent at the end of 1950 to 3.46 percent at the end of 2000, so the percent capital gain of the last 50 years is 3.87 times the percent growth in earnings. (Interestingly, almost all of the excess capital gain occurs in the last 20 years; Figure 1 shows that the 1979 earnings–price ratio, 13.40 percent, is nearly identical to the 13.39 percent value of 1950.)

All valuation models say that D_t/P_t and E_t/P_t are driven by expected future returns (discount rates) and expectations about future dividend and earnings growth. Our evidence suggests that rational forecasts of long-term dividend and earnings growth rates are not unusually high in 2000. We conclude that the large spread of capital gains for 1951 to 2000 over dividend and earnings growth is largely due to a decline in the expected stock return.

Some of the decline in D_t/P_t and E_t/P_t during 1951 to 2000 is probably anticipated in 1950. The dividend–price ratio for 1950, 7.18 percent, is high (Figure 1). The average for 1872 to 2000 is 4.64 percent. If D_t/P_t is mean-reverting, the expectation in 1950 of the yield in 2000 is close to the unconditional mean, say 4.64 percent. The actual dividend–price ratio for 2000 is

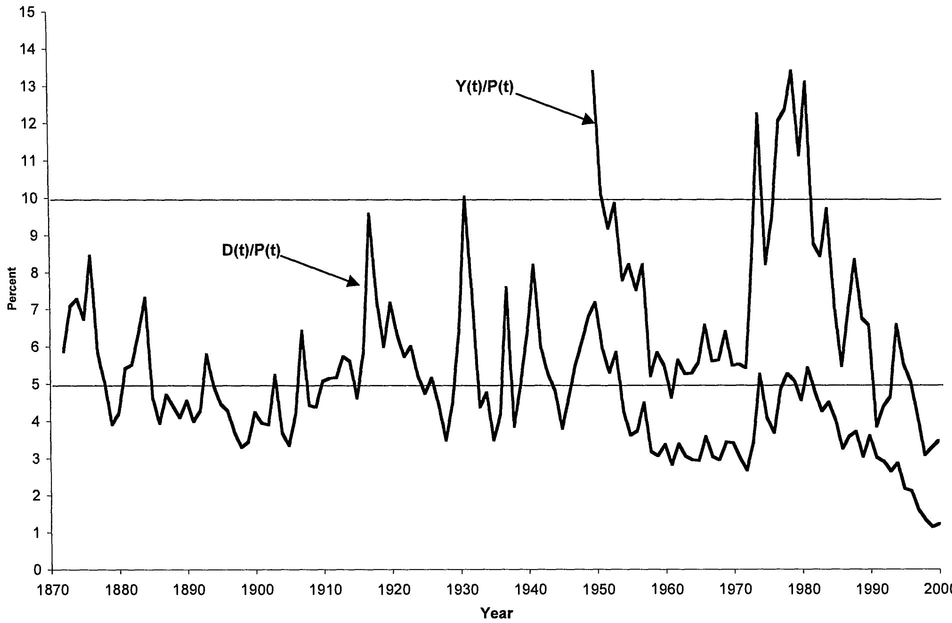


Figure 1. Dividend–price and earnings–price ratios.

1.22 percent. The 2000 stock price is thus 4.64/1.22 = 3.80 times what it would be if the dividend yield for 2000 hit the historical mean. Roughly speaking, this unexpected capital gain adds about 2.67 percent to the compound annual return for 1951 to 2000.

Similarly, part of the large difference between the 1951 to 2000 capital gain and the growth in earnings is probably anticipated in 1950. The 13.39 percent value of Y_t/P_t in 1950 is high relative to the mean for 1951 to 2000, 7.14 percent. If the earnings–price ratio is stationary, the expectation in 1950 of Y_t/P_t for 2000 is close to the unconditional mean, say 7.14 percent. The actual Y_t/P_t for 2000 is 3.46 percent. Thus, the 2000 stock price is 7.14/3.46 = 2.06 times what it would be if the ratio for 2000 hit the 7.14 percent average value for 1951 to 2000. Roughly speaking, this estimate of the unexpected capital gain adds about 1.45 percent to the compound annual return for the 50-year period.

In short, the percent capital gain for 1951 to 2000 is several times the growth of dividends or earnings. The result is historically low dividend–price and earnings–price ratios at the end of the period. Since the ratios are high in 1950, some of their subsequent decline is probably expected, but much of it is unexpected. Given the evidence that rational forecasts of long-term growth rates of dividends and earnings are not high in 2000, we conclude that the unexpected capital gains for 1951 to 2000 are largely due to a decline in the discount rate. In other words, the low end-of-sample price ratios imply low (rationally assessed, or true) expected future returns.

Like us, Campbell (1991), Cochrane (1994), and Campbell and Shiller (1998) find that, for recent periods, dividend and earnings growth are largely unpredictable, so variation in dividend–price and earnings–price ratios is largely due to the expected stock return. The samples in Campbell (1991) and Cochrane (1994) end in 1988 (before the strong subsequent returns that produce sharp declines in the price ratios), and they focus on explaining, in general terms, how variation in D_t/P_t splits between variation in the expected stock return and expected dividend growth. Campbell and Shiller (1998) focus on the low expected future returns implied by the low price ratios of recent years.

In contrast, we are more interested in what the decline in the price ratios says about past returns, specifically, that the average return for 1951 to 2000 is above the expected return. And this inference does not rest solely on the information in price ratios. We buttress it with two types of novel evidence. (a) The perspective from our long sample period that, although the average stock return for 1951 to 2000 is much higher than expected return estimates from fundamentals, the two approaches produce similar estimates for 1872 to 1950. (b) Evidence from Sharpe ratios, the book-to-market ratio, and the income return on investment, which also suggests that the average return for 1951 to 2000 is above the expected value.

III. Estimating the Expected Stock Return: Issues

There are two open questions about our estimates of the expected stock return. (a) In recent years the propensity of firms to pay dividends declines and stock repurchases surge. How do these changes in dividend policy affect our estimates of the expected return? (b) Under rather general conditions, the dividend and earnings growth models (2) and (3) provide estimates of the expected stock return. Are the estimates biased and does the bias depend on the return horizon? This section addresses these issues.

A. Repurchases and the Declining Incidence of Dividend Payers

Share repurchases surge after 1983 (Bagwell and Shoven (1989) and Dunsby (1995)), and, after 1978, the fraction of firms that do not pay dividends steadily increases (Fama and French (2001)). More generally, dividends are a policy variable, and changes in policy can raise problems for estimates of the expected stock return from the dividend growth model. There is no problem in the long-term, as long as dividend policies stabilize and the dividend–price ratio resumes its mean-reversion, though perhaps to a new mean. (An Appendix, available on request, provides an example involving repurchases.) But there can be problems during transition periods. For example, if the fraction of firms that do not pay dividends steadily increases, the market dividend–price ratio is probably nonstationary; it is likely to decline over time, and the dividend growth model is likely to underestimate the expected stock return.

Fortunately, the earnings growth model is not subject to the problems posed by drift in dividend policy. The earnings growth model provides an estimate of the expected stock return when the earnings–price ratio is stationary. And as discussed earlier, the model provides an estimate of the average expected return during the sample period when there are permanent shifts in the expected value of Y_t/P_t, as long as the ratio mean-reverts within regimes.

The earnings growth model is not, however, clearly superior to the dividend growth model. The standard deviation of annual earnings growth rates for 1951 to 2000 (13.79 percent, versus 5.09 percent for dividends) is similar to that of capital gains (16.77 percent), so much of the precision advantage of using fundamentals to estimate the expected stock return is lost. We see next that the dividend growth model has an advantage over the earnings growth model and the average stock return if the goal is to estimate the long-term expected growth of wealth.

B. The Investment Horizon

The return concept in discrete time asset pricing models is a one-period simple return, and our empirical work focuses on the one-year return. But many, if not most, investors are concerned with long-term returns, that is, terminal wealth over a long holding period. Do the advantages and disadvantages of different expected return estimates depend on the return horizon? This section addresses this question.

B.1. The Expected Annual Simple Return

There is downward bias in the estimates of the expected annual simple return from the dividend and earnings growth models—the result of a variance effect. The expected value of the dividend growth estimate of the expected return, for example, is the expected value of the dividend yield plus the expected value of the annual simple dividend growth rate. The expected annual simple return is the expected value of the dividend yield plus the expected annual simple rate of capital gain. If the dividend–price ratio is stationary, the compound rate of capital gain converges to the compound dividend growth rate as the sample period increases. But because the dividend growth rate is less volatile than the rate of capital gain, the expected simple dividend growth rate is less than the expected simple rate of capital gain.

The standard deviation of the annual simple rate of capital gain for 1951 to 2000 is 3.29 times the standard deviation of the annual dividend growth rate (Table I). The resulting downward bias of the average dividend growth rate as an estimate of the expected annual simple rate of capital gain is roughly 1.28 percent per year (half the difference between the variances of the two growth rates). Corrected for this bias, the dividend growth estimate of the equity premium in the simple returns of 1951 to 2000 rises from 2.55 to 3.83 percent (Table IV), which is still far below the estimate from the average return, 7.43 percent. Since the earnings growth rate and the annual rate of capital gain have similar standard deviations for 1951 to 2000,

Table IV

Estimates of the Real Equity Premium in Simple Annual and Long-term Returns: 1951 to 2000

The inflation rate for year t is $Inf_t = L_t/L_{t-1}$, where L_t is the price level at the end of year t. The real return for year t on six-month (three-month for the year 2000) commercial paper (rolled over at midyear) is F_t. The nominal value of the S&P index at the end of year t is p_t. Nominal S&P dividends and earnings for year t are d_t and y_t. Real rates of growth of dividends, earnings, and the stock price are $GD_t = (d_t/d_{t-1})*(L_{t-1}/L_t) - 1$, $GY_t = (y_t/y_{t-1})*(L_{t-1}/L_t) - 1$, and $GP_t = (p_t/p_{t-1})*(L_{t-1}/L_t) - 1$. The real dividend yield is $D_t/P_{t-1} = (d_t/p_{t-1})*(L_{t-1}/L_t)$. The dividend growth estimate of the real S&P return for t is $RD_t = D_t/P_{t-1} + GD_t$, the earnings growth estimate is $RY_t = D_t/P_{t-1} + GY_t$, and R_t is the realized real S&P return. The dividend and earnings growth estimates of the real equity premium for year t are $RXD_t = RD_t - F_t$ and $RXY_t = RY_t - F_t$, and $RX_t = R_t - F_t$ is the real equity premium from the realized real return. The average values of the equity premium estimates are $A(RXD_t)$, $A(RXY_t)$, and $A(RX_t)$. The first column of the table shows unadjusted estimates of the annual simple equity premium. The second column shows bias-adjusted estimates of the annual premium. The bias adjustment is one-half the difference between the variance of the annual rate of capital gain and the variance of either the dividend growth rate or the earnings growth rate. The third column shows bias-adjusted estimates of the expected equity premium relevant if one is interested in the long-term growth rate of wealth. The bias adjustment is one-half the difference between the variance of the annual dividend growth rate and the variance of either the growth rate of earnings or the rate of capital gain. The equity premiums are expressed as percents.

		Bias-adjusted	
	Unadjusted	Annual	Long-term
$A(RXD_t)$	2.55	3.83	2.55
$A(RXY_t)$	4.32	4.78	3.50
$A(RX_t)$	7.43	7.43	6.16

13.79 percent and 16.77 percent (Table I), the bias of the earnings growth estimate of the expected return is smaller (0.46 percent). Corrected for bias, the estimate of the equity premium for 1951 to 2000 from the earnings growth model rises from 4.32 to 4.78 percent (Table IV), which again is far below the 7.43 percent estimate from the average return.

B.2. Long-term Expected Wealth

The (unadjusted) estimate of the expected annual simple return from the dividend growth model is probably the best choice if we are concerned with the long-term expected wealth generated by the market portfolio. The annual dividend growth rates of 1951 to 2000 are essentially unpredictable. If the dividend growth rate is serially uncorrelated, the expected value of the compounded dividend growth rate is the compounded expected simple growth rate:

$$E\left[\prod_{t=1}^{T}(1 + GD_t)\right] = [1 + E(GD)]^T. \tag{5}$$

And if the dividend–price ratio is stationary, for long horizons the expected compounded dividend growth rate is the expected compounded rate of capital gain:

$$E\left[\prod_{t=1}^{T}(1+GD_t)\right]=E\left[\prod_{t=1}^{T}(1+GP_t)\right]. \tag{6}$$

Thus, when the horizon T is long, compounding the true expected annual simple return from the dividend growth model produces an unbiased estimate of the expected long-term return:

$$[1+E(RD)]^T=E\left[\prod_{t=1}^{T}(1+R_t)\right]. \tag{7}$$

In contrast, if the dividend growth rate is unpredictable and the dividend–price ratio is stationary, part of the higher volatility of annual rates of capital gain is transitory, the result of a mean-reverting expected annual return (Cochrane (1994)). Thus, compounding even the true unconditional expected annual simple return, $E(R)$, yields an upward biased measure of the expected compounded return:

$$[1+E(R)]^T>E\left[\prod_{t=1}^{T}(1+R_t)\right]. \tag{8}$$

There is a similar problem in using the average (simple) earnings growth rate to estimate long-term expected wealth. The regressions in Table III suggest that the predictability of earnings growth for 1951 to 2000 is due to transitory variation in earnings. As a result, annual earnings growth is 2.71 times more volatile than dividend growth (Table I). The compound growth rate of earnings for 1951 to 2000, 1.89 percent, is 2.05 times the compound dividend growth rate, 0.92 percent. But because earnings are more volatile, the average simple growth rate of earnings, 2.82 percent, is 2.69 times the average simple growth rate of dividends, 1.05 percent. As a result, the average simple growth rate of earnings produces an upward biased estimate of the compound rate of growth of long-term expected wealth.

We can correct the bias by subtracting half the difference between the variance of earnings growth and the variance of dividend growth (0.82 percent) from the average earnings growth rate. The estimate of the expected rate of capital gain provided by this adjusted average growth rate of earnings is 2.00 percent per year. Using this adjusted average growth rate of earnings, the earnings growth estimate of the expected real stock return for 1951 to 2000 falls from 6.51 to 5.69 percent. The estimate of the equity premium falls from 4.32 to 3.50 percent (Table IV), which is closer to the 2.55 percent obtained when the average dividend growth rate is used to estimate the expected rate of capital gain. Similarly, adjusting for the effects of transitory return volatility causes the estimate of the equity premium from realized stock returns to fall from 7.43 to 6.16 percent, which is still far above the bias-adjusted estimate of the earnings growth model (3.50 percent) and the estimate from the dividend growth model (2.55 percent).

Finally, we only have estimates of the expected growth rates of dividends and earnings and the expected rate of capital gain. Compounding estimates rather than true expected values adds upward bias to measures of expected long-term wealth (Blume (1974)). The bias increases with the imprecision of the estimates. This is another reason to favor the more precise estimate of the expected stock return from the dividend growth model over the earnings growth estimate or the estimate from the average stock return.

IV. Conclusions

There is a burgeoning literature on the equity premium. Our main additions are on two fronts. (a) A long (1872 to 2000) perspective on the competing estimates of the unconditional expected stock return from fundamentals (the dividend and earnings growth models) and the average stock return. (b) Evidence (estimates of precision, Sharpe ratios, and the behavior of the book-to-market ratio and the income return on investment) that allows us to choose between the expected return estimates from the two approaches.

Specifically, the dividend growth model and the realized average return produce similar real equity premium estimates for 1872 to 1950, 4.17 percent and 4.40 percent. For the half-century from 1951 to 2000, however, the equity premium estimates from the dividend and earnings growth models, 2.55 percent and 4.32 percent, are far below the estimate from the average return, 7.43 percent.

We argue that the dividend and earnings growth estimates of the equity premium for 1951 to 2000 are closer to the true expected value. This conclusion is based on three results.

(a) The estimates from fundamentals, especially the estimate from the dividend growth model, are more precise; they have lower standard errors than the estimate from the average return.

(b) The appealing message from the dividend and earnings growth models is that aggregate risk aversion (as measured by the Sharpe ratio for the equity premium) is on average roughly similar for the 1872 to 1949 and 1950 to 1999 periods. In contrast, the Sharpe ratio for the equity premium from the average return just about doubles from the 1872 to 1950 period to the 1951 to 2000 period.

(c) Most important, the average stock return for 1951 to 2000 is much greater than the average income return on book equity. Taken at face value, this says that investment during the period is on average unprofitable (its expected return is less than the cost of capital). In contrast, the lower estimates of the expected stock return from the dividend and earnings growth models are less than the income return on investment, so the message is that investment is on average profitable. This is more consistent with book-to-market ratios that are rather consistently less than one during the period.

If the average stock return for 1951 to 2000 exceeds the expected return, stocks experience unexpected capital gains. What is the source of the gains? Growth rates of dividends and earnings are largely unpredictable, so there is no basis for extrapolating unusually high long-term future growth. This leaves a decline in the expected stock return as the prime source of the unexpected capital gain. In other words, the high return for 1951 to 2000 seems to be the result of low expected future returns.

Many papers suggest that the decline in the expected stock return is in part permanent, the result of (a) wider equity market participation by individuals and institutions, and (b) lower costs of obtaining diversified equity portfolios from mutual funds (Diamond (1999), Heaton and Lucas (1999), and Siegel (1999)). But there is also evidence that the expected stock return is slowly mean reverting (Fama and French (1989) and Cochrane (1994)). Moreover, there are two schools of thought on how to explain the variation in expected returns. Some attribute it to rational variation in response to macroeconomic factors (Fama and French (1989), Blanchard (1993), and Cochrane (1994)), while others judge that irrational swings in investor sentiment are the prime moving force (e.g., Shiller (1989)). Whatever the story for variation in the expected return, and whether it is temporary or partly permanent, the message from the low end-of-sample dividend–price and earnings–price ratios is that we face a period of low (true) expected returns.

Our main concern, however, is the unconditional expected stock return, not the end-of-sample conditional expected value. Here there are some nuances. If we are interested in the unconditional expected annual simple return, the estimates for 1951 to 2000 from fundamentals are downward biased. The bias is rather large when the average growth rate of dividends is used to estimate the expected rate of capital gain, but it is small for the average growth rate of earnings. On the other hand, if we are interested in the long-term expected growth of wealth, the dividend growth model is probably best, and the average stock return and the earnings growth estimate of the expected return are upward biased. But our bottom line inference does not depend on whether one is interested in the expected annual simple return or long-term expected wealth. In either case, the bias-adjusted expected return estimates for 1951 to 2000 from fundamentals are a lot (more than 2.6 percent per year) lower than bias-adjusted estimates from realized returns. (See Table IV.) Based on this and other evidence, our main message is that the unconditional expected equity premium of the last 50 years is probably far below the realized premium.

REFERENCES

Bagwell, Laurie S., and John B. Shoven, 1989, Cash distributions to shareholders, *Journal of Economic Perspectives* 3, 129–149.

Blanchard, Olivier J., 1993, Movements in the equity premium, *Brooking Papers on Economic Activity* 2, 75–138.

Blume, Marshall, 1974, Unbiased estimators of long-run expected rates of return, *Journal of the American Statistical Association* 69, 634–638.

Brown, Stephen J., William N. Goetzmann, and Stephen Ross, 1995, Survival, *Journal of Finance* 50, 853–873.

Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.

Campbell, John Y., and Robert J. Shiller, 1989, The dividend–price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Campbell, John Y., and Robert J. Shiller, 1998, Valuation ratios and the long-run stock market outlook, *Journal of Portfolio Management* 24, 11–26.

Claus, James, and Jacob Thomas, 2001, Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets, *Journal of Finance* 56, 1629-1666.

Cochrane, John, 1991, Volatility tests and efficient markets: A review essay, *Journal of Monetary Economics* 27, 463–487.

Cochrane, John H., 1994, Permanent and transitory components of GNP and stock prices, *Quarterly Journal of Economics* 109, 241–265.

Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns, *Journal of Finance* 55, 389–406.

Diamond, Peter A., 1999, What stock market returns to expect for the future? Center of Retirement Research at Boston College, September, Number 2.

Dunsby, Adam, 1995, *Share Repurchases, Dividends, and Corporate Distribution Policy* (Ph.D. thesis, The Wharton School, University of Pennsylvania).

Fama, Eugene F., and Kenneth R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.

Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene F., and Kenneth R. French, 2001, Disappearing dividends: Changing firm characteristics or lower propensity to pay, *Journal of Financial Economics* 60, 3–43.

Gebhardt, William R., Charles M. C. Lee, and Bhaskaram Swaminathan, 2001, Toward an implied cost of capital, *Journal of Accounting Research* 39, June, 135–176.

Gordon, Myron, 1962, *The Investment Financing and Valuation of the Corporation* (Irwin, Homewood, IL).

Heaton, John, and Deborah Lucas, 1999, Stock prices and fundamentals, in Ben Bernanke and Julio Rotemberg, eds.: *Macroeconomics Annual 1999* (National Bureau of Economic Research, MIT Press, Cambridge, MA).

Keane, Michael P., and David E. Runkle, 1998, Are financial analysts' forecasts of corporate profits rational? *Journal of Political Economy* 106, 768–805.

Lamont, Owen, 1998, Earnings and expected returns, *Journal of Finance* 53, 1563–1587.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849.

Mehra, Rajnish, and Edward Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.

Shiller, Robert, 1989, *Market Volatility* (MIT Press, Cambridge, MA).

Siegel, Jeremy J., 1999, The shrinking equity premium, *Journal of Portfolio Management* 26, 10–17.

Vuolteenaho, Tuomo, 2000, Understanding the aggregate book-to-market ratio, Manuscript, University of Chicago.

LINKED CITATIONS

You have printed the following article:

The Equity Premium
Eugene F. Fama; Kenneth R. French
The Journal of Finance, Vol. 57, No. 2. (Apr., 2002), pp. 637-659.
Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28200204%2957%3A2%3C637%3ATEP%3E2.0.CO%3B2-%23

[Footnotes]

[1] **Consumption, Aggregate Wealth, and Expected Stock Returns**
Martin Lettau; Sydney Ludvigson
The Journal of Finance, Vol. 56, No. 3. (Jun., 2001), pp. 815-849.
Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28200106%2956%3A3%3C815%3ACAWAES%3E2.0.CO%3B2-N

References

Cash Distributions to Shareholders
Laurie Simon Bagwell; John B. Shoven
The Journal of Economic Perspectives, Vol. 3, No. 3. (Summer, 1989), pp. 129-140.
Stable URL:
http://links.jstor.org/sici?sici=0895-3309%28198922%293%3A3%3C129%3ACDTS%3E2.0.CO%3B2-9

Unbiased Estimators of Long-Run Expected Rates of Return
Marshall E. Blume
Journal of the American Statistical Association, Vol. 69, No. 347. (Sep., 1974), pp. 634-638.
Stable URL:
http://links.jstor.org/sici?sici=0162-1459%28197409%2969%3A347%3C634%3AUEOLER%3E2.0.CO%3B2-V

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

Survival
Stephen J. Brown; William N. Goetzmann; Stephen A. Ross
The Journal of Finance, Vol. 50, No. 3, Papers and Proceedings Fifty-Fifth Annual Meeting, American Finance, Association, Washington, D.C., January 6-8, 1995. (Jul., 1995), pp. 853-873.
Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28199507%2950%3A3%3C853%3AS%3E2.0.CO%3B2-0

A Variance Decomposition for Stock Returns
John Y. Campbell
The Economic Journal, Vol. 101, No. 405. (Mar., 1991), pp. 157-179.
Stable URL:
http://links.jstor.org/sici?sici=0013-0133%28199103%29101%3A405%3C157%3AAVDFSR%3E2.0.CO%3B2-X

The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors
John Y. Campbell; Robert J. Shiller
The Review of Financial Studies, Vol. 1, No. 3. (Autumn, 1988), pp. 195-228.
Stable URL:
http://links.jstor.org/sici?sici=0893-9454%28198823%291%3A3%3C195%3ATDRAEO%3E2.0.CO%3B2-O

Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets
James Claus; Jacob Thomas
The Journal of Finance, Vol. 56, No. 5. (Oct., 2001), pp. 1629-1666.
Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28200110%2956%3A5%3C1629%3AEPALAT%3E2.0.CO%3B2-6

Permanent and Transitory Components of GNP and Stock Prices
John H. Cochrane
The Quarterly Journal of Economics, Vol. 109, No. 1. (Feb., 1994), pp. 241-265.
Stable URL:
http://links.jstor.org/sici?sici=0033-5533%28199402%29109%3A1%3C241%3APATCOG%3E2.0.CO%3B2-K

Characteristics, Covariances, and Average Returns: 1929 to 1997
James L. Davis; Eugene F. Fama; Kenneth R. French
The Journal of Finance, Vol. 55, No. 1. (Feb., 2000), pp. 389-406.
Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28200002%2955%3A1%3C389%3ACCAAR1%3E2.0.CO%3B2-V

NOTE: *The reference numbering from the original has been maintained in this citation list.*

LINKED CITATIONS

Toward an Implied Cost of Capital
William R. Gebhardt; Charles M. C. Lee; Bhaskaran Swaminathan
Journal of Accounting Research, Vol. 39, No. 1. (Jun., 2001), pp. 135-176.
Stable URL:
http://links.jstor.org/sici?sici=0021-8456%28200106%2939%3A1%3C135%3ATAICOC%3E2.0.CO%3B2-4

Are Financial Analysts' Forecasts of Corporate Profits Rational?
Michael P. Keane; David E. Runkle
The Journal of Political Economy, Vol. 106, No. 4. (Aug., 1998), pp. 768-805.
Stable URL:
http://links.jstor.org/sici?sici=0022-3808%28199808%29106%3A4%3C768%3AAFAFOC%3E2.0.CO%3B2-%23

Earnings and Expected Returns
Owen Lamont
The Journal of Finance, Vol. 53, No. 5. (Oct., 1998), pp. 1563-1587.
Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28199810%2953%3A5%3C1563%3AEAER%3E2.0.CO%3B2-G

Consumption, Aggregate Wealth, and Expected Stock Returns
Martin Lettau; Sydney Ludvigson
The Journal of Finance, Vol. 56, No. 3. (Jun., 2001), pp. 815-849.
Stable URL:
http://links.jstor.org/sici?sici=0022-1082%28200106%2956%3A3%3C815%3ACAWAES%3E2.0.CO%3B2-N

NOTE: *The reference numbering from the original has been maintained in this citation list.*

Markets & Finance June 10, 2010, 5:00PM EST

Bloomberg Businessweek

For Analysts, Things Are Always Looking Up

They're raising earnings estimates for U.S. companies at a record pace

By Roben Farzad

For years, the rap on Wall Street securities analysts was that they were shills, reflexively producing upbeat research on companies they cover to help their employers win investment banking business. The dynamic was well understood: Let my bank take your company public, or advise it on this acquisition, and—wink, wink—I will recommend your stock through thick or thin. After the Internet bubble burst, that was supposed to change. In April 2003 the Securities & Exchange Commission reached a settlement with 10 Wall Street firms in which they agreed, among other things, to separate research from investment banking.

Seven years on, Wall Street analysts remain a decidedly optimistic lot. Some economists look at the global economy and see troubles—the European debt crisis, persistently high unemployment worldwide, and housing woes in the U.S. Stock analysts as a group seem unfazed. Projected 2010 profit growth for companies in the Standard & Poor's 500-stock index has climbed seven percentage points this quarter, to 34 percent, data compiled by Bloomberg show. According to Sanford C. Bernstein (AB), that's the fastest pace since 1980, when the Dow Jones industrial average was quoted in the hundreds and Nancy Reagan was getting ready to order new window treatments for the Oval Office.

Among the companies analysts expect to excel: Intel (INTL) is projected to post an increase in net income of 142 percent this year. Caterpillar, a multinational that gets much of its revenue abroad, is expected to boost its net income by 47 percent this year. Analysts have also hiked their S&P 500 profit estimate for 2011 to $95.53 a share, up from $92.45 at the beginning of January, according to Bloomberg data. That would be a record, surpassing the previous high reached in 2007.

With such prospects, it's not surprising that more than half of S&P 500-listed stocks boast overall buy ratings. It is telling that the proportion has essentially held constant at both the market's October 2007 high and March 2009 low, bookends of a period that saw stocks fall by more than half. If the analysts are correct, the market would appear to be attractively priced right now. Using the $95.53 per share figure, the price-to-earnings ratio of the S&P 500 is a modest 11 as of June 9. If, however, analysts end up being too high by, say, 20 percent, the P/E would jump to almost 14.

If history is any guide, chances are good that the analysts are wrong. According to a recent McKinsey report by Marc Goedhart, Rishi Raj, and Abhishek Saxena, "Analysts have been persistently over-optimistic for 25 years," a stretch that saw them peg earnings growth at 10 percent to 12 percent a year when the actual number was ultimately 6 percent. "On average," the researchers note, "analysts' forecasts have been almost 100 percent too high," even after regulations were enacted to weed out conflicts and improve the rigor of their calculations. As the chart below shows, in most years analysts have been forced to lower their estimates after it became apparent they had set them too high.

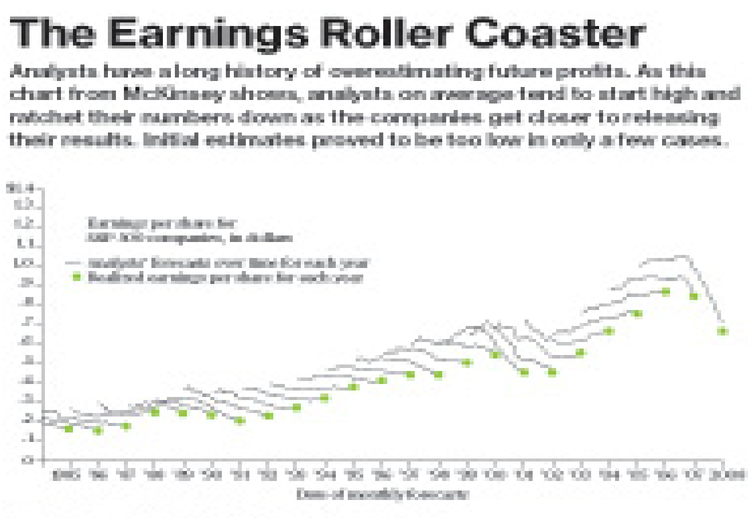
While a few analysts, like Meredith Whitney, have made their names on bearish calls, most are chronically bullish. Part of the problem is that despite all the reforms they remain too aligned with the companies they cover. "Analysts still need to get the bulk of their information from companies, which have an incentive to be over-optimistic," says Stephen Bainbridge, a professor at UCLA Law School who specializes in the securities industry. "Meanwhile, analysts don't want to threaten that ongoing access by being too negative." Bainbridge says that with the era of the overpaid, superstar analyst long over, today's job description calls for resisting the urge to be an iconoclast. "It's a matter of herd behavior," he says.

So what's a more plausible estimate of companies' earning power? Looking at factors including the strengthening dollar, which hurts exports, and higher corporate borrowing costs, David Rosenberg, chief economist at Toronto-based investment shop Gluskin Sheff + Associates, says "disappointment looms." Bernstein's Adam Parker says every 10 percent drop in the value of the euro knocks U.S. corporate earnings down by 2.5 percent to 3 percent. He sees the S&P 500 earning $86 a share next year.

As realities hit home, "It's only natural that analysts will have to revise down their views," says Todd Salamone, senior vice-president at Schaeffer's Investment Research. The market may be making its own downward adjustment, as the S&P 500 has already fallen 14 percent from its high in April. If precedent holds, analysts are bound to curb their enthusiasm belatedly, telling us next year what we really needed to know this year.

The bottom line: *Despite reforms intended to improve Wall Street research, stock analysts seem to be promoting an overly rosy view of profit prospects.*

Bloomberg Businessweek Senior Writer Farzad covers Wall Street and international finance.



Markets & Finance June 10, 2010, 5:00PM EST

For Analysts, Things Are Always Looking Up

They're raising earnings estimates for U.S. companies at a record pace

By Roben Farzad

GETTY IMAGES

For years, the rap on Wall Street securities analysts was that they were shills, reflexively producing upbeat research on companies they cover to help their employers win investment banking business. The dynamic was well understood: Let my bank take your company public, or advise it on this acquisition, and—wink, wink—I will recommend your stock through thick or thin. After the Internet bubble burst, that was supposed to change. In April 2003 the Securities & Exchange Commission reached a settlement with 10 Wall Street firms in which they agreed, among other things, to separate research from investment banking.

Seven years on, Wall Street analysts remain a decidedly optimistic lot. Some economists look at the global economy and see troubles—the European debt crisis, persistently high unemployment worldwide, and housing woes in the U.S. Stock analysts as a group seem unfazed. Projected 2010 profit growth for companies in the Standard & Poor's 500-stock index has climbed seven percentage points this quarter, to

34 percent, data compiled by Bloomberg show. According to Sanford C. Bernstein (AB), that's the fastest pace since 1980, when the Dow Jones industrial average was quoted in the hundreds and Nancy Reagan was getting ready to order new window treatments for the Oval Office.

Among the companies analysts expect to excel: Intel (INTL) is projected to post an increase in net income of 142 percent this year. Caterpillar, a multinational that gets much of its revenue abroad, is expected to boost its net income by 47 percent this year. Analysts have also hiked their S&P 500 profit estimate for 2011 to $95.53 a share, up from $92.45 at the beginning of January, according to Bloomberg data. That would be a record, surpassing the previous high reached in 2007.

With such prospects, it's not surprising that more than half of S&P 500-listed stocks boast overall buy ratings. It is telling that the proportion has essentially held constant at both the market's October 2007 high and March 2009 low, bookends of a period that saw stocks fall by more than half. If the analysts are correct, the market would appear to be attractively priced right now. Using the $95.53 per share figure, the price-to-earnings ratio of the S&P 500 is a modest 11 as of June 9. If, however, analysts end up being too high by, say, 20 percent, the P/E would jump to almost 14.

If history is any guide, chances are good that the analysts are wrong. According to a recent McKinsey report by Marc Goedhart, Rishi Raj, and Abhishek Saxena, "Analysts have been persistently over-optimistic for 25 years," a stretch that saw them peg earnings growth at 10 percent to 12 percent a year when the actual number was ultimately 6 percent. "On average," the researchers note, "analysts' forecasts have been almost 100 percent too high," even after regulations were enacted to weed out conflicts and improve the rigor of their calculations. As the chart below shows, in most years analysts have been forced to lower their estimates after it became apparent they had set them too high.

While a few analysts, like Meredith Whitney, have made their names on bearish calls, most are chronically bullish. Part of the problem is that despite all the reforms they remain too aligned with the companies they cover. "Analysts still need to get the bulk of their information from companies, which have an incentive to be over-optimistic," says Stephen Bainbridge, a professor at UCLA Law School who specializes in the securities industry. "Meanwhile, analysts don't want to threaten that ongoing access by being too negative." Bainbridge says that with the era of the overpaid, superstar analyst long over, today's job description calls for resisting the urge to be an iconoclast. "It's a matter of herd behavior," he says.

So what's a more plausible estimate of companies' earning power? Looking at factors including the strengthening dollar, which hurts exports, and higher corporate borrowing costs, David Rosenberg, chief economist at Toronto-based investment shop Gluskin Sheff + Associates, says "disappointment looms." Bernstein's Adam Parker says every 10 percent drop in the value of the euro knocks U.S. corporate earnings down by 2.5 percent to 3 percent. He sees the S&P 500 earning $86 a share next year.

As realities hit home, "It's only natural that analysts will have to revise down their views," says Todd Salamone, senior vice-president at Schaeffer's Investment Research. The market may be making its own downward adjustment, as the S&P 500 has already fallen 14 percent from its high in April. If precedent holds, analysts are bound to curb their enthusiasm belatedly, telling us next year what we really needed to know this year.

The bottom line: *Despite reforms intended to improve Wall Street research, stock analysts seem to be promoting an overly rosy view of profit prospects.*

Bloomberg Businessweek Senior Writer Farzad covers Wall Street and international finance.

Bernanke: Low interest-rate-policies benefit trade

By MARTIN CRUTSINGER | Associated Press – Mon, Mar 25, 2013 4:20 PM EDT

WASHINGTON (AP) — Chairman Ben Bernanke said Monday that the Federal Reserve's low-interest-rate policies are helping to boost growth around the world, rejecting criticism that they could lead to a global currency war.

In a speech at the London School of Economics, Bernanke staunchly defended the Fed's policies and similar stimulus efforts pursued by other central banks since the 2008 financial crisis.

Last week, the Fed stood by its policies to keep borrowing costs at record lows, saying the U.S. economy still required the support to help lower high unemployment.

Critics have argued that the low-interest-rate policies could lower a country's currency value and make its products more competitive on global markets.

Some have blamed such policies for making the Great Depression worse during the 1930s. Countries devalued their currencies and raised tariffs, which made foreign-made goods more expensive and stunted trade. They became known as "beggar-thy-neighbor" policies.

Bernanke argued that the situation is different today because the low-interest rate policies have the primary aim of boosting domestic growth, not trying to lower the value of a nation's currency.

"Because stronger growth in each economy confers beneficial spillovers to trading policies, these policies are not 'beggar-thy-neighbor' but rather ... 'enrich-thy-neighbor' actions," Bernanke said.

The current efforts should support stronger trade flows, Bernanke said. By boosting growth in major economies, consumers can buy more imported goods from developing countries.

In addition to concerns about currency wars, critics have also said that the policies adopted by the Fed and other central banks could increase the risk of inflation and destabilize financial markets.

Panelist Axel Weber, a former president of Germany's central bank and now chairman of the board of Swiss bank UBS, spoke to those concerns. He said central banks will be pressed to develop policies that wind down their stimulus without triggering "even bigger problems."

During a question-and-answer session after the speech, Bernanke neither addressed Weber's concerns directly nor offered any hints about the direction of U.S. interest rates.

But former U.S. Treasury Secretary Larry Summers, another member of the panel and a supporter of the Fed's policies, said the biggest threats to the economy now are high unemployment.

"The risks of stagnation are an inherently greater concern than inflation," Summers said.

The Fed's policies are aimed at lowering unemployment, which has fallen to 7.7 percent but is still above healthy levels.

After its two-day meeting last week, the Fed said it would stick with its plan to keep short-term interest rates at record lows at least until unemployment falls to 6.5 percent.

Bernanke told reporters that the Fed saw the 6.5 percent unemployment level as a threshold and not a "trigger," for a possible rate increase.

The Fed also said it would keep buying $85 billion a month in bonds to keep long-term borrowing costs. Lower rates encourage more borrowing and spending, which leads to faster growth and lower unemployment.

Bernanke told reporters at a news conference that the Fed might vary the size of its monthly purchases depending on whether the job market is improving and by how much.

In its policy statement, the Fed noted that the U.S. job market has improved, consumer spending and business investment have increased and the housing market has strengthened. But in an updated economic forecast also released last week, the Fed said it still did not expect unemployment to reach 6.5 percent until 2015.

The Fed's economic projections showed that 13 Fed officials still think the first Fed rate hike will not occur until 2015. That was the same number as in December. One Fed official thinks the first boost in the short-term lending rate won't occur until 2016.

___

Discount Rate (Risk-Free Rate and Market Risk Premium) used for 41 countries in 2015: a survey

Pablo Fernandez. Professor of Finance. IESE Business School, fernandezpa@iese.edu
Alberto Ortiz. Research Assistant. IESE Business School. AOrtiz@iese.edu
Isabel F. Acín. Independent researcher. University of Navarra. ifernandez.28@alumni.unav.es

ABSTRACT

This paper contains the statistics of a survey about the Risk-Free Rate (RF) and of the Market Risk Premium (MRP) used in 2015 for **41 countries**. We got answers for 68 countries, but we only report the results for 41 countries with more than 25 answers.

The average (RF) used in 2015 was smaller than the one used in 2013 in 26 countries (in 11 of them the difference was more than 1%). In 8 countries the average (RF) used in 2015 was more than a 1% higher than the one used in 2013 (see figure 3).

The change between 2013 and 2015 of the average Market risk premium used was higher than 1% for 13 countries (see figure 4).

Most of the respondents use for US, Europe and UK a Risk-Free Rate (RF) higher than the yield of the 10-year Government bonds.

JEL Classification: G12, G31, M21

Keywords: equity premium; required equity premium; expected equity premium; risk-free rate; heterogeneous expectations

April 23, 2015

xPpLnmlsjmO

1. Market Risk Premium (MRP), Risk Free Rate (RF) and Km [RF + MRP)] used in 2015 in 41 countries

We sent a short email (see exhibit 1) on the period March 15- April 10, 2015 to about 22,500 email addresses of finance and economic professors, analysts and managers of companies obtained from previous correspondence, papers and webs of companies and universities. We asked about the Risk Free Rate and the Market Risk Premium (**MRP**) **used** *"to calculate the required return to equity in different countries"*.

By April 22, 2015, we had received 2,396 emails. 216 persons answered that they do not use MRP for different reasons (see table 1). The remaining 2,758 emails had specific Risk Free Rates and MRPs used in 2015 for one or more countries.[1] We would like to sincerely thank everyone who took the time to answer us.

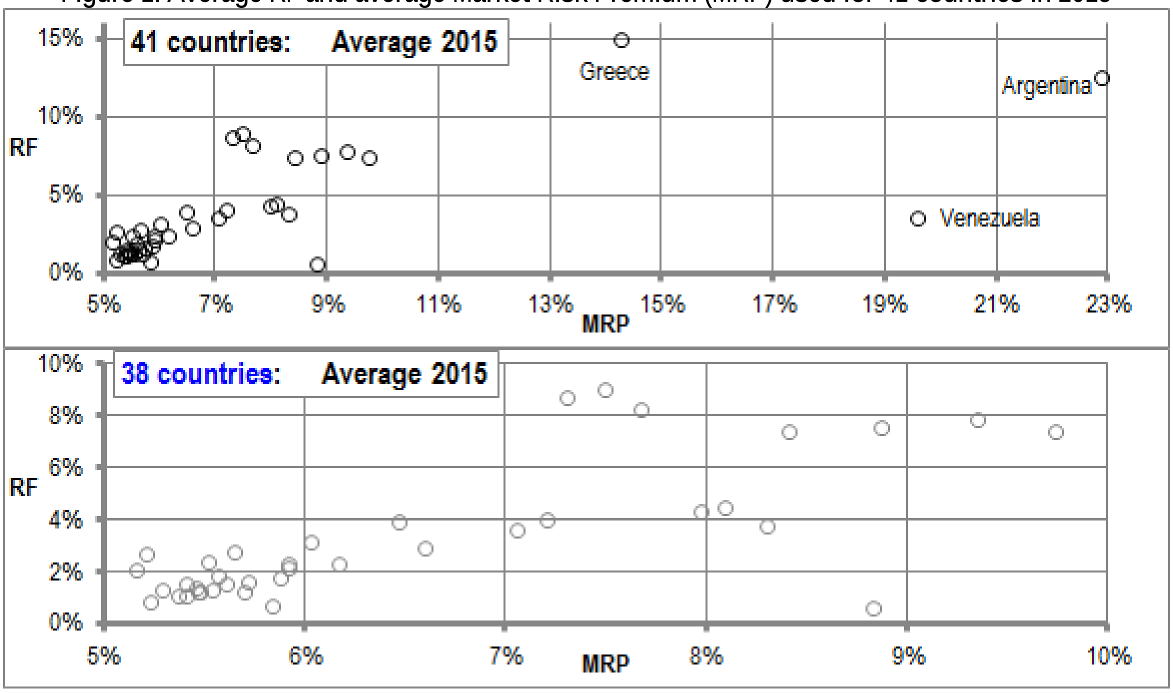
Table 1. MRP and RF used in 2015: 2,396 emails with 5,056 answers

	Total
Answers reported (MRP figures)	4,573
Outliers	48
Answers for 27 countries with less than 25 answers	177
Only MRP or RF (not both)	42
Answers that do not provide figures	216
Total	5,056

Table 2 contains the statistics of the **MRP** used in 2015 **for 41 countries**. We got answers for 68 countries, but we only report the results for 41 countries with more than 25 answers. **Table 3** contains the statistics of the Risk-Free Rate (**RF**) used in 2015 in the 41 countries and **Table 4** contains the statistics of **Km** (required return to equity: Km = Risk-Free Rate + MRP).

Figures 1, 2 and 3 are graphic representations of the MRPs reported in table 2.

Figure 1. Average RF and average Market Risk Premium (MRP) used for 41 countries in 2015



[1] We considered 48 of them as outliers because they provided a very small MRP (for example, -2% and 0% for the USA) or a very high MRP (for example, 30% for the USA).

Table 2. Market Risk Premium (MRP) used for 41 countries in 2015

MRP	Number of answers	average	Median	St. Dev.	max	min	Av-Median
USA	1983	5,5%	5,3%	1,4%	15,0%	2,0%	0,2%
Spain	443	5,9%	5,5%	1,6%	12,0%	3,0%	0,4%
Germany	252	5,3%	5,1%	1,5%	11,3%	2,0%	0,2%
France	122	5,6%	5,5%	1,4%	10,0%	2,0%	0,1%
United Kingdom	101	5,2%	5,0%	1,7%	10,5%	1,3%	0,2%
Italy	83	5,4%	5,2%	1,5%	10,0%	2,0%	0,2%
Canada	81	5,9%	6,0%	1,3%	12,0%	4,0%	-0,1%
Portugal	72	5,7%	5,5%	1,5%	10,0%	3,0%	0,2%
Switzerland	71	5,4%	5,0%	1,2%	10,0%	3,0%	0,4%
Belgium	70	5,5%	5,4%	1,3%	10,0%	1,0%	0,1%
Sweden	68	5,4%	5,1%	1,3%	10,0%	3,0%	0,3%
Denmark	65	5,5%	5,5%	1,2%	10,0%	3,0%	0,0%
Finland	64	5,7%	5,8%	1,1%	10,0%	3,0%	0,0%
Japan	61	5,8%	6,0%	2,0%	15,0%	2,0%	-0,2%
Norway	61	5,5%	5,2%	1,2%	10,0%	3,0%	0,2%
Brazil	59	7,5%	7,0%	2,1%	13,0%	4,1%	0,5%
Ireland	59	5,5%	5,2%	1,3%	10,0%	3,0%	0,3%
China	55	8,1%	7,0%	5,6%	40,2%	2,8%	1,1%
Mexico	55	8,0%	8,0%	1,5%	12,0%	4,0%	0,0%
Russia	48	9,7%	10,0%	2,9%	14,0%	3,0%	-0,3%
India	43	8,4%	8,3%	2,5%	14,0%	5,0%	0,1%
South Africa	43	7,7%	7,3%	2,3%	13,0%	4,5%	0,4%
Australia	40	6,0%	5,1%	4,0%	19,0%	1,5%	1,0%
Chile	37	6,5%	6,5%	0,9%	8,4%	5,0%	0,0%
Uruguay	37	7,1%	7,0%	0,9%	8,8%	5,5%	0,1%
Poland	36	5,2%	5,0%	1,0%	7,5%	3,5%	0,2%
Peru	35	7,2%	7,4%	1,2%	8,9%	3,6%	-0,2%
Czech Republic	34	5,6%	5,4%	0,7%	7,2%	4,4%	0,2%
Indonesia	34	8,9%	9,0%	1,2%	12,5%	6,0%	-0,1%
Israel	34	5,2%	5,0%	1,1%	7,0%	3,2%	0,2%
Korea (South)	33	6,2%	6,0%	1,5%	10,0%	4,0%	0,2%
Netherlands	33	5,9%	6,0%	0,6%	6,8%	4,0%	-0,1%
New Zealand	31	6,6%	6,0%	1,3%	10,0%	5,5%	0,6%
Thailand	31	7,3%	7,5%	0,9%	8,2%	4,0%	0,0%
Turkey	31	9,3%	9,1%	2,5%	16,0%	5,7%	0,3%
Austria	29	5,7%	5,6%	0,3%	6,5%	5,2%	0,0%
Greece	29	14,3%	15,0%	5,8%	22,8%	0,3%	-0,7%
Colombia	28	8,3%	8,0%	1,4%	12,0%	6,0%	0,3%
Hungary	28	8,8%	8,9%	0,8%	10,1%	6,0%	-0,1%
Venezuela	28	19,6%	19,3%	3,7%	32,4%	14,3%	0,3%
Argentina	26	22,9%	20,1%	12,3%	46,2%	5,0%	2,7%

Table 3. Risk Free Rate (RF) used for 41 countries in 2015

RF	Number of answers	average	Median	St. Dev.	max	min	Av-Median
USA	1983	2,4%	2,3%	1,1%	8,0%	0,0%	0,1%
Spain	443	2,2%	2,0%	1,2%	7,0%	0,0%	0,2%
Germany	252	1,3%	1,1%	0,8%	5,1%	-0,2%	0,2%
France	122	1,5%	1,3%	1,0%	5,1%	0,0%	0,3%
United Kingdom	101	2,1%	2,1%	0,8%	6,0%	0,4%	0,0%
Italy	83	1,5%	1,3%	1,1%	5,0%	0,0%	0,2%
Canada	81	2,3%	2,0%	1,0%	6,0%	0,8%	0,3%
Portugal	72	1,6%	1,5%	0,9%	5,0%	0,0%	0,1%
Switzerland	71	1,1%	1,0%	0,7%	4,0%	0,0%	0,1%
Belgium	70	1,3%	1,0%	0,9%	5,0%	0,0%	0,3%
Sweden	68	1,1%	1,0%	0,8%	4,0%	0,0%	0,1%
Denmark	65	1,3%	1,0%	1,0%	5,0%	-0,1%	0,3%
Finland	64	1,2%	1,0%	0,9%	5,0%	0,0%	0,2%
Japan	61	0,7%	0,5%	1,0%	7,0%	-0,7%	0,2%
Norway	61	1,4%	1,0%	1,1%	5,0%	-0,2%	0,4%
Brazil	59	9,0%	9,0%	2,8%	13,1%	3,2%	0,0%
Ireland	59	1,3%	1,0%	0,9%	5,0%	-0,1%	0,3%
China	55	4,5%	4,0%	2,1%	13,0%	2,0%	0,5%
Mexico	55	4,3%	4,0%	1,0%	7,0%	2,8%	0,3%
Russia	48	7,4%	7,0%	2,7%	15,0%	3,0%	0,4%
India	43	7,4%	7,5%	1,1%	9,0%	4,0%	-0,1%
South Africa	43	8,2%	8,4%	0,9%	9,4%	5,5%	-0,1%
Australia	40	3,1%	3,0%	1,1%	6,0%	1,5%	0,2%
Chile	37	3,9%	3,8%	1,1%	6,1%	2,1%	0,1%
Uruguay	37	3,6%	3,6%	0,4%	4,1%	3,0%	0,0%
Poland	36	2,7%	2,7%	0,5%	3,8%	1,6%	0,0%
Peru	35	4,0%	4,0%	0,9%	6,0%	2,1%	0,0%
Czech Republic	34	1,8%	1,5%	1,1%	4,4%	0,2%	0,3%
Indonesia	34	7,5%	7,5%	0,4%	8,8%	7,0%	0,0%
Israel	34	0,9%	0,6%	1,0%	3,0%	-0,6%	0,3%
Korea (South)	33	2,3%	2,3%	0,6%	4,0%	1,5%	0,1%
Netherlands	33	1,8%	1,8%	0,6%	3,0%	0,5%	0,0%
New Zealand	31	2,9%	3,2%	0,9%	4,0%	1,0%	-0,3%
Thailand	31	8,7%	9,0%	2,1%	11,7%	3,0%	-0,3%
Turkey	31	7,8%	7,6%	0,7%	9,4%	6,0%	0,2%
Austria	29	2,8%	2,8%	1,2%	5,0%	0,5%	0,0%
Greece	29	15,0%	15,0%	5,5%	25,0%	5,0%	0,0%
Colombia	28	3,8%	3,6%	1,2%	6,0%	2,2%	0,2%
Hungary	28	0,6%	0,5%	1,0%	4,0%	-0,9%	0,1%
Venezuela	28	3,5%	3,2%	1,6%	8,0%	1,0%	0,3%
Argentina	26	12,6%	12,4%	4,5%	19,0%	3,2%	0,1%

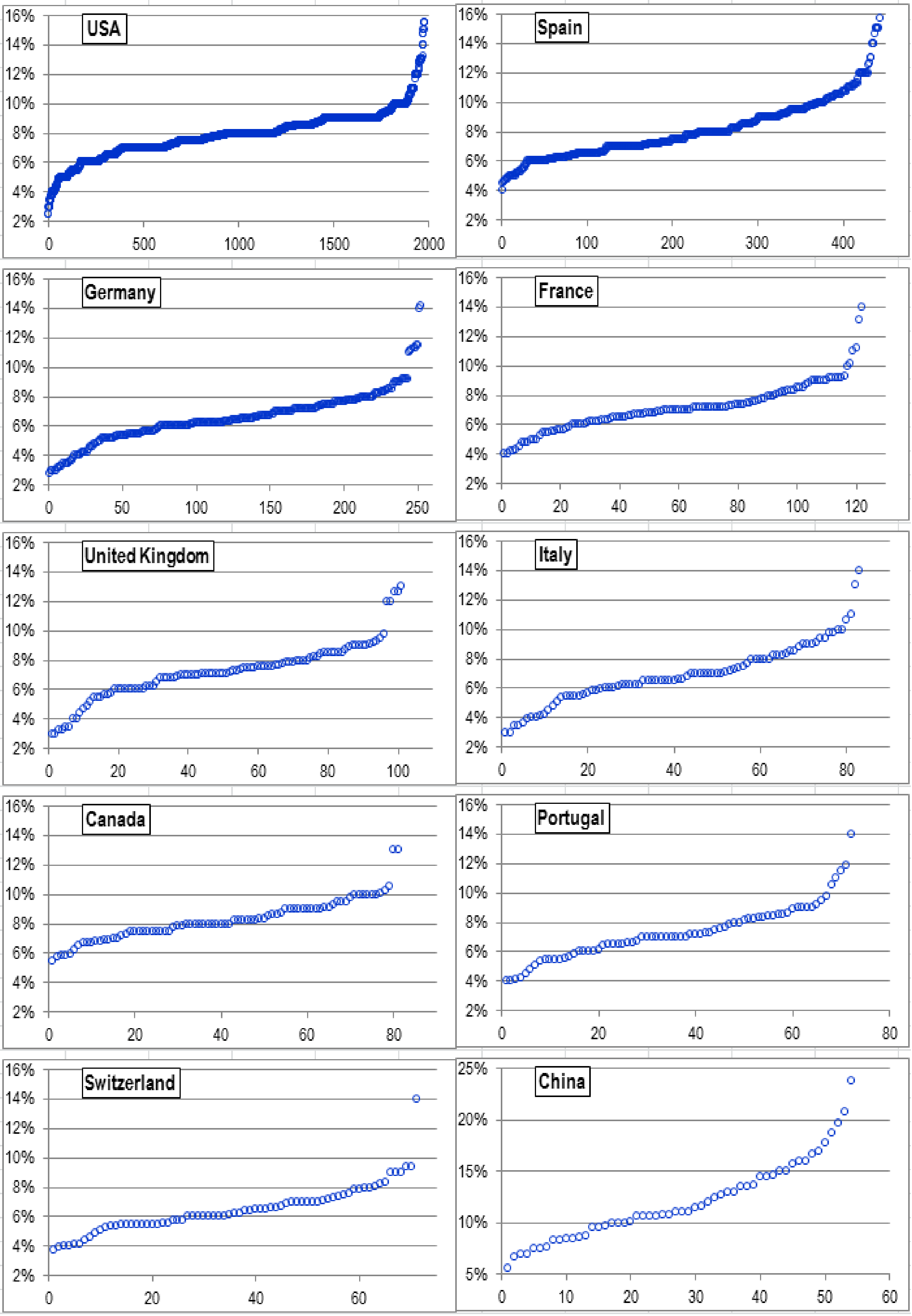
Table 4. Km [Required return to equity (market): RF + MRP)] used for 41 countries in 2015

Km	Number of answers	average	Median	St. Dev.	max	min	Av-Median
USA	1983	7,9%	8,0%	1,7%	22,0%	2,5%	-0,1%
Spain	443	8,1%	7,8%	2,0%	15,7%	4,1%	0,3%
Germany	252	6,6%	6,4%	1,7%	14,2%	2,8%	0,1%
France	122	7,2%	7,0%	1,6%	14,0%	4,0%	0,2%
United Kingdom	101	7,2%	7,1%	1,9%	13,0%	3,0%	0,2%
Italy	83	7,0%	6,7%	2,1%	14,0%	3,0%	0,3%
Canada	81	8,3%	8,0%	1,4%	13,0%	5,5%	0,3%
Portugal	72	7,3%	7,0%	1,9%	14,0%	4,0%	0,3%
Switzerland	71	6,5%	6,2%	1,6%	14,0%	3,7%	0,2%
Belgium	70	6,7%	6,9%	1,8%	14,0%	1,2%	-0,1%
Sweden	68	6,5%	6,4%	1,7%	14,0%	3,7%	0,1%
Denmark	65	6,8%	6,6%	1,8%	14,0%	4,0%	0,2%
Finland	64	6,9%	7,0%	1,6%	14,0%	4,0%	0,0%
Japan	61	6,6%	6,5%	2,4%	15,1%	2,3%	0,1%
Norway	61	6,8%	6,7%	1,9%	14,0%	3,6%	0,2%
Brazil	59	16,5%	16,6%	3,8%	23,0%	8,0%	-0,1%
Ireland	59	6,7%	6,6%	1,8%	14,0%	3,6%	0,1%
China	55	12,6%	11,0%	5,7%	43,0%	5,6%	1,6%
Mexico	55	12,2%	12,5%	1,6%	15,8%	9,0%	-0,3%
Russia	48	17,1%	17,0%	4,3%	25,0%	6,0%	0,1%
India	43	15,8%	15,8%	3,0%	21,0%	10,0%	0,0%
South Africa	43	15,9%	15,2%	2,9%	22,0%	11,7%	0,7%
Australia	40	9,2%	8,2%	4,3%	23,0%	3,4%	1,0%
Chile	37	10,4%	10,3%	1,3%	12,7%	7,8%	0,1%
Uruguay	37	10,6%	10,6%	1,2%	12,9%	8,5%	0,0%
Poland	36	7,9%	7,8%	1,4%	10,7%	5,1%	0,2%
Peru	35	11,2%	11,2%	1,6%	13,8%	8,0%	0,1%
Czech Republic	34	7,4%	7,2%	1,4%	10,5%	4,6%	0,2%
Indonesia	34	16,4%	16,4%	1,4%	21,3%	14,0%	0,0%
Israel	34	6,1%	6,1%	1,9%	9,0%	2,6%	0,0%
Korea (South)	33	8,5%	8,3%	2,0%	14,0%	5,6%	0,3%
Netherlands	33	7,6%	7,6%	0,9%	9,3%	6,2%	0,0%
New Zealand	31	9,5%	9,3%	0,8%	11,5%	8,5%	0,2%
Thailand	31	16,0%	16,3%	2,4%	19,8%	10,0%	-0,3%
Turkey	31	17,2%	16,6%	2,3%	24,0%	14,0%	0,6%
Austria	29	8,4%	8,4%	1,4%	11,0%	6,0%	0,0%
Greece	29	29,3%	30,0%	8,7%	45,4%	14,6%	-0,7%
Colombia	28	12,1%	12,1%	1,6%	15,3%	9,1%	0,0%
Hungary	28	9,5%	9,4%	1,4%	11,9%	6,9%	0,1%
Venezuela	28	23,1%	22,5%	4,7%	36,1%	15,3%	0,6%
Argentina	26	35,5%	29,6%	14,7%	63,2%	15,8%	5,9%

Table 5. Market Risk Premium (MRP), Risk Free Rate (R_F) and Km
(Required return to equity: Km = R_F + MRP) used for 41 countries in 2015

	n	average			st dev			Median		
		Km	R_F	MRP	Km	R_F	MRP	Km	R_F	MRP
USA	1983	7,9%	2,4%	5,5%	1,7%	1,1%	1,4%	8,0%	2,3%	5,3%
Spain	443	8,1%	2,2%	5,9%	2,0%	1,2%	1,6%	7,8%	2,0%	5,5%
Germany	252	6,6%	1,3%	5,3%	1,7%	0,8%	1,5%	6,4%	1,1%	5,1%
France	122	7,2%	1,5%	5,6%	1,6%	1,0%	1,4%	7,0%	1,3%	5,5%
United Kingdom	101	7,2%	2,1%	5,2%	1,9%	0,8%	1,7%	7,1%	2,1%	5,0%
Italy	83	7,0%	1,5%	5,4%	2,1%	1,1%	1,5%	6,7%	1,3%	5,2%
Canada	81	8,3%	2,3%	5,9%	1,4%	1,0%	1,3%	8,0%	2,0%	6,0%
Portugal	72	7,3%	1,6%	5,7%	1,9%	0,9%	1,5%	7,0%	1,5%	5,5%
Switzerland	71	6,5%	1,1%	5,4%	1,6%	0,7%	1,2%	6,2%	1,0%	5,0%
Belgium	70	6,7%	1,3%	5,5%	1,8%	0,9%	1,3%	6,9%	1,0%	5,4%
Sweden	68	6,5%	1,1%	5,4%	1,7%	0,8%	1,3%	6,4%	1,0%	5,1%
Denmark	65	6,8%	1,3%	5,5%	1,8%	1,0%	1,2%	6,6%	1,0%	5,5%
Finland	64	6,9%	1,2%	5,7%	1,6%	0,9%	1,1%	7,0%	1,0%	5,8%
Japan	61	6,6%	0,7%	5,8%	2,4%	1,0%	2,0%	6,5%	0,5%	6,0%
Norway	61	6,8%	1,4%	5,5%	1,9%	1,1%	1,2%	6,7%	1,0%	5,2%
Brazil	59	16,5%	9,0%	7,5%	3,8%	2,8%	2,1%	16,6%	9,0%	7,0%
Ireland	59	6,7%	1,3%	5,5%	1,8%	0,9%	1,3%	6,6%	1,0%	5,2%
China	55	12,6%	4,5%	8,1%	5,7%	2,1%	5,6%	11,0%	4,0%	7,0%
Mexico	55	12,2%	4,3%	8,0%	1,6%	1,0%	1,5%	12,5%	4,0%	8,0%
Russia	48	17,1%	7,4%	9,7%	4,3%	2,7%	2,9%	17,0%	7,0%	10,0%
India	43	15,8%	7,4%	8,4%	3,0%	1,1%	2,5%	15,8%	7,5%	8,3%
South Africa	43	15,9%	8,2%	7,7%	2,9%	0,9%	2,3%	15,2%	8,4%	7,3%
Australia	40	9,2%	3,1%	6,0%	4,3%	1,1%	4,0%	8,2%	3,0%	5,1%
Chile	37	10,4%	3,9%	6,5%	1,3%	1,1%	0,9%	10,3%	3,8%	6,5%
Uruguay	37	10,6%	3,6%	7,1%	1,2%	0,4%	0,9%	10,6%	3,6%	7,0%
Poland	36	7,9%	2,7%	5,2%	1,4%	0,5%	1,0%	7,8%	2,7%	5,0%
Peru	35	11,2%	4,0%	7,2%	1,6%	0,9%	1,2%	11,2%	4,0%	7,4%
Czech Republic	34	7,4%	1,8%	5,6%	1,4%	1,1%	0,7%	7,2%	1,5%	5,4%
Indonesia	34	16,4%	7,5%	8,9%	1,4%	0,4%	1,2%	16,4%	7,5%	9,0%
Israel	34	6,1%	0,9%	5,2%	1,9%	1,0%	1,1%	6,1%	0,6%	5,0%
Korea (South)	33	8,5%	2,3%	6,2%	2,0%	0,6%	1,5%	8,3%	2,3%	6,0%
Netherlands	33	7,6%	1,8%	5,9%	0,9%	0,6%	0,6%	7,6%	1,8%	6,0%
New Zealand	31	9,5%	2,9%	6,6%	0,8%	0,9%	1,3%	9,3%	3,2%	6,0%
Thailand	31	16,0%	8,7%	7,3%	2,4%	2,1%	0,9%	16,3%	9,0%	7,5%
Turkey	31	17,2%	7,8%	9,3%	2,3%	0,7%	2,5%	16,6%	7,6%	9,1%
Austria	29	8,4%	2,8%	5,7%	1,4%	1,2%	0,3%	8,4%	2,8%	5,6%
Greece	29	29,3%	15,0%	14,3%	8,7%	5,5%	5,8%	30,0%	15,0%	15,0%
Colombia	28	12,1%	3,8%	8,3%	1,6%	1,2%	1,4%	12,1%	3,6%	8,0%
Hungary	28	9,5%	0,6%	8,8%	1,4%	1,0%	0,8%	9,4%	0,5%	8,9%
Venezuela	28	23,1%	3,5%	19,6%	4,7%	1,6%	3,7%	22,5%	3,2%	19,3%
Argentina	26	35,5%	12,6%	22,9%	14,7%	4,5%	12,3%	29,6%	12,4%	20,1%

Figure 2. Km (Km = RF + MRP) used in 2015 for some countries (plot of answers)



2. Changes from 2013 to 2015

In this section, we compare the results of 2015 with the results of a similar survey collected in 2013 (see http://ssrn.com/abstract=914160).

Table 6. Market Risk Premium (MRP), Risk Free Rate (RF) and Km
Difference of the averages of the surveys of 2015 and 2013

	Average 2015			Average 2015 - Average 2013		
	Km	RF	MRP	Km	RF	MRP
USA	7,9%	2,4%	5,5%	-0,1%	0,0%	-0,2%
Spain	8,1%	2,2%	5,9%	-2,3%	-2,2%	-0,1%
Germany	6,6%	1,3%	5,3%	-0,9%	-0,6%	-0,2%
France	7,2%	1,5%	5,6%	-0,9%	-0,5%	-0,5%
United Kingdom	7,2%	2,1%	5,2%	-0,7%	-0,3%	-0,3%
Italy	7,0%	1,5%	5,4%	-3,0%	-2,9%	-0,3%
Canada	8,3%	2,3%	5,9%	0,9%	0,3%	0,5%
Portugal	7,3%	1,6%	5,7%	-3,9%	-3,5%	-0,4%
Switzerland	6,5%	1,1%	5,4%	-0,4%	-0,2%	-0,2%
Belgium	6,7%	1,3%	5,5%	-1,8%	-1,1%	-0,6%
Sweden	6,5%	1,1%	5,4%	-1,8%	-1,2%	-0,6%
Denmark	6,8%	1,3%	5,5%	-1,2%	-0,3%	-0,9%
Finland	6,9%	1,2%	5,7%	-1,6%	-0,5%	-1,1%
Japan	6,6%	0,7%	5,8%	-1,0%	-0,4%	-0,8%
Norway	6,8%	1,4%	5,5%	-1,6%	-1,0%	-0,5%
Brazil	16,5%	9,0%	7,5%	4,1%	3,1%	1,0%
Ireland	6,7%	1,3%	5,5%	-2,8%	-2,0%	-0,7%
China	12,6%	4,5%	8,1%	1,1%	0,7%	0,4%
Mexico	12,2%	4,3%	8,0%	1,0%	-0,6%	1,3%
Russia	17,1%	7,4%	9,7%	3,7%	1,8%	2,4%
India	15,8%	7,4%	8,4%	0,4%	0,5%	-0,1%
South Africa	15,9%	8,2%	7,7%	2,7%	1,8%	0,9%
Australia	9,2%	3,1%	6,0%	-1,0%	-0,2%	-0,8%
Chile	10,4%	3,9%	6,5%	0,6%	-0,9%	1,5%
Poland	7,9%	2,7%	5,2%	-2,3%	-1,2%	-1,1%
Peru	11,2%	4,0%	7,2%	0,0%	-0,7%	0,7%
Czech Republic	7,4%	1,8%	5,6%	-0,7%	0,1%	-0,9%
Indonesia	16,4%	7,5%	8,9%	2,7%	1,6%	1,1%
Israel	6,1%	0,9%	5,2%	-3,6%	-2,4%	-1,2%
Korea (South)	8,5%	2,3%	6,2%	-1,7%	-0,8%	-0,8%
Netherlands	7,6%	1,8%	5,9%	-0,4%	-0,2%	-0,1%
New Zealand	9,5%	2,9%	6,6%	0,6%	-0,6%	1,2%
Thailand	16,0%	8,7%	7,3%	4,7%	5,0%	-0,3%
Turkey	17,2%	7,8%	9,3%	2,8%	1,7%	1,1%
Austria	8,4%	2,8%	5,7%	0,1%	0,4%	-0,3%
Greece	29,3%	15,0%	14,3%	12,5%	5,4%	7,0%
Colombia	12,1%	3,8%	8,3%	-1,0%	-0,8%	-0,1%
Hungary	9,5%	0,6%	8,8%	-3,2%	-4,0%	0,6%
Venezuela	23,1%	3,5%	19,6%	3,9%	-4,6%	8,4%
Argentina	35,5%	12,6%	22,9%	16,1%	3,9%	12,3%

Table 7. Market Risk Premium (MRP), Risk Free Rate (RF) and Km
Difference of the averages and of the St. Dev. of the surveys of 2015 and 2013

	Average 2015 -Average 2013			average	St. Dev. 2015 -St. Dev. 2013		
	Km	RF	MRP	Km 2015	Km	RF	MRP
Argentina	16,1%	3,9%	12,3%	35,5%	8,4%	1,7%	4,2%
Greece	12,5%	5,4%	7,0%	29,3%	4,9%	4,2%	1,7%
Thailand	4,7%	5,0%	-0,3%	16,0%	2,2%	1,5%	0,3%
Brazil	4,1%	3,1%	1,0%	16,5%	0,4%	0,4%	0,0%
Venezuela	3,9%	-4,6%	8,4%	23,1%	1,1%	-1,4%	1,9%
Russia	3,7%	1,8%	2,4%	17,1%	0,5%	1,3%	-1,2%
Turkey	2,8%	1,7%	1,1%	17,2%	-1,3%	0,0%	-0,4%
Indonesia	2,7%	1,6%	1,1%	16,4%	0,2%	-0,1%	-0,2%
South Africa	2,7%	1,8%	0,9%	15,9%	1,4%	0,3%	0,9%
China	1,1%	0,7%	0,4%	12,6%	3,3%	1,5%	3,3%
Mexico	1,0%	-0,6%	1,3%	12,2%	-1,0%	-0,2%	-0,9%
Canada	0,9%	0,3%	0,5%	8,3%	0,0%	0,5%	0,0%
New Zealand	0,6%	-0,6%	1,2%	9,5%	-0,9%	0,3%	-0,5%
Chile	0,6%	-0,9%	1,5%	10,4%	-1,3%	-1,5%	-1,3%
India	0,4%	0,5%	-0,1%	15,8%	-0,7%	-0,2%	-0,4%
Austria	0,1%	0,4%	-0,3%	8,4%	-0,8%	0,0%	-1,6%
Peru	0,0%	-0,7%	0,7%	11,2%	-2,0%	-1,1%	-0,9%
USA	-0,1%	0,0%	-0,2%	7,9%	-0,3%	0,1%	-0,2%
Netherlands	-0,4%	-0,2%	-0,1%	7,6%	-0,4%	-0,2%	-0,7%
Switzerland	-0,4%	-0,2%	-0,2%	6,5%	0,1%	0,4%	-0,3%
United Kingdom	-0,7%	-0,3%	-0,3%	7,2%	0,2%	-0,2%	0,3%
Czech Republic	-0,7%	0,1%	-0,9%	7,4%	0,2%	0,6%	-0,4%
Germany	-0,9%	-0,6%	-0,2%	6,6%	-0,1%	0,2%	-0,2%
France	-0,9%	-0,5%	-0,5%	7,2%	-0,3%	0,0%	-0,2%
Colombia	-1,0%	-0,8%	-0,1%	12,1%	-2,2%	0,5%	-2,0%
Japan	-1,0%	-0,4%	-0,8%	6,6%	-0,2%	0,2%	-0,7%
Australia	-1,0%	-0,2%	-0,8%	9,2%	-0,6%	0,5%	-0,9%
Denmark	-1,2%	-0,3%	-0,9%	6,8%	0,9%	0,9%	0,4%
Finland	-1,6%	-0,5%	-1,1%	6,9%	0,4%	0,5%	-0,1%
Norway	-1,6%	-1,0%	-0,5%	6,8%	-0,2%	-0,1%	-0,6%
Korea (South)	-1,7%	-0,8%	-0,8%	8,5%	0,0%	0,0%	-0,3%
Belgium	-1,8%	-1,1%	-0,6%	6,7%	-0,3%	-0,3%	-0,5%
Sweden	-1,8%	-1,2%	-0,6%	6,5%	-0,4%	-0,4%	-0,4%
Poland	-2,3%	-1,2%	-1,1%	7,9%	0,5%	0,1%	0,0%
Spain	-2,3%	-2,2%	-0,1%	8,1%	-0,1%	0,3%	-0,1%
Ireland	-2,8%	-2,0%	-0,7%	6,7%	-1,6%	0,6%	-2,0%
Italy	-3,0%	-2,9%	-0,3%	7,0%	0,5%	0,5%	0,0%
Hungary	-3,2%	-4,0%	0,6%	9,5%	-1,0%	-1,0%	-0,8%
Israel	-3,6%	-2,4%	-1,2%	6,1%	0,1%	0,1%	0,0%
Portugal	-3,9%	-3,5%	-0,4%	7,3%	-0,5%	0,1%	-0,8%

Figure 3. Km (RF + MRP) and Risk Free Rate (RF)
Difference of the averages of the surveys of 2015 and 2013. *Source: Table 7*

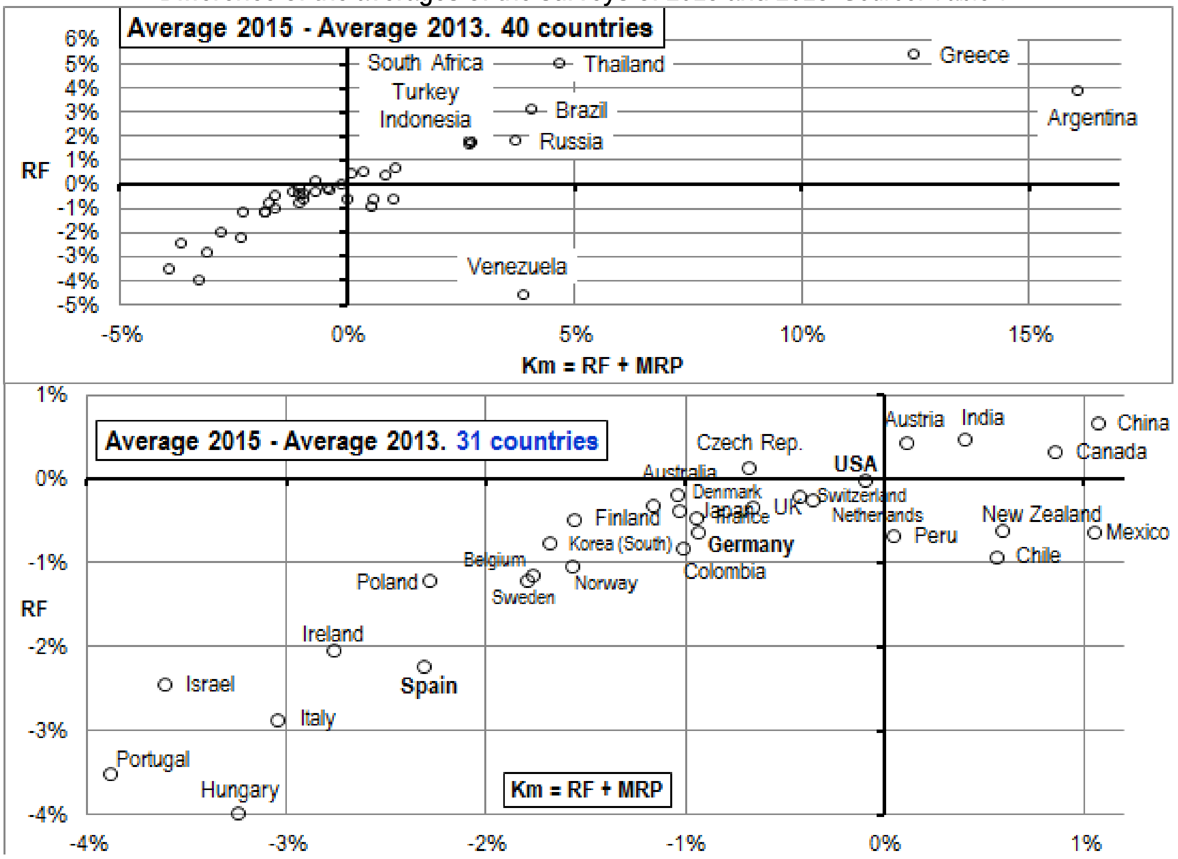
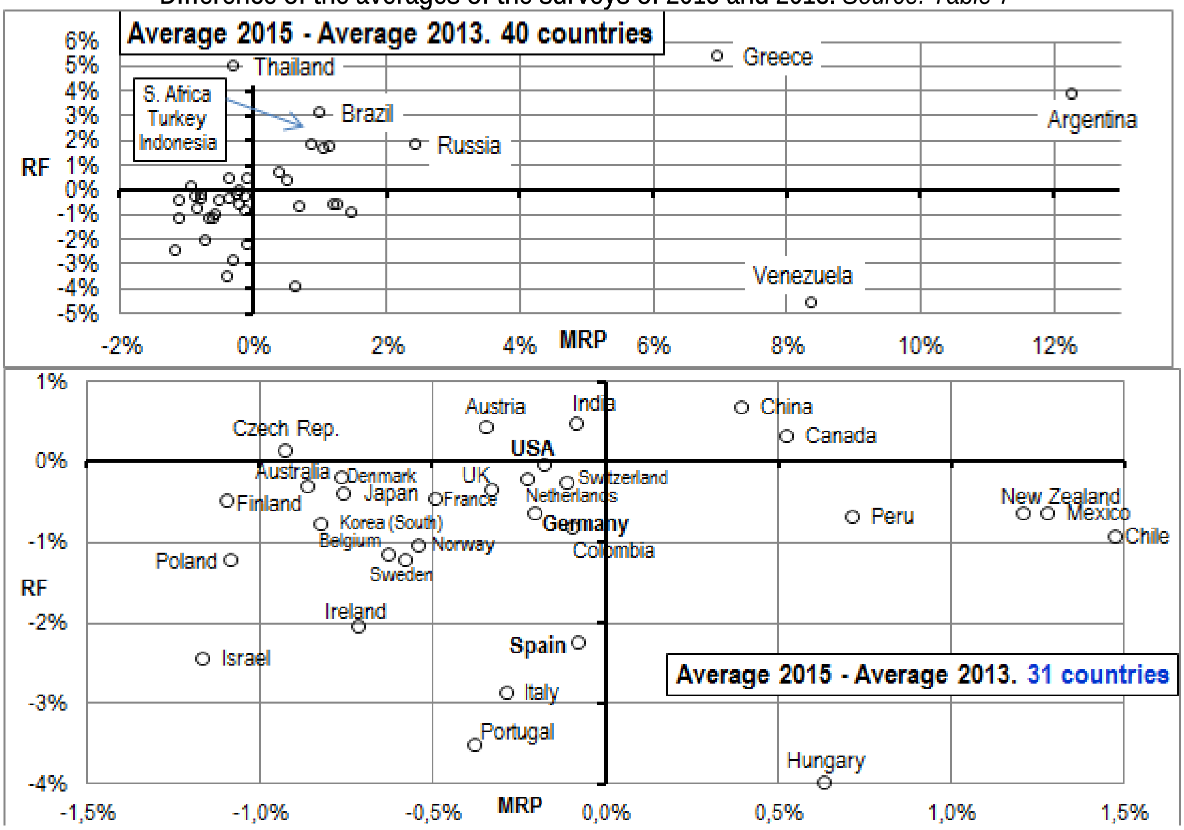


Figure 4. MRP and Risk Free Rate (RF)
Difference of the averages of the surveys of 2015 and 2013. *Source: Table 7*



3. RF used in 2013 and 2015 for US, Europe and UK vs. yield of the 10-year Government bonds

Figure 5. Yield on 10-year Gov. Bonds. 4 Countries

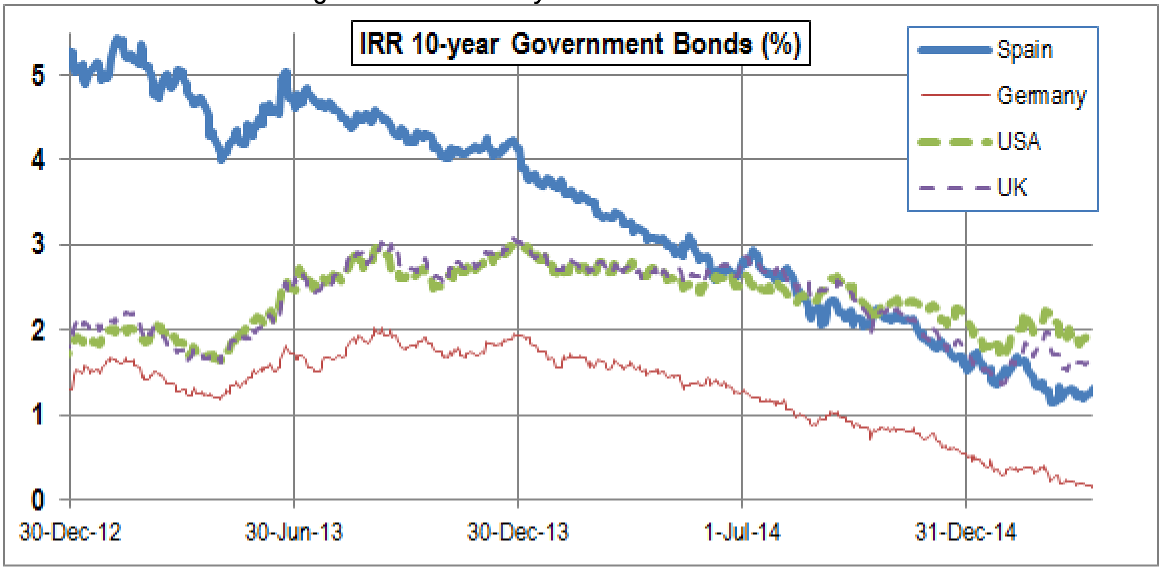


Table 8 shows that most of the respondents use for US, Europe and UK a Risk-Free Rate (RF) higher than the yield of the 10-year Government bonds.

Table 8. Yield on 10-year Gov. Bonds and RF used in 2013 and 2015
4 Countries: USA, Germany, Spain and UK

		USA	Germany	Spain	UK
Average 10-year Government Bonds	May 2013	1,9%	1,4%	4,2%	1,9%
	March-april 2015	2,0%	0,2%	1,3%	1,7%
RF used in May 2013	average	2.4%	1.9%	4.4%	2.4%
	St. Dev.	1.0%	0.6%	0.9%	1.0%
	max	6.0%	6.5%	6.0%	7.0%
	min	0.1%	0.1%	0.5%	0.2%
RF used in March-April 2015	average	2,4%	1,3%	2,2%	2,1%
	St. Dev.	1,1%	0,8%	1,2%	0,8%
	max	8,0%	5,1%	7,0%	6,0%
	min	0,0%	-0,2%	0,0%	0,4%

4. Previous surveys

Previous surveys. Market risk premium used

2008	http://ssrn.com/abstract=1344209
2010	http://ssrn.com/abstract=1606563; http://ssrn.com/abstract=1609563
2011	http://ssrn.com/abstract=1822182; http://ssrn.com/abstract=1805852
2012	http://ssrn.com/abstract=2084213
2013	http://ssrn.com/abstract=914160
2014	http://ssrn.com/abstract=1609563

Welch (2000) performed two surveys with finance professors in 1997 and 1998, asking them what they thought the Expected MRP would be over the next 30 years. He obtained 226 replies, ranging from 1% to 15%, with an average arithmetic EEP of 7% above T-Bonds.[2] Welch (2001) presented the results of a survey of 510 finance and economics professors performed in August 2001 and the consensus for the 30-year arithmetic EEP was 5.5%, much lower than just 3 years earlier. In an update published in 2008 Welch reports that the MRP "used in class" in December 2007 by about 400 finance professors was on average 5.89%, and 90% of the professors used equity premiums between 4% and 8.5%.

Johnson et al (2007) report the results of a survey of 116 finance professors in North America done in March 2007: 90% of the professors believed the Expected MRP during the next 30 years to range from 3% to 7%.

Graham and Harvey (2007) indicate that U.S. CFOs reduced their average EEP from 4.65% in September 2000 to 2.93% by September 2006 (st. dev. of the 465 responses = 2.47%). In the 2008 survey, they report an average EEP of 3.80%, ranging from 3.1% to 11.5% at the tenth percentile at each end of the spectrum. They show that average EEP changes through time. Goldman Sachs (O'Neill, Wilson and Masih 2002) conducted a survey of its global clients in July 2002 and the average long-run EEP was 3.9%, with most responses between 3.5% and 4.5%.

Ilmanen (2003) argues that surveys tend to be optimistic: *"survey-based expected returns may tell us more about hoped-for returns than about required returns".* Damodaran (2008) points out that *"the risk premiums in academic surveys indicate how far removed most academics are from the real world of valuation and corporate finance and how much of their own thinking is framed by the historical risk premiums... The risk premiums that are presented in classroom settings are not only much higher than the risk premiums in practice but also contradict other academic research".*

Table 4 of Fernandez et al (2011a) shows the evolution of the Market Risk Premium used for the USA in 2011, 2010, 2009 and 2008 according to previous surveys (Fernandez et al, 2009, 2010a and 2010b).

Table 9. Comparison of previous surveys

	Surveys of Ivo Welch					Fernandez et al (2009, 2010)			
	Oct 97– Feb 98*	Jan-May 99+	Sep 2001**	Dec. 2007#	January 2009++	US 2008	Europe 2008	US 2009	Europe 2009
Number of answers	226	112	510	360	143	487	224	462	194
Average	7.2	6.8	4.7	5.96	6.2	6.3	5.3	6.0	5.3
Std. Deviation	*2.0*	*2.0*	*2.2*	*1.7*	*1.7*	*2.2*	*1.5*	*1.7*	*1.7*
Max	15	15	20	20		19.0	10.0	12.0	12.0
Q3	8.4	8	6	7.0	7	7.2	6.0	7.0	6.0
Median	7	7	4.5	6.0	6	6.0	5.0	6.0	5.0
Q1	6	5	3	5.0	5	5.0	4.1	5.0	5.3
Min	1.5	1.5	0	2		0.8	1.0	2.0	2.0

* 30-Year Forecast. Welch (2000) First survey + 30-Year Forecast. Welch (2000) Second survey
** 30 year Equity Premium Forecast (Geometric). "The Equity Premium Consensus Forecast Revisited" (2001)
\# 30-Year Geo Eq Prem Used in class. Welch, I. (2008), "The Consensus Estimate for the Equity Premium by Academic Financial Economists in December 2007". http://ssrn.com/abstract=1084918
++ In your classes, what is the main number you are recommending for long-term CAPM purposes? "Short Academic Equity Premium Survey for January 2009". http://welch.econ.brown.edu/academics/equpdate-results2009.html

Table 10. Estimates of the EEP (Expected Equity Premium) according to other surveys

Authors	Conclusion about EEP	Respondents
Pensions and Investments (1998)	3%	Institutional investors
Graham and Harvey (2007)	Sep. 2000. Mean: 4.65%. Std. Dev. = 2.7%	CFOs
Graham and Harvey (2007)	Sep. 2006. Mean: 2.93%. Std. Dev. = 2.47%	CFOs
Welch update	December 2007. Mean: 5.69%. Range 2% to 12%	Finance professors
O'Neill, Wilson and Masih (2002)	3.9%	Global clients Goldman

[2] At that time, the most recent Ibbotson Associates Yearbook reported an arithmetic HEP versus T-bills of 8.9% (1926–1997).

The magazine *Pensions and Investments* (12/1/1998) carried out a survey among professionals working for institutional investors: the average EEP was 3%. Shiller[3] publishes and updates an index of investor sentiment since the crash of 1987. While neither survey provides a direct measure of the equity risk premium, they yield a broad measure of where investors or professors expect stock prices to go in the near future. The 2004 survey of the Securities Industry Association (SIA) found that the median EEP of 1500 U.S. investors was about 8.3%. Merrill Lynch surveys more than 300 institutional investors globally in July 2008: the average EEP was 3.5%.

A main difference of this survey with previous ones is that this survey asks about the **Required** MRP, while most surveys are interested in the **Expected** MRP.

5. Expected and Required Equity Premium: different concepts

Fernandez and F.Acín (2015) claim and show that Expected Return and Required Return are two very different concepts.

Fernandez (2007, 2009b) claims that the term “equity premium” is used to designate four different concepts:

1. **Historical** equity premium (HEP): historical differential return of the stock market over treasuries.
2. **Expected** equity premium (EEP): expected differential return of the stock market over treasuries.
3. **Required** equity premium (REP): incremental return of a diversified portfolio (the market) over the risk-free rate required by an investor. It is used for calculating the required return to equity.
4. **Implied** equity premium (IEP): the required equity premium that arises from assuming that the market price is correct.

The four concepts (HEP, REP, EEP and IEP) designate different realities. The **HEP** is easy to calculate and is equal for all investors, provided they use the same time frame, the same market index, the same risk-free instrument and the same average (arithmetic or geometric). But the **EEP,** the **REP** and the **IEP** may be different for different investors and are not observable.

The **HEP** is the historical average differential return of the market portfolio over the risk-free debt. The most widely cited sources are Ibbotson Associates and Dimson *et al.* (2007).

Numerous papers and books assert or imply that there is a “market” EEP. However, it is obvious that investors and professors do not share “homogeneous expectations” and have different assessments of the **EEP**. As Brealey et al. (2005, page 154) affirm, *“Do not trust anyone who claims to know what returns investors expect”.*

The **REP** is the answer to the following question: What incremental return do I require for investing in a diversified portfolio of shares over the risk-free rate? It is a crucial parameter because the REP is the key to determining the company’s required return to equity and the WACC. Different companies may use, and in fact do use, different **REPs**.

The **IEP** is the implicit REP used in the valuation of a stock (or market index) that matches the current market price. The most widely used model to calculate the IEP is the dividend discount model: the current price per share (P_0) is the present value of expected dividends discounted at the required rate of return (Ke). If d_1 is the dividend per share expected to be received in year 1, and g the expected long term growth rate in dividends per share,

$P_0 = d_1 / (Ke - g)$, which implies: $IEP = d_1/P_0 + g - R_F$ (1)

The estimates of the IEP depend on the particular assumption made for the expected growth (g). Even if market prices are correct for all investors, there is not an IEP common for all investors: there are many pairs (IEP, g) that accomplish equation (1). Even if equation (1) holds for every investor, there are many *required* returns (as many as expected growths, g) in the market. Many papers in the financial literature report different estimates of the IEP with great dispersion, as for example, Claus and Thomas

[3] See http://icf.som.yale.edu/Confidence.Index

(2001, IEP = 3%), Harris and Marston (2001, IEP = 7.14%) and Ritter and Warr (2002, IEP = 12% in 1980 and -2% in 1999). There is no a common **IEP** for all investors.

For a particular investor, the **EEP** is not necessary equal to the REP (unless he considers that the market price is equal to the value of the shares). Obviously, an investor will hold a diversified portfolio of shares if his EEP is higher (or equal) than his REP and will not hold it otherwise.

We can find out the REP and the EEP of an investor by asking him, although for many investors the REP is not an explicit parameter but, rather, it is implicit in the price they are prepared to pay for the shares. However, it is not possible to determine the REP for the market as a whole, because it does not exist: even if we knew the REPs of all the investors in the market, it would be meaningless to talk of a REP for the market as a whole. There is a distribution of REPs and we can only say that some percentage of investors have REPs contained in a range. The average of that distribution cannot be interpreted as the REP of the market nor as the REP of a representative investor.

Much confusion arises from not distinguishing among the four concepts that the phrase *equity premium* designates: Historical equity premium, Expected equity premium, Required equity premium and Implied equity premium. 129 of the books reviewed by Fernandez (2009b) identify Expected and Required equity premium and 82 books identify Expected and Historical equity premium.

Finance textbooks should clarify the MRP by incorporating distinguishing definitions of the four different concepts and conveying a clearer message about their sensible magnitudes.

6. Conclusion

Most previous surveys have been interested in the Expected MRP, but this survey asks about the Required MRP.

This paper contains the statistics of a survey about the Risk-Free Rate (R_F) and of the Market Risk Premium (MRP) used in 2015 for **41 countries**. We got answers for 68 countries, but we only report the results for 41 countries with more than 25 answers.

The average (R_F) used in 2015 was smaller than the one used in 2013 in 26 countries (in 11 of them the difference was more than 1%). In 8 countries the average (R_F) used in 2015 was more than a 1% higher than the one used in 2013.

The change between 2013 and 2015 of the average Market risk premium used was higher than 1% for 13 countries.

Most of the respondents use for US, Europe and UK a Risk-Free Rate (R_F) higher than the yield of the 10-year Government bonds.

This survey links with the *Equity Premium Puzzle*: Fernandez et al (2009), argue that the equity premium puzzle may be explained by the fact that many market participants (equity investors, investment banks, analysts, companies…) do not use standard theory (such as a standard representative consumer asset pricing model…) for determining their Required Equity Premium, but rather, they use historical data and advice from textbooks and finance professors. Consequently, ex-ante equity premia have been high, market prices have been consistently undervalued, and the ex-post risk premia has been also high. Many investors use historical data and textbook prescriptions to estimate the required and the expected equity premium, the undervaluation and the high ex-post risk premium are self fulfilling prophecies.

EXHIBIT 1. Mail sent on March 2015

Survey Market Risk Premium and Risk-Free Rate 2015

We are doing a **survey** about the **Market Risk Premium** (MRP or Equity Premium) and **Risk Free Rate** that companies, analysts, regulators and professors use to calculate the **required return on equity (Ke)** in different countries.

I would be grateful if you would kindly answer the following 2 questions. No companies, individuals or universities will be identified, and only aggregate data will be made public. I will send you the results in a month.

Best regards and thanks,
Pablo Fernandez. Professor of Finance. IESE Business School. Spain.

2 questions:
1. The Market Risk Premium that I am using in 2015
for USA is: _______ %
for___________ is: _______ %
for___________ is: _______ %
for___________ is: _______ %

2. The Risk Free rate that I am using in 2015
for USA is: _______ %
for ___________ is: _______ %
for ___________ is: _______ %
for ___________ is: _______ %

EXHIBIT 2. Some comments and webs recommended by respondents

Equity premium: http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/ctryprem.html
http://www.market-risk-premia.com/market-risk-premia.html
http://www.marktrisikoprämie.de/marktrisikopraemien.html

US risk free rate: http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldYear&year=2015

risk free rate: http://www.basiszinskurve.de/basiszinssatz-gemaess-idw.html

http://www.econ.yale.edu/~shiller/

http://www.cfosurvey.org/pastresults.htm

http://alephblog.com/

In my DCF valuation I use a global perspective of the marginal investor hence a global MRP.

I match rf with currency/inflation of cash flows being discounted and do not rely too much on current interest rates due to imperfections in the market. The MRP is made consistent with the level of interest rate I use in my model (E(Rm)-Rf) end end up with 6%

For equities we use a 10% as a cost of opportunity independently of the level of interest.

Rf: average last 5-year USA 10 year Treasury

I would like to help you with these two questions, but the problem is that in no any literature sources or analytical reports I met the calculation of Market Risk Premium and Risk Free rate for Uzbekistan.

The risk free rate that I use depends upon the timing of the future cash flows. I refer to the interest rate swap market and the US treasury market for starters. These days, one has to bear in mind currency volatility as that has a bigger effect on PV than market cost-of-capital.

We use the same Market Risk Premium for any country: 5,75% (source: Damodaran). Only Rf changes.

I am happy that you are asking the second question, because it accounts for what I consider to be a historical anomaly in the reply to the first question. I've concluded that the ERP was recently 3-4 percent. But I think US monetary policy (the various "QE" programs) have in the past couple of years distorted the traditional relationship between expected total market returns and the risk free rate. QE has been driving the US Treasury rate down, while the expected total market return has held steady, leading to a larger than usual market risk premium. This higher market risk premium is not a sign of higher market equity risk, but of the perverse impact of aggressive monetary policy.

In my most recent work on this, I made use of relatively new Standard and Poor indexes that attempts to track the equity risk premium from the spread between expected returns on index futures for the S&P 500 and Treasury bond futures. There are two forms of this index, which go by the symbols “SPUSERPP” and “SPUSERPT.” The first is the index for the equity risk premium proper, and the second is an index for the expected total market return. The historical data for these indices only goes back to early 2011, but I think they may turn out to be a useful objective measure for variables that are frequently contentious because they have heretofore been estimated using less objective or more controversial methods.

For the US in 2015: MRP: 14% (as US equities are even more highly priced than last year).

Rf: 8% (the long term average growth rate in the money supply).

I use cash flows that have been adjusted for risk, and apply to these cash flows the risk-free Rate.

Both No. 1 and No. 2: 100%

Interest rates are artificially well below historic levels. Thus, bonds and equities values are artificially inflated.

I do not use "canned" rates applicable for a whole year. The rates I use are time-specific and case-specific, depending on conditions prevailing as of the valuation date.

For the 1st question, definition of investopedia http://www.investopedia.com/terms/m/marketriskpremium.asp for market risk premium can be utilized, , which is defined as the difference between expected rate of return on market portfolio and risk free rate. An exemplary portfolio consist of US bonds can be deemed to be market portfolio. Consequently, the difference will be market risk premium.

Implied ERP on March 1, 2015= 5.67% (Trailing 12 month cash yield); 6.09% (Normalized cash flow); 5.11% (Net cash yield)

1. The Market Risk Premium that I am using in 2015 for USA is: 5,5 %; for 2016 is: 5,5 %; for 2017 is: 5,5 %

2. The Risk Free rate that I am using in 2015 for USA is: 2 %; for 2016 is: 3 %; for 2017 is: 3,5 %; for 2018 is: 4 %

References

Brealey, R.A., S.C. Myers and F. Allen (2005), *Principles of Corporate Finance*, 8th edition, McGraw-Hill/Irwin.

Claus, J.J. and J.K. Thomas (2001), “Equity Premia as Low as Three Percent? Evidence from Analysts’ Earnings Forecasts for Domestic and International Stock Markets,” *Journal of Finance*. 55, (5), pp. 1629-66.

Damodaran, A. (2008), “Equity Risk Premiums (ERP): Determinants, Estimation and Implications”, Working Paper.

Dimson, E., P. Marsh and M. Staunton (2007), “The Worldwide Equity Premium: A Smaller Puzzle,” in *Handbook of investments: Equity risk premium*, R. Mehra, Elsevier.

Fernandez, P. (2007), “Equity Premium: Historical, Expected, Required and Implied”, http://ssrn.com/abstract=933070

Fernandez, P. (2009a), “Market Risk Premium Used in 2008 by Professors: A Survey with 1,400 Answers”, http://ssrn.com/abstract=1344209

Fernandez, P. (2009b), “The Equity Premium in 150 Textbooks”, http://ssrn.com/abstract=1473225

Fernandez, P., J. Aguirreamalloa and H. Liechtenstein (2009), “The Equity Premium Puzzle: High Required Premium, Undervaluation and Self Fulfilling Prophecy”. IESE Business School WP. http://ssrn.com/abstract=1274816

Fernandez, P. and J. del Campo (2010a), "Market Risk Premium used in 2010 by Analysts and Companies: a survey with 2,400 answers”, downloadable in http://ssrn.com/abstract=1609563

Fernandez, P. and J. del Campo (2010b), “Market Risk Premium Used in 2010 by Professors: A Survey with 1,500 Answers”, downloadable in http://ssrn.com/abstract=1606563

Fernandez, P., J. Aguirreamalloa and L. Corres (2011a), “US Market Risk Premium Used in 2011 by Professors, Analysts and Companies: A Survey with 5.731 Answers”, downloadable in http://ssrn.com/abstract=1805852

Fernandez, P., J. Aguirreamalloa and L. Corres (2011b), “The Equity Premium in Spain: Survey 2011 (in Spanish)”, downloadable in http://ssrn.com/abstract=1822422

Fernandez, P., J. Aguirreamalloa and L. Corres (2011c), "Market Risk Premium Used in 56 Countries in 2011: A Survey with 6,014 Answers", downloadable in http://ssrn.com/abstract=1822182

Fernandez, P., J. Aguirreamalloa and P. Linares (2014), "Market Risk Premium and Risk Free Rate Used for 51 Countries in 2013: A Survey with 6,237 Answers", downloadable in http://ssrn.com/abstract=914160

Fernandez, P., J. Aguirreamalloa and L. Corres (2012), "Market Risk Premium Used in 82 Countries in 2012: A Survey with 7,192 Answers", downloadable in http://ssrn.com/abstract=2084213

Fernandez, P. and I. F. Acín (2015), "Expected and Required Returns: Very Different Concepts", downloadable in http://ssrn.com/abstract=2591319

Fernandez, P., P. Linares and I. F. Acín (2014), "Market Risk Premium Used in 88 Countries in 2014: A Survey with 8,228 Answers", downloadable in http://ssrn.com/abstract=2450452

Graham, J.R. and C.R. Harvey (2007), "The Equity Risk Premium in January 2007: Evidence from the Global CFO Outlook Survey," *Icfai Journal of Financial Risk Management,* Vol. IV, No. 2, pp. 46-61.

Harris, R.S. and F.C. Marston (2001), "The Market Risk Premium: Expectational Estimates Using Analysts' Forecasts," *Journal of Applied Finance*, Vol. 11.

Ilmanen, A. (2003), "Expected returns on stocks and bonds", *Journal of Portfolio Management* 29, pp. 7-27.

Johnson, D. T., T. Kochanek, T and J. Alexander (2007), "The Equity Premium Puzzle: A New Look", *Journal of the Academy of Finance,* Vol. 5, No. 1, pp. 61-71.

O'Neill, J., D. Wilson and R. Masih (2002), "The Equity Risk Premium from an Economics Perspective", Goldman Sachs, Global Economics Paper No. 84.

Ritter, J.R. and R. Warr (2002), "The Decline of Inflation and the Bull Market of 1982 to 1999," *Journal of Financial and Quantitative Analysis*, Vol. 37, No. 1, pp. 29-61.

Welch, I. (2000), "Views of Financial Economists on the Equity Premium and on Professional Controversies", *Journal of Business*, Vol. 73, No. 4, pp. 501-537.

Welch, I. (2001), "The Equity Premium Consensus Forecast Revisited", Cowles Foundation Discussion Paper No. 1325.

Welch, I. (2007), "A Different Way to Estimate the Equity Premium (for CAPM and One-Factor Model Use Only)," SSRN n. 1077876.

Huge dispersion of the Risk-Free Rate and Market Risk Premium used by analysts in USA and Europe in 2015

Pablo Fernandez. Professor of Finance. IESE Business School, fernandezpa@iese.edu
Alberto Ortiz. Research Assistant. IESE Business School. AOrtiz@iese.edu
Isabel F. Acín. Independent researcher. University of Navarra. ifernandez.28@alumni.unav.es

ABSTRACT

We look at the Risk-Free Rate (R_F) and the Market Risk Premium (MRP) used by analysts in 2015 to value companies of six countries.

The dispersion of both, the RF and the MRP used, is huge, and the most unexpected result is that the dispersion is higher for the RF than for the MRP.

We also find that some analysts have more freedom than others do.

The data permits other comparisons. For example: Does it make sense that the average MRP used for Germany is higher than the average MRP used for France, Italy, Spain or the UK?

Most of the analysts use a Risk-Free Rate (R_F) higher than the yield of the 10-year Government bonds. A reason for it and for the huge dispersion may be the activity of the European Central Bank (ECB). The risk-free rate (R_F) is the required return to Government bonds when nobody (not even the ECB) manipulates the market. A question arises: May we consider the Quantitative Easing (QE) implemented by the ECB in 2014 and 2015 "market abuse", "market manipulation", a way of "altering competitive markets"…?

JEL Classification: G12, G31, M21

Keywords: analyst, market risk premium; required equity premium; risk-free rate

November 2, 2015

xPpLhmlsjmOo

1. RF and MRP used in 156 valuation reports

We revised more than 1,000 analyst reports about companies with headquarters in six countries: France, Germany, Italy, Spain, UK and USA. We looked for reports that indicated the Risk-Free Rate (**RF**) and the Market Risk Premium (**MRP**) used by the analyst in the valuation. We found only 156. Exhibit 1 contains the date, the company of the financial analyst, the company valued, and the RF and MRP used. The analysts belong to 35 different companies and the reports refer to 99 different companies.

Figures 1 and 2 contain the Risk-Free Rate (**RF**) and the Market Risk Premium (**MRP**) used in 2015 in by the financial analysts in the 156 reports. The dispersion is huge.

Table 1 contains the statistics of the **RF**, **MRP** and (**RF** + **MRP**) that appear in Figure 1. The most unexpected result is that the (Standard deviation / average) is higher for **RF** than for **MRP** in the six countries: the dispersion is higher for the **RF** used than for the **MRP** used.

The reader can do also other comparisons and assessments. For example: Does it make sense that the average MRP used for Germany is higher than the average MRP used for France, Italy, Spain or the UK? Does it make sense that the MRP and the RF used have positive correlation only in France?

Figure 1. RF, MRP and (RF + MRP) used in 2015 by 156 analysts in their valuations of companies of 6 countries

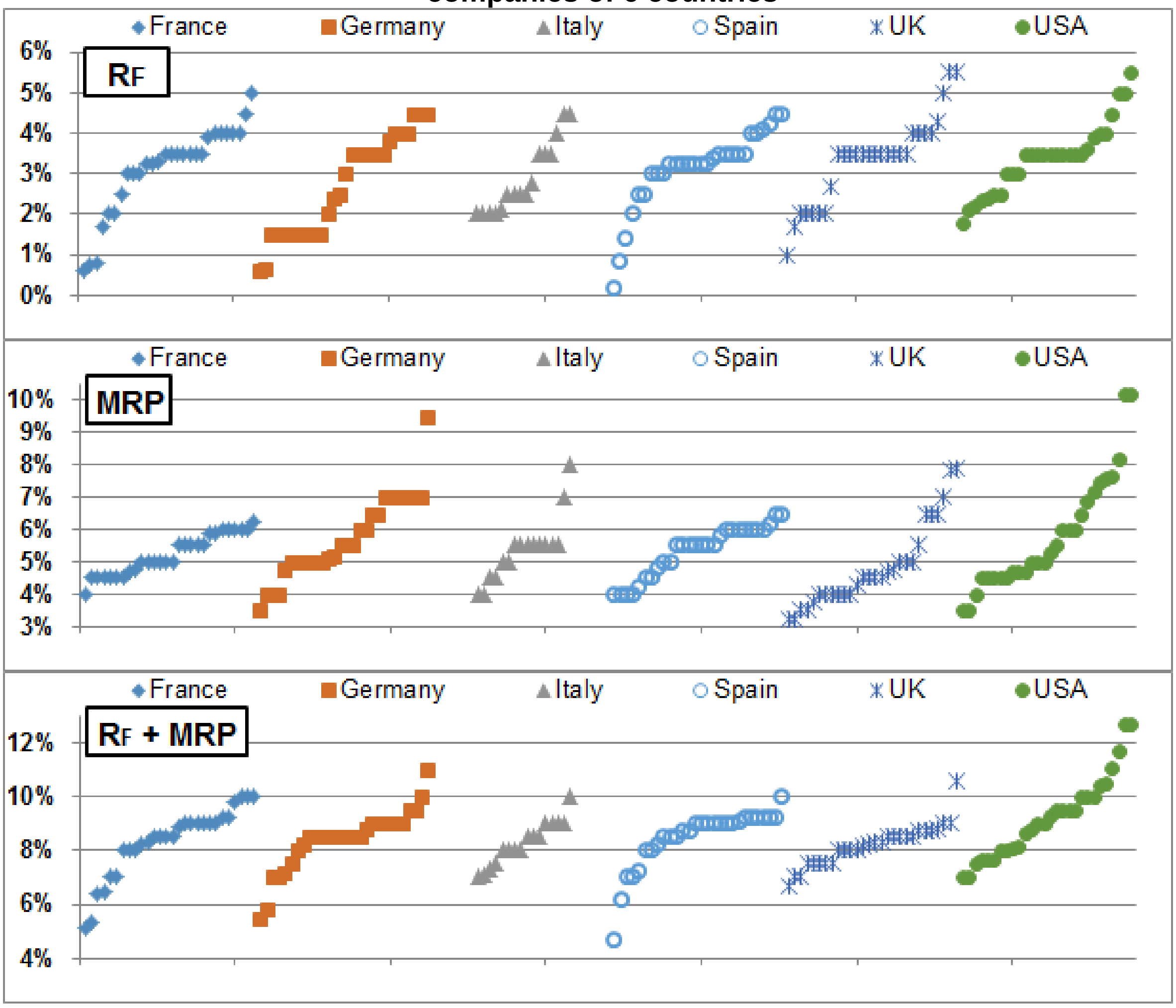


Figure 2. RF and MRP used in 2015 by 156 analysts in their valuations of companies of six countries

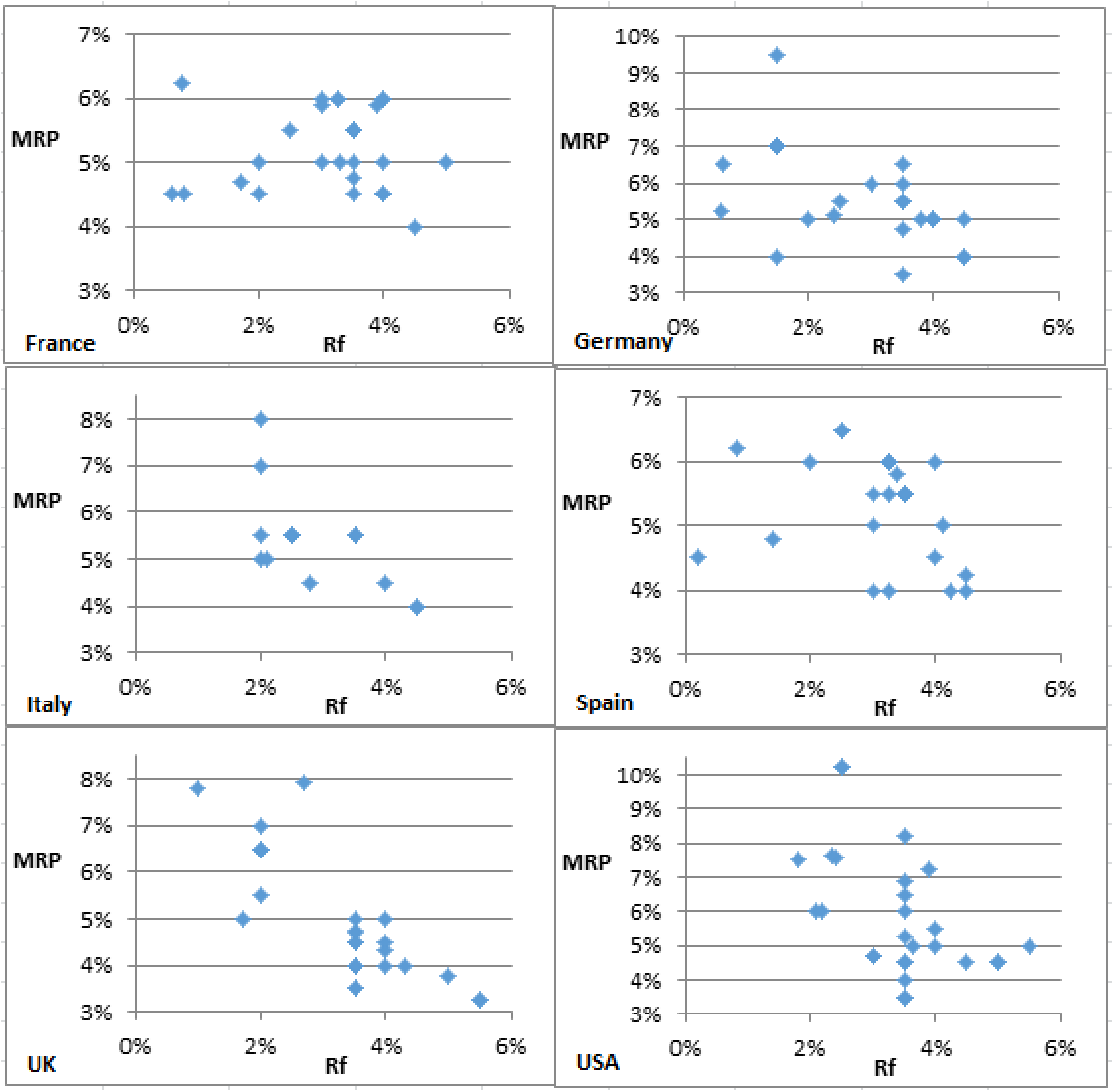


Table 1. RF and MRP used in 2015 by 156 analysts in their valuations

	RF				MRP				RF + MRP			
	average	max	min	StDev	average	max	min	StDev	average	max	min	StDev
France	3.1%	5.0%	0.6%	1.1%	5.2%	6.3%	4.0%	0.6%	8.3%	10.0%	5.1%	1.3%
Germany	2.7%	4.5%	0.6%	1.2%	5.7%	9.5%	3.5%	1.3%	8.4%	11.0%	5.5%	1.2%
Italy	2.8%	4.5%	2.0%	0.8%	5.4%	8.0%	4.0%	1.0%	8.3%	10.0%	7.0%	0.8%
Spain	3.1%	4.5%	0.2%	1.0%	5.3%	6.5%	4.0%	0.8%	8.5%	10.0%	4.7%	1.1%
UK	3.3%	5.5%	1.0%	1.1%	4.8%	7.9%	3.3%	1.3%	8.2%	10.6%	6.7%	0.8%
USA	3.4%	5.5%	1.8%	0.9%	5.8%	10.2%	3.5%	1.8%	9.2%	12.7%	7.0%	1.5%

	StDev / average		Correlation		
	RF	MRP	RF, RF + MRP	RF, MRP	MRP, RF + MRP
France	0.67	0.25	*87.5%*	*4.9%*	52.7%
Germany	0.62	0.28	41.4%	-59.3%	48.7%
Italy	0.73	0.28	31.3%	-56.2%	54.0%
Spain	0.55	0.27	*72.4%*	-24.2%	49.4%
UK	0.38	0.33	5.8%	-80.3%	54.8%
USA	0.46	0.25	0.7%	-49.7%	86.4%

The statistics of table 1 can be compared with the statistics of a survey that was conducted on April 2015 (see **Table 2).** It can be seen that:

- The (average RF) used by analysts is substantially higher than the (average RF) of the survey.
- The (average MRP) is not substantially different,
- The average (RF + MRP) used by analysts is substantially higher than the average (RF + MRP) of the survey.

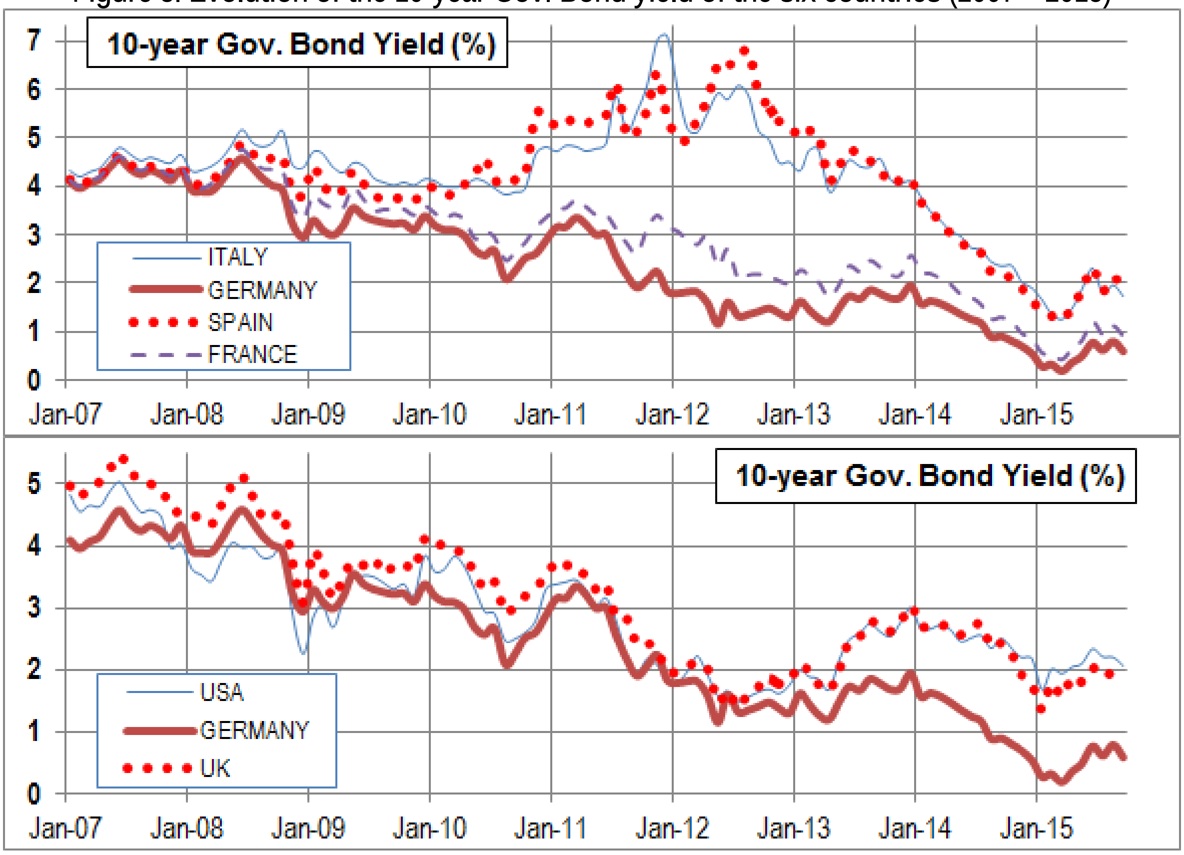
Table 2. RF and MRP used in 2015 according to a survey[1]

	RF				MRP				RF + MRP			
	average	max	min	StDev	average	max	min	StDev	average	max	min	StDev
France	1.5%	5.1%	0.0%	1.0%	5.6%	10.0%	2.0%	1.4%	7.2%	14.0%	4.0%	1.6%
Germany	1.3%	5.1%	-0.2%	0.8%	5.3%	11.3%	2.0%	1.5%	6.6%	14.2%	2.8%	1.7%
Italy	1.5%	5.0%	0.0%	1.1%	5.4%	10.0%	2.0%	1.5%	7.0%	14.0%	3.0%	2.1%
Spain	2.2%	7.0%	0.0%	1.2%	5.9%	12.0%	3.0%	1.6%	8.1%	15.7%	4.1%	2.0%
UK	2.1%	6.0%	0.4%	0.8%	5.2%	10.5%	1.3%	1.7%	7.2%	13.0%	3.0%	1.9%
USA	2.4%	8.0%	0.0%	1.1%	5.5%	15.0%	2.0%	1.4%	7.9%	22.0%	2.5%	1.7%

2. Evolution of the 10-year Government bonds yield for the six countries

The anomalous low yields on the Government bonds in 2014 and 2015 (see **Figure 3**) may have some influence on the results presented on the previous section. Figure 3 suggests three pairs of countries with RF moving quite close: Italy-Spain, Germany-France and US-UK.

Figure 3. Evolution of the 10-year Gov. Bond yield of the six countries (2007 – 2015)



[1] "RF and Market Risk Premium Used for 41 Countries in 2015: A Survey", http://ssrn.com/abstract=2598104

A comment about the **Quantitative Easing (QE)** implemented by the ECB in 2014, 2015... It is just a strange synonym for "print a lot of money (euros) and buy many, many bonds of the countries in the EU". By doing so, bond prices increase (and bond yields decrease) dramatically. Some people refer to this "QE" as "market abuse of the ECB", "market manipulation", "altering competitive markets", "expropriation of savings"… We agree with all this definitions: they are clearer than "QE".

3. Degrees of freedom of different analysts

A closer look at Exhibit 1 permits to find four different patterns of analyst houses. We can see that there companies with their analysts using (for companies of the same country):

a) The same RF and same MRP
b) The same RF and different MRP
c) Different RF and same MRP
d) Different RF and different MRP. Among these we find the analysts of Spanish companies that belong to Deutsche Bank (see table 3). Other companies where the analysts have a lot of freedom are Jefferies, Morgan Stanley, Natixis, Societe Generale (although not in the USA) and UBS.

Table 3. RF and MRP used in 2015 in reports done by analysts of Spanish companies that belong to Deutsche Bank.

	Company	RF	MRP	RF + MRP
24/02/2015	Amadeus	4.1%	5.0%	9.1%
25/02/2015	Gamesa	1.8%	4.0%	5.8%
26/02/2015	Grifols	3.5%	5.5%	9.0%
24/03/2015	Inditex	2.0%	6.0%	8.0%
26/05/2015	Red Eléctrica	1.8%	4.0%	5.8%
17/06/2015	Aena	3.4%	5.8%	9.2%
30/07/2015	Acerinox	2.5%	6.5%	9.0%

4. MRP in 2015 according to Damodaran

Damodaran does a strange calculation of the MRP in 2015 of 116 countries. **Table 4** contains the 57 countries with MRP smaller than 9%.

Table 4. MRP in 2015 according to Damodaran

Source: *Damodaran* http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/ctryprem.html

5.75%; Australia; Austria; Brunei; Canada; Denmark; Finland; **Germany**; Luxembourg; Netherlands; New Zealand; Norway; Singapore; Sweden; Switzerland; **United States**
6.35%; **France**; Hong Kong; **United Kingdom**
6.50%; Kuwait; Qatar; United Arab Emirates
6.65%; Belgium; Chile; China. Peoples' Rep.; Korea. Republic; Saudi Arabia; Taiwan
6.80%; Czech Republic; Estonia; Israel; Japan; Oman
7.03%; Botswana; Poland; Slovakia
7.55%; Malaysia; Malta; Mexico; Peru
8.15%; Ireland; Latvia; Lithuania; Thailand; Trinidad & Tob.
8.60%; Bahamas; Bahrain; Brazil; Bulgaria; Colombia; **Italy**; Kazakhstan; Panama; Philippines; Russia; South Africa; **Spain;** Uruguay

Exhibit 3 reproduces how Damodaran gets the MRP of table 4. We agree with Damodaran in many valuation issues, but we cannot agree with table 4 or with Exhibit 3 for two reasons:

1. It does not make much sense to calculate the MRP, something related to the perceived risk of shares of a market, using only a characteristic of the Government bonds of that market.
2. The comparison of the MRP of different countries does not agree with common sense. Does it make sense, for example, to affirm that in December 2014 the risk of investing in Italian

companies was higher than the risk of investing in companies from Estonia, Oman, Botswana, Poland, Malaysia, Malta, Mexico, Peru…?

All analysts in our sample used a MRP smaller than the suggested by Damodaran for Spain (8.6%), Italy (8.6%) and France (6.35%). 12 analysts used a MRP smaller (and 16 a MRP higher) than the suggested by Damodaran for USA and Germany (5.75%). 6 analysts used a MRP smaller (and 22 a MRP higher) than the suggested by Damodaran for UK (6.35%).

5. MRP and RF. Where they come from?[2]

The valuation of companies using discounted cash flows is an extension of the valuation of Government bonds. The Value of a Government bond (VGB) is the present value of the cash flows promised in the bond (CFgb) using the so called "**risk-free rate**" (R_F):

Value of a Government bond = VGB = PV (CFgb; R_F) (1)

The risk-free rate (R_F) is the **required return** to Government bonds (when the ECB does not manipulate the market)

Valuation of the Debt. The **Debt cash flows** (CFd) are interest payments and repayments of debt (∇N).[3] CFd = Interest + ∇N (2)

As the (CFd) promised by a company are usually riskier[4] than the cash flows promised by the Government (CFgb), the **required return to Debt** (Kd) is usually higher than the risk-free rate (R_F)

Required return to debt = Kd = R_F + RPd (debt risk premium) (3)

The **debt risk premium** (RPd) depends on the perceived risk on the Debt (expectations of getting less money than the promised Debt cash flows) by every investor. Applying Equation (1) to the Debt of the company, we get: Value of debt = D = PV (CFd; Kd) (4)

Cash	Debt (N) *Bank debt, bonds…*
Working Capital Requirements (WCR)	
Net Fixed Assets (NFA)	Book value of Equity (Ebv) *Shares*

Debt Cash Flow (CFd): money (cash) that goes from the Cash of the company to the pockets of bondholders

Equity Cash Flow (ECF): money (cash) that goes from the Cash of the company to the pockets of shareholders

Valuation of the shares. A share of a company is a piece of paper that, contrary to debt, has not dates nor amounts that will receive its owner, the shareholder. We need, first, to estimate the expected cash flows for the owners of the shares in the following years, named **Equity Cash Flows** (ECF). A usual way of estimating the ECF is to start with the expected Balance Sheets and P&Ls. Equation (5) is the basic accounting identity: assets are equal to liabilities and equity:

Cash + WCR + NFA = N + Ebv (5)

Equation (6) is the annual change of Equation (5). The increase of the cash of the company before giving anything to the shareholders will be divided between the **ECF** and the increase of cash (ΔCash) decided by the managers: ECF + ΔCash + ΔWCR + ΔNFA = ΔN + ΔEbv (6)

If the (ΔEbv) is due only to the Profit after Tax (PAT) of the year, then[5]:

ECF = PAT - ΔWCR - ΔNFA + ΔN - ΔCash (7)

---

[2] *Source: "Cash flow discounting: fundamental relationships and unnecessary complications"*, http://ssrn.com/abstract=2117765

[3] If the company does not repay debt (∇N) but increases its debt (ΔN), Equation (2) would be CFd = Interest - ΔN

[4] The risk of the debt is the probability that the company will not pay some of the promised cash flows. Risk-free debt means that we believe that the issuer will pay all promised cash flows for sure.

[5] As NFA = GFA (gross fixed assets) – depreciation, equation (7) can be written:
ECF = PAT + depreciation - ΔNOF - ΔGFA + ΔN - ΔCash

As the expected ECF are riskier than the cash flows promised by the Government bonds (CFgb) and also riskier than the cash flows promised by the Debt of the company (CFd), the **required return to equity (shares)** (Ke) is higher than risk-free rate (R_F) and also higher than the required return to Debt (Kd): Ke = R_F + RPs (shares risk premium) (8)

The so-called **shares risk premium** (RPs) depends on the estimated (expected) risk of the expected equity cash flows (ECF). Obviously, this parameter depends on the expectations of each investor. Applying Equation (1) to the equity (the shares of the company), we get:
Value of the shares (equity value) = E = PV (ECF; Ke) (9)

With equations (2) to (9) we can value any company. But, as equations (2) to (9) are relatively easy to understand, it is quite common to complicate the valuation with new 'concepts' and new equations. With these unnecessary complications, the valuation becomes more difficult to understand and acquires a more "scientific", "serious", "intriguing", "impenetrable"… appearance.

'Invention' of the beta and the market risk premium (MRP). It consists in calculating RPs (shares risk premium) as a product: RPs = β MRP (10)

The MRP (**market risk premium**) is the "shares risk premium" applied to the whole market (or to a portfolio with shares of most of the companies traded in the stock markets). The market risk Premium (MRP) is the answer to the following question: Knowing that your money invested in long-term Government bonds will provide you a return of R_F% almost for sure, which additional return you require to another investment (in a portfolio with shares of most of the companies with shares traded in the financial markets) for feeling compensated for the extra risk that you assume?

The "market risk premium" is also called "*equity premium*", "*equity risk premium*", "*market premium*" and "*risk premium*".

The β (**beta**) is a specific parameter for each company. β = 0 corresponds to Government bonds (no risk) and β = 1 to an investment with a perceived risk similar to that of the market.

With the 'invention' of the beta, Equation (8) becomes equation (11)

Ke = R_F + β MRP (11)

6. Two common errors about β and MRP

First error: To maintain that the β may be calculated with a regression of historical data

This lack of common sense consists first, in assuming that "the market" assigns a beta to every company and second, in maintaining that the levered beta may be calculated with a regression of historical data. According to the followers of this new "complication", the beta has nothing to do with the expectations of risk, the experience of the valuator… but rather every investor should use the same beta: the calculated beta. You can get that beta running a regression of the past returns of the company this the returns of some market index.

We show that it is an enormous error to use calculated betas (see *Are Calculated Betas Worth for Anything?* http://ssrn.com/abstract=504565). First, because it is almost impossible to calculate a meaningful beta because historical betas change dramatically from one day to the next. Second, because very often we cannot say with a relevant statistical confidence that the beta of one company is smaller or bigger than the beta of another. Third, because historical betas do not make much sense in many cases: high-risk companies very often have smaller historical betas than low-risk companies do. Fourth, because historical betas depend very much on which index we use to calculate them.

Some authors and companies publish calculated betas. For example, Damodaran publish industry betas in http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/Betas.html

Second error: To maintain that "the market" has "a MRP" and that it is possible to estimate it

This new "complication" consists in assuming that "the market" has a MRP (market risk premium). Then, the MRP would be a parameter "of the market" and not a parameter that is different for different investors.

We reviewed[6] 150 textbooks on corporate finance and valuation written by authors such as Brealey, Myers, Copeland, Damodaran, Merton, Ross, Bruner, Bodie, Penman, Arzac… and finds that their recommendations regarding the equity premium range from 3% to 10%, and that 51 books use different equity premia in various pages. Some confusion arises from not distinguishing among the concepts that the phrase equity premium designates: the Historical, the Expected and the Required equity premium (incremental return of a diversified portfolio over the risk-free rate required by an investor). 129 out of the 150 textbooks identify Expected and Required equity premium and 82 books identify Expected and Historical equity premium.

We maintain that the CAPM is an absurd model[7].

7. Expected, Required and Historical Market Risk Premium: different concepts

Fernandez and F.Acín (2015)[8] claim and show that Expected Return and Required Return are two very different concepts.

We also claim[9] that the term "MRP" is used to designate three different concepts:

1. **Historical** equity premium (HEP): historical differential return of the stock market over treasuries.
2. **Expected** equity premium (EEP): expected differential return of the stock market over treasuries.
3. **Required** equity premium (REP): incremental return of a diversified portfolio (the market) over the risk-free rate required by an investor. It is used for calculating the required return to equity.

The three concepts (HEP, REP, EEP) designate different realities. The **HEP** is easy to calculate and is equal for all investors, provided they use the same period, the same market index, the same risk-free instrument and the same average (arithmetic or geometric). However, the **EEP** and the **REP** may be different for different investors and are not observable.

The **HEP** is the historical average differential return of the market portfolio over the risk-free debt. The most widely cited sources are Ibbotson Associates and Dimson *et al.* (2007).

Numerous papers and books assert or imply that there is a "market" EEP. However, it is obvious that investors and professors do not share "homogeneous expectations" and have different assessments of the **EEP**. As Brealey et al. (2005, page 154) affirm, *"Do not trust anyone who claims to know what returns investors expect".*

The **REP** is the answer to the following question: What incremental return do I require for investing in a diversified portfolio of shares over the risk-free rate? It is a crucial parameter because the REP is the key to determining the company's required return to equity and the WACC. Different companies may use, and in fact do use, different **REPs**.

For a particular investor, the **EEP** is not necessary equal to the REP (unless he considers that the market price is equal to the value of the shares). Obviously, an investor will hold a diversified portfolio of shares if his EEP is higher (or equal) than his REP and will not hold it otherwise.

We can find out the REP and the EEP of an investor by asking him, although for many investors the REP is not an explicit parameter but, rather, it is implicit in the price they are prepared to pay for the shares. However, it is not possible to determine the REP for the market as a whole, because it does not exist: even if we knew the REPs of all the investors in the market, it would be meaningless to talk of a REP for the market as a whole. There is a distribution of REPs and we can only say that some percentage of investors have REPs contained in a range. The average of that distribution cannot be interpreted as the REP of the market nor as the REP of a representative investor.

[6] "The Equity Premium in 150 Textbooks", http://ssrn.com/abstract=1473225
[7] "CAPM: an absurd model", http://ssrn.com/abstract=2505597
[8] "Expected and Required Returns: Very Different Concepts", http://ssrn.com/abstract=2591319
[9] "Equity premium: historical, expected, required and implied", http://ssrn.com/abstract=933070

8. Conclusion

We look at the RF and the MRP used by analysts in 2015 to value companies of six countries. The dispersion of both, the RF and the MRP used, is huge, and the most unexpected result is that the dispersion is higher for the RF than for the MRP.

Most of the analyst use a RF higher than the yield of the 10-year Government bonds. A reason for it and for the huge dispersion may be the activity of the European Central Bank (ECB). The risk-free rate (RF) is the required return to Government bonds when nobody (not even the ECB) manipulates the market. A question arises: May we consider the Quantitative Easing (QE) implemented by the ECB in 2014, 2015... "market abuse", "market manipulation", a way of "altering competitive markets"…?

We also find that some analysts have more freedom than others.

The data permits other comparisons. For example: Does it make sense that the average MRP used for Germany is higher than the average MRP used for France, Italy, Spain or the UK?

References

Claus, J.J. and J.K. Thomas (2001), "Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets," *Journal of Finance*. 55, (5), pp. 1629-66.

Damodaran, A. (2008), "Equity Risk Premiums (ERP): Determinants, Estimation and Implications", Working Paper.

Dimson, E., P. Marsh and M. Staunton (2007), "The Worldwide Equity Premium: A Smaller Puzzle," in *Handbook of investments: Equity risk premium*, R. Mehra, Elsevier.

Fernandez, P. (2007), "Equity Premium: Historical, Expected, Required and Implied", http://ssrn.com/abstract=933070

Fernandez, P. (2009a), "Market Risk Premium Used in 2008 by Professors: A Survey with 1,400 Answers", http://ssrn.com/abstract=1344209

Fernandez, P. (2009b), "The Equity Premium in 150 Textbooks", http://ssrn.com/abstract=1473225

Fernandez, P., J. Aguirremalloa and H. Liechtenstein (2009), "The Equity Premium Puzzle: High Required Premium, Undervaluation and Self Fulfilling Prophecy". IESE Business School WP. http://ssrn.com/abstract=1274816

Fernandez, P. and J. del Campo (2010a), "Market Risk Premium used in 2010 by Analysts and Companies: a survey with 2,400 answers", downloadable in http://ssrn.com/abstract=1609563

Fernandez, P. and J. del Campo (2010b), "Market Risk Premium Used in 2010 by Professors: A Survey with 1,500 Answers", downloadable in http://ssrn.com/abstract=1606563

Fernandez, P., J. Aguirreamalloa and L. Corres (2011a), "US Market Risk Premium Used in 2011 by Professors, Analysts and Companies: A Survey with 5.731 Answers", downloadable in http://ssrn.com/abstract=1805852

Fernandez, P., J. Aguirreamalloa and L. Corres (2011b), "The Equity Premium in Spain: Survey 2011 (in Spanish)", downloadable in http://ssrn.com/abstract=1822422

Fernandez, P., J. Aguirreamalloa and L. Corres (2011c), "Market Risk Premium Used in 56 Countries in 2011: A Survey with 6,014 Answers", downloadable in http://ssrn.com/abstract=1822182

Fernandez, P., J. Aguirreamalloa and P. Linares (2014), "Market Risk Premium and Risk Free Rate Used for 51 Countries in 2013: A Survey with 6,237 Answers", downloadable in http://ssrn.com/abstract=914160

Fernandez, P., J. Aguirreamalloa and L. Corres (2012), "Market Risk Premium Used in 82 Countries in 2012: A Survey with 7,192 Answers", downloadable in http://ssrn.com/abstract=2084213

Fernandez, P. and I. F. Acín (2015), "Expected and Required Returns: Very Different Concepts", downloadable in http://ssrn.com/abstract=2591319

Fernandez, P., P. Linares and I. F. Acín (2014), "Market Risk Premium Used in 88 Countries in 2014: A Survey with 8,228 Answers", downloadable in http://ssrn.com/abstract=2450452

Exhibit 1. Analyst reports analyzed in this document

US COMPANIES

Date	Analyst	Company	Rf	MRP	Rf+MRP
02/02/2015	HSBC	Chevron	3,5%	6,0%	9,5%
06/08/2015	JP Morgan	Tesla	1,8%	7,5%	9,3%
22/07/2015	Jefferies	Microsoft	3,5%	8,2%	11,7%
06/10/2015	Deutsche Bank	Pepsico	3,5%	4,0%	7,5%
23/09/2015	Morgan Stanley	JP Morgan Chase	5,0%	4,5%	9,5%
14/07/2015	RBC Capital	JP Morgan Chase	2,4%	7,6%	10,0%
16/04/2015	Societe Generale	JP Morgan Chase	3,0%	4,7%	7,7%
17/07/2015	RBC Capital	Citigroup	2,4%	7,7%	10,0%
17/07/2015	Morgan Stanley	Citigroup	5,0%	4,5%	9,5%
18/09/2015	Barclays	Altria (ex Philip Morris)	3,5%	4,5%	8,0%
12/05/2015	HSBC	Verizon	3,5%	3,5%	7,0%
22/07/2015	Cowen and Company	Verizon	2,5%	10,2%	12,7%
18/02/2015	Piper Jaffray	Amazon	2,1%	6,0%	8,1%
15/04/2015	Morgan Stanley	Wells Fargo	4,5%	4,5%	9,0%
09/03/2015	Societe Generale	Wells Fargo	3,0%	4,7%	7,7%
13/10/2015	Piper Jaffray	Johnson & Johnson	3,6%	5,0%	8,6%
30/07/2015	Brean Capital	Facebook	3,5%	6,5%	10,0%
17/09/2015	Wedbush Securities	Oracle	2,2%	6,0%	8,2%
17/09/2015	Jefferies	Oracle	3,5%	5,3%	8,8%
15/06/2015	UBS	Walt Disney	3,5%	6,9%	10,4%
12/05/2015	Deutsche Bank	Walt Disney	4,0%	5,0%	9,0%
24/07/2015	Cowen and Company	AT&T	2,5%	10,2%	12,7%
03/06/2015	Jefferies	AT&T	4,0%	5,5%	9,5%
23/04/2015	HSBC	AT&T	3,5%	3,5%	7,0%
19/06/2015	Societe Generale	Bank of America	3,0%	4,7%	7,7%
13/08/2015	Deutsche Bank	Cisco Systems	5,5%	5,0%	10,5%
12/10/2015	UBS	Anheuser-Busch	3,9%	7,2%	11,1%
08/10/2015	Barclays	Anheuser-Busch	3,5%	4,5%	8,0%

	Rf	MRP	Rf+MRP
Average	3,4%	5,8%	9,2%
Max	5,5%	10,2%	12,7%
min	1,8%	3,5%	7,0%
St. Dev	0,9%	1,8%	1,5%

GERMAN COMPANIES

Date	Analyst	Company	Rf	MRP	Rf+MRP
30/09/2015	HSBC	BASF	0,6%	5,2%	5,8%
11/03/2015	Warburg Research Gmbh	BASF	1,5%	7,0%	8,5%
07/05/2015	Warburg Research Gmbh	BMW	1,5%	7,0%	8,5%
29/04/2015	Warburg Research Gmbh	DAIMLER	1,5%	7,0%	8,5%
14/04/2015	Natixis	Deutsche Bank	2,4%	5,1%	7,5%
03/08/2015	Esn/Equinet Bank	Deutsche Telekom	4,5%	4,0%	8,5%
13/05/2015	Warburg Research Gmbh	Deutsche Telekom	1,5%	7,0%	8,5%
03/12/2014	Kepler Cheuvreux	Siemens	3,8%	5,0%	8,8%
30/04/2015	Warburg Research Gmbh	Bayer	1,5%	7,0%	8,5%
18/02/2015	Commerzbank	Bayer	1,5%	4,0%	5,5%
02/10/2015	HSBC	Merck	3,5%	3,5%	7,0%
12/08/2015	Deutsche Bank	Merck	4,0%	5,0%	9,0%
19/05/2015	Warburg Research Gmbh	Merck	1,5%	7,0%	8,5%
02/10/2015	HSBC	Lufthansa	3,5%	5,5%	9,0%
28/09/2015	ESN/Equinet Bank	Lufthansa	4,5%	5,0%	9,5%
18/09/2015	Macquarie Research	E.ON	4,0%	5,0%	9,0%
10/09/2015	HSBC	E.ON	3,5%	5,5%	9,0%
12/03/2015	Warburg Research Gmbh	E.ON	1,5%	7,0%	8,5%
07/10/2015	ESN/Equinet Bank	Volkswagen	4,5%	4,0%	8,5%
02/10/2015	Baader Helvea Equity Res.	Volkswagen	2,0%	5,0%	7,0%
30/03/2015	Redburn / Automotive	Volkswagen	3,5%	6,0%	9,5%
12/08/2015	HSBC	Henkel	3,5%	4,8%	8,3%
09/03/2015	Jefferies	Henkel	2,5%	5,5%	8,0%
26/08/2015	JP Morgan	ThyssenKrupp	0,7%	6,5%	7,2%

31/03/2015	Societe Generale	ThyssenKrupp	1,5%	9,5%	11,0%
22/04/2015	HSBC	SAP	3,0%	6,0%	9,0%
21/04/2015	Jefferies	SAP	3,5%	6,5%	10,0%
20/03/2015	Deutsche Bank	SAP	4,0%	5,0%	9,0%

Average	2,7%	5,7%	8,4%
Max	4,5%	9,5%	11,0%
min	0,6%	3,5%	5,5%
St. Dev	1,2%	1,3%	1,2%

FRENCH COMPANIES

Date	Analyst	Company	Rf	MRP	Rf+MRP
10/04/2015	ESN	AXA	4,5%	4,0%	8,5%
09/09/2015	BPI	BNP Paribas	3,3%	6,0%	9,3%
27/04/2015	Kepler Cheuvreux	BNP Paribas	4,0%	4,5%	8,5%
15/09/2015	HSBC	Carrefour	3,5%	5,5%	9,0%
04/09/2015	Natixis	Credit Agricole	3,0%	6,0%	9,0%
23/09/2015	HSBC	EDF	3,5%	5,5%	9,0%
22/06/2015	Macquarie Research	EDF	3,3%	5,0%	8,3%
07/10/2015	RBC Capital Markets	LVMH	3,0%	5,0%	8,0%
15/09/2015	Societe Generale	LVMH	2,0%	5,0%	7,0%
07/08/2015	HSBC	LVMH	3,5%	5,5%	9,0%
14/09/2015	Raymond James	L´Oreal	3,5%	5,0%	8,5%
31/07/2015	HSBC	L´Oreal	3,5%	4,8%	8,3%
27/07/2015	Deutsche Bank	Saint Gobain	4,0%	6,0%	10,0%
07/10/2015	Liberum	Total	4,0%	5,0%	9,0%
16/09/2015	BPI	Societe Generale	3,3%	6,0%	9,3%
30/09/2015	UBS	Pernod Ricard	3,0%	5,9%	8,9%
28/08/2015	Raymond James	Pernod Ricard	3,5%	5,5%	9,0%
28/08/2015	Barclays	Pernod Ricard	3,5%	4,5%	8,0%
18/05/2015	UBS	Essilor	2,0%	4,5%	6,5%
06/04/2015	UBS	Legrand	0,6%	4,5%	5,1%
17/12/2015	Deutsche Bank	Lafarge	3,9%	5,9%	9,8%
14/07/2015	Liberum	Technip	5,0%	5,0%	10,0%
19/02/2015	Raymond James	Technip	4,0%	6,0%	10,0%
02/10/2015	UBS	Schneider Electric	0,8%	4,5%	5,3%
12/06/2015	Deutsche Bank	Air Liquide	4,0%	4,5%	8,5%
20/02/2015	Liberum	Air Liquide	2,5%	5,5%	8,0%
23/09/2015	Societe Generale	Unibail-Rodamco	1,7%	4,7%	6,4%
23/07/2015	Natixis	Unibail-Rodamco	0,8%	6,3%	7,0%

Average	3,1%	5,2%	8,3%
Max	5,0%	6,3%	10,0%
min	0,6%	4,0%	5,1%
St. Dev	1,1%	0,6%	1,3%

UK COMPANIES

Date	Analyst	Company	Rf	MRP	Rf+MRP
30/07/2015	HSBC	GlaxoSmithKline	3,5%	3,5%	7,0%
24/07/2015	Renaissance Capital	Anglo American Minas	4,0%	5,0%	9,0%
09/09/2015	HSBC	Astra Zeneca	3,5%	3,5%	7,0%
31/07/2015	Barclays	Diageo	3,5%	4,5%	8,0%
27/01/2015	Deutsche Bank	Intercont.Hotels Group	4,0%	4,0%	8,0%
12/01/2015	BNP Paribas	British Americ. Tobacco	2,7%	7,9%	10,6%
27/07/2015	HSBC	Reckitt Benckiser	3,5%	4,8%	8,3%
29/06/2015	Kepler Cheuvreux	Reckitt Benckiser	1,0%	7,8%	8,8%
14/01/2015	Raymond James	Reckitt Benckiser	3,5%	5,0%	8,5%
24/06/2015	Barclays	Next	3,5%	4,5%	8,0%
10/07/2015	Morgan Stanley	Rolls Royce	4,0%	4,5%	8,5%
08/07/2015	Kepler Cheuvreux	Rolls Royce	2,0%	5,5%	7,5%
18/08/2015	UBS	Aviva	1,7%	5,0%	6,7%
16/07/2015	Credit Suisse	Michael Page	2,0%	6,5%	8,5%
16/04/2015	Kepler Cheuvreux	Michael Page	2,0%	7,0%	9,0%
08/10/2015	Morgan Stanley	Sports Direct	5,5%	3,3%	8,8%

26/01/2015	Morgan Stanley	Dixons Carphone	5,0%	3,8%	8,8%
03/09/2015	HSBC	International Airlines Gr	3,5%	4,0%	7,5%
26/06/2015	HSBC	Aberdeen Assets Manag.	3,5%	4,0%	7,5%
10/09/2015	Peel Hunt	Admiral Group	4,3%	4,0%	8,3%
01/07/2015	SEB	4GS	3,5%	4,0%	7,5%
12/01/2015	Credit Suisse	4GS	2,0%	6,5%	8,5%
30/09/2015	HSBC	Sainsbury	3,5%	4,0%	7,5%
17/03/2015	Morgan Stanley	Sainsbury	5,5%	3,3%	8,8%
26/03/2015	Credit Suisse	Serco	2,0%	6,5%	8,5%
24/04/2015	HSBC	WPP	3,5%	4,7%	8,2%
22/06/2015	Barclays	Imperial Tobacco	3,5%	4,5%	8,0%
15/06/2015	Deutsche Bank	Imperial Tobacco	4,0%	4,3%	8,3%

Average	3,3%	4,8%	8,2%
Max	5,5%	7,9%	10,6%
min	1,0%	3,3%	6,7%
St. Dev	1,1%	1,3%	0,8%

ITALIAN COMPANIES

Date	Analyst	Company	Rf	MRP	Rf+MRP
20/01/2015	Barclays	Luxottica	4,0%	4,5%	8,5%
22/01/2015	Banca IMI	Natural Gas	2,5%	5,5%	8,0%
22/01/2015	Banca IMI	R&M	2,5%	5,5%	8,0%
22/01/2015	Banca IMI	Chem	2,5%	5,5%	8,0%
22/01/2015	Banca IMI	PowerGen	2,5%	5,5%	8,0%
27/01/2015	UBS	Finmeccanica	2,0%	7,0%	9,0%
30/01/2015	Deutsche Bank	Grupo Mediolanum	2,0%	8,0%	10,0%
05/02/2015	ICBPI	Mediaset	2,0%	5,0%	7,0%
17/02/2015	ICBPI	Pirelli	2,1%	5,0%	7,1%
12/03/2015	HSBC	Telecom Italia	3,5%	5,5%	9,0%
13/03/2015	Banca IMI	Luxottica	2,0%	5,5%	7,5%
16/03/2015	Morgan Stanley	Enel	4-5%	4,0%	8,5%
19/03/2015	ESN	Eni	4,5%	4,0%	8,5%
20/03/2015	HSBC	Enel	3,5%	5,5%	9,0%
23/03/2015	Barclays	Luxottica	2,8%	4,5%	7,3%
25/03/2015	HSBC	Mediaset	3,5%	5,5%	9,0%

Average	2,8%	5,4%	8,3%
Max	4,5%	8,0%	10,0%
Min	2,0%	4,0%	7,0%
St Dev	0,8%	1,0%	0,8%

SPANISH COMPANIES

Date	Analyst	Company	Rf	MRP	Rf+MRP
20/04/2015	HSBC	Telefónica	3,5%	5,5%	9,0%
18/09/2015	HSBC	Inditex	3,5%	5,5%	9,0%
04/09/2015	Societe Generale	Inditex	1,4%	4,8%	6,2%
03/06/2015	UBS	Inditex	0,2%	4,5%	4,7%
26/05/2015	Morgan Stanley	Inditex	3,3%	5,5%	8,8%
24/03/2015	Deutsche Bank	Inditex	2,0%	6,0%	8,0%
08/09/2015	BPI	OHL	3,3%	6,0%	9,3%
17/09/2015	UBS	Banco Sabadell	3,5%	5,5%	9,0%
30/07/2015	Deutsche Bank	Acerinox	2,5%	6,5%	9,0%
19/01/2015	BPI	Acerinox	3,3%	6,0%	9,3%
03/07/2015	HSBC	Enagás	3,5%	5,5%	9,0%
29/06/2015	BPI	Enagás	3,3%	6,0%	9,3%
27/07/2015	HSBC	Iberdrola	3,0%	5,5%	8,5%
18/02/2015	Morgan Stanley	Iberdrola	3,0%	4,0%	7,0%
06/10/2015	HSBC	Gas Natural	3,5%	5,5%	9,0%
16/09/2015	Morgan Stanley	Gas Natural	4,5%	4,0%	8,5%
03/07/2015	Santander	Gamesa	4,5%	4,3%	8,8%
13/08/2015	Santander	Amadeus	4,3%	4,0%	8,3%
13/07/2015	JP Morgan	Amadeus	0,8%	6,2%	7,0%
24/02/2015	Deutsche Bank	Amadeus	4,1%	5,0%	9,1%

29/06/2015	BPI	Prosegur	3,3%	6,0%	9,3%
14/05/2015	Santander	Abertis	3,3%	4,0%	7,3%
10/09/2015	UBS	Aena	3,0%	5,0%	8,0%
11/08/2015	Societe Generale	Aena	4,0%	6,0%	10,0%
17/06/2015	Deutsche Bank	Aena	3,4%	5,8%	9,2%
14/08/2015	RBC Capital Markets	Endesa	2,5%	6,5%	9,0%
06/07/2015	ESN/Beka Finance	Mapfre	4,0%	4,5%	8,5%
08/09/2015	BPI	FCC	3,3%	6,0%	9,3%

Average	3,1%	5,3%	8,5%
Max	4,5%	6,5%	10,0%
min	0,2%	4,0%	4,7%
St. Dev	1,0%	0,8%	1,1%

Exhibit *2.* Details of some valuation reports

20/04/2015 – HSBC – Telefónica

A risk-free rate of 3.5%, an equity risk premium of 5.5% and an asset beta of 1.0 (vs 1.1 previously to reflect the normalising Spanish macro), leading to a WACC of 7.8% to discount the FCF of

18/09/2015 – HSBC – Inditex risk-free rate of 3.5%, an equity risk premium of 5.5%, a beta of 0.8

03/06/2015 – UBS – Inditex		08/09/2015 – BPI – OHL		04/9/2015 – Societe Generale - Inditex	
WACC		Rf	3.25%	Risk free rate	1.4%
Risk free rate	0.20%	Beta of Equity	0.9	Equity risk premium	4.8%
ERP	4.5%	Mkt Premium	6.0%		

26/05/2015 – Morgan Stanley – Inditex

- A cost of equity of 8.8% and a WACC of 8.8% (calculated on the basis of a risk-free rate of return of 5.5%, an equity risk premium of 3.25% and a beta of 1.0)

24/03/2015 – Deutsche Bank – Inditex

RFR 2%, ERP of 6%, beta of 1.0, and terminal nominal growth rate of 2.5%.

17/09/2015 – UBS – Banco Sabadell		19/01/2015 – BPI – Acerinox		29/06/2015 – BPI – Enagás	
Long Term Risk Free Rate	3.50%	Rf	3.25%	Rf	3.25%
Equity Risk Premium	5.5%	Beta Equity	1.09	CRP	0.35%
		Market Premium	6.0%	Be	0.88
				Mkt premium	6%

30/07/2015 – Deutsche Bank – Acerinox

growth rate (to reflect limited growth prospects), a risk-free rate of 2.5%, an equity risk premium of 6.5%, and a 1.1 industry beta. Risks: demand-price

03/07/2015 – HSBC – Enagás

growth rate of 1%. Our WACC of 5.2% is derived from cost of equity of 7.2% (risk free rate of 3.5%, country risk premium of 5.5%, and beta of 0.7) and cost of debt of 3.7% (pre-tax).

27/07/2015 – HSBC – Iberdrola

Our DCF is based on WACC of 5.8% (unchanged) derived from cost of equity of 7.7% (risk free rate of 3.0%, country risk premium of 5.5%, and beta of 0.85) and pre-tax cost of debt of 4.2%. Our terminal

18/02/2015 – Morgan Stanley – Iberdrola

rate case. We use a different WACC, on a 3% risk-free
rate and an equity risk premium of 4.0%, depending on

06/10/2015 – HSBC – Gas Natural (risk free rate of 3.5%, country risk premium of 5.5%, and beta of 0.9)

16/09/2015 – Morgan Stanley – Gas Natural

main businesses (gas and electricity distribution and gas supply). We use a different WACC, on a 4.5% risk-free
rate and an equity risk premium of 4.0%, depending on the risk and cost of debt of each business, ranging from

13/8/2015 – Santander – Amadeus		13/07/2015 – JP Morgan – Amadeus		03/07/2015 – Santander – Gamesa	
Risk-free rate	4.25%	Risk free rate	0.83%	Risk free rate	4.5%
Market premium	**4.00%**	Equity risk premium	6.2%	Equity risk premium	4.25%

24/02/2015 – Deutsche Bank – Amadeus

of 8.4%, with a 4.1% risk free rate, 5% equity risk premium, and beta of 1. Our
3% terminal growth rate is slightly below the 5% long-term air traffic growth

29/6/2015 – BPI – Prosegur		14/05/2015 – Santander – Abertis		10/09/2015 – UBS – Aena	
Rf	3.25%	Risk free rate	3.25%	Equity risk premium	5.0%
CRP	0.3%	Beta adjusted	0.700	Risk free rate	3.0%
Be	0.8	Net debt	11,943		
MRP	6.0%	Market cap	15,000		
		D	0.44		
		E	0.56		
		D/E	0.80		
		Beta levered	1.10		
		MRP	4.00%		

11/08/2015 – Societe Generale – Aena

parameters: WACC 6.5%, terminal growth 2%, cost of equity 10% (risk-free rate 4%, risk
premium 6%, beta 1), cost of debt 4%. The stock is trading close to our target price with a

14/08/2015 – RBC Capital Markets – Endesa

Our WACC which has nudged up from 7.2% is calculated using a risk free rate of 2.5%, an
equity risk premium of 6.5%, and a geared beta of 1.0.

06/07/2015 – ESN/Beka Finance – Mapfre

For the implied P/NAV ratio we used the following variables: i) 50% pay-out; ii) 4.5% risk
premium; iii) 4.0% risk-free rate; and iv) 1.15x beta resulting in CoE 9.18%.

08/09/2015 – BPI – FCC		17/06/2015 – Deutsche Bank – Aena	
Rf	3.25%	Risk-free rate (Spain 30-year)	3.4%
CRP [1]	1.5%	Market premium	5.8%
Beta Equity	0.95	Beta	1.10
Mkt Prem.	6.0%	Risk premium	6.3%

Exhibit 3. MRP according to Damodaran

Source: Damodaran http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/ctryprem.html

In Damodarans' page, we can read for Italy:

Country	Italy
Moody's sovereign rating	Baa2
S&P sovereign rating	BBB-
CDS spread	2,34%
Excess CDS spread (over US CDS)	2,03%
Country Risk Premium (Rating)	2,85%
Equity Risk Premium (Rating)	**8,60%**

This table summarizes the latest **bond ratings** and appropriate **default spreads** for different countries. While you can use these numbers as rough estimates of country risk premiums, you may want to modify the premia to reflect the additional risk of equity markets. To estimate the long term country equity risk premium, I start with a default spread, which I obtain in one of two ways:
(1) I use the local currency sovereign rating (from Moody's: www.moodys.com) and estimate the default spread for that rating (based upon traded country bonds) over a default free government bond rate. For countries without a Moody's rating but with an S&P rating, I use the Moody's equivalent of the S&P rating. To get the default spreads by sovereign rating, I use the CDS spreads and compute the average CDS spread by rating. Using that number as a basis, I extrapolate for those ratings for which I have no CDS spreads.
(2) I start with the CDS spread for the country, if one is available and subtract out the US CDS spread, since my mature market premium is derived from the US market. That difference becomes the country spread. For the few countries that have CDS spreads that are lower than the US, I will get a negative number.
You can add just this default spread to the mature market premium to arrive at the total equity risk premium. I add an additional step. In the short term especially, the equity country risk premium is likely to be greater than the country's default spread. You can estimate an adjusted country risk premium by multiplying the default spread by the relative equity market volatility for that market (Std dev in country equity market/Std dev in country bond). I have used the **emerging market average of 1.5** (equity markets are about 1.5 times more volatile than bond markets) to estimate country risk premium. I have added this to my estimated risk premium of 5.75% for mature markets (obtained by looking at the implied premium for the S&P 500) to get the total risk premium.
860=575+285 **285=190*1,5**

Surveys about Market Risk Premium

2008	http://ssrn.com/abstract=1344209
2010	http://ssrn.com/abstract=1606563; http://ssrn.com/abstract=1609563
2011	http://ssrn.com/abstract=1822182; http://ssrn.com/abstract=1805852
2012	http://ssrn.com/abstract=2084213
2013	http://ssrn.com/abstract=914160
2014	http://ssrn.com/abstract=1609563

Some comments and webs about MRP and Rf

MRP: http://pages.stern.nyu.edu/~adamodar/New_Home_Page/datafile/ctryprem.html
http://www.market-risk-premia.com/market-risk-premia.html http://www.marktrisikoprämie.de/marktrisikopraemien.html

US risk free rate: http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=yieldYear&year=2015

risk free rate: http://www.basiszinskurve.de/basiszinssatz-gemaess-idw.html
http://www.econ.yale.edu/~shiller/ http://www.cfosurvey.org/pastresults.htm http://alephblog.com/

Equity Premium:
Historical, Expected, Required and Implied

Pablo Fernandez
Professor of Finance. IESE Business School, University of Navarra
Camino del Cerro del Aguila 3. 28023 Madrid, Spain
e-mail: fernandezpa@iese.edu

January 29, 2013

The equity premium designates four different concepts: Historical Equity Premium (HEP); Expected Equity Premium (EEP); Required Equity Premium (REP); and Implied Equity Premium (IEP). We highlight the confusing message in the literature regarding the equity premium and its evolution. The confusion arises from not distinguishing among the four concepts and from not recognizing that although the HEP is equal for all investors, the REP, the EEP and the IEP differ for different investors.

A unique IEP requires assuming homogeneous expectations for the expected growth (g), but we show that there are several pairs (IEP, g) that satisfy current prices. We claim that different investors have different REPs and that it is impossible to determine the REP for the market as a whole, because it does not exist. We also investigate the relationship between (IEP – g) and the risk free rate.

There is a kind of schizophrenic approach to valuation: while all authors admit different expectations of equity cash flows, most authors look for a unique discount rate. It seems as if the expectations of equity cash flows are formed in a democratic regime, while the discount rate is determined in a dictatorship.

1. Introduction

The equity premium (also called *market risk premium*, *equity risk premium*, *market premium* and *risk premium*) is one of the most important, but elusive parameters in finance. Some confusion arises from the fact that the term equity premium is used to designate four different concepts:

1. **Historical** Equity Premium (HEP): historical differential return of the stock market over treasuries.
2. **Expected** Equity Premium (EEP): expected differential return of the stock market over treasuries.
3. **Required** Equity Premium (REP): incremental return of the market portfolio over the risk-free rate required by an investor in order to hold the market portfolio[1]. It is needed for calculating the required return to equity (cost of equity). The CAPM assumes that REP and EEP are unique and that **REP = EEP.**
4. **Implied** Equity Premium (IEP): the required equity premium that arises from a pricing model and from assuming that the market price is correct.

The four concepts are different[2]. The **HEP** is easy to calculate and is equal for all investors[3], but the **REP,** the **EEP** and the **IEP** are different for each investor and are not observable magnitudes. We also claim that there is not an **IEP** for the market as a whole: different investors have different **IEPs** and use different **REPs**. A unique IEP requires assuming homogeneous expectations for the expected growth (g), but there are several pairs (IEP, g) that satisfy current prices.

An anecdote from Merton Miller (2000, page 3) about the expected market return in the Nobel context: *"I still remember the teasing we financial economists, Harry Markowitz, William Sharpe, and I, had to put up with from the physicists and chemists in Stockholm when we conceded that the basic unit of our research, the expected rate of return, was not actually observable. I tried to tease back by reminding them of their neutrino –a particle with no mass whose presence was inferred only as a missing residual from the interactions of other particles. But that was eight years ago. In the meantime, the neutrino has been detected"*.

Different authors claim different relations among the four equity premiums defined above. These relationships vary widely:

- **HEP = EEP = REP** according to Brealey and Myers (1996); Copeland *et al* (1995); Ross *et al* (2005); Stowe *et al* (2002); Pratt (2002); Bruner (2004); Bodie *et al* (2003); Damodaran (2006); Goyal and Welch (2007); Ibbotson Ass. (2006).
- **EEP is smaller than HEP** according to Copeland *et al* (2000, HEP-1.5 to 2%); Goedhart *et al* (2005, HEP-1 to 2%); Bodie *et al* (1996, HEP-1%); Mayfield (2004, HEP-2.4%); Booth (1999, HEP-2%); Bostock (2004, 0.6 to 1.8%); Dimson *et al* (2006c, 3 to 3.5%); Siegel (2005b, 2 to 3%); Ibbotson (2002, < 4%); Campbell (2002, 1.5 to 2%); Campbell (2007, 4%)[4].
- **EEP is near zero** according to McGrattan and Prescott (2001); Arnott and Ryan (2001); Arnott and Bernstein (2002).
- Authors that try to find the **EEP doing surveys,** as Welch (2000, 7%); Welch (2001, 5.5%); Graham and Harvey (2007: 4.65% in 2000; 2.39% in nov. 05; 3.21% in nov. 06); O'Neill *et al* (2002, 3.9%).
- There is **a unique IEP and REP = IEP,** according to Damodaran (2001a); Arzac (2005); Jagannathan *et al* (2000); Harris and Marston (2001); Claus and Thomas (2001); Fama and French (2002); Goedhart *et al* (2002); Harris *et al* (2003); Vivian (2005).
- Authors that ***"have no official position"***, as Brealey and Myers (2000, 2003, 2005).
- Authors that claim ***"that no one knows what the REP is"***, as Penman (2003).
- Authors that claim that ***"it is impossible to determine the REP for the market as a whole, because it does not exist"***, as Fernandez (2002).
- Authors that claim that ***"different investors have different REPs"***, as Fernandez (2004).

[1] Or the extra return that the overall stock market must provide over the Government Bonds to compensate for the extra risk.

[2] We agree with Bostock (2004) when he says that *"understanding the equity premium is largely a matter of using clear terms"*.

[3] Provided they use the same time frame, the same market index, the same risk-free instrument and the same average (arithmetic or geometric).

[4] However, his figure 4 shows a world equity premium lower than 2% in the period 1985-2002.

The rest of this paper is organized as follows. In section 2 we revise different estimates of the Historical Equity Premium (HEP), note that not all the authors get the same result for the HEP, and analyze the data. We highlight the change in the market around 1960. Before that date, the dividend yield was higher than the risk-free rate, but after that date has been always smaller. In sections 3 and 4 we discuss different estimates of the Expected Equity Premium (EEP) and of the Required Equity Premium (REP). In section 5 we revise the equity premium puzzle. Section 6 is a revision of the prescriptions of the main finance textbooks about the risk premium. We highlight the confusing message of the textbooks regarding the equity premium and its evolution. In section 7, we show that there are several pairs (IEP, g) that explain current market prices and we argue that there is no a REP for the market as a whole, but rather different investors use different REPs. We also show a positive relationship between (IEP – g) and the risk free rate after 1960. Section 8 explains which REP uses the author. Finally, section 9 concludes.

2. Historical Equity Premium (HEP)

The HEP is the historical average differential return of the market portfolio over the risk-free debt[5]. The most widely cited source is Ibbotson Associates whose U.S. database starts in 1926. Another frequently used source is the Center for Research in Security Prices (CRSP) at the University of Chicago.

2.1. First studies of the historical equity return

Smith (1926) made the first empirical estimate of the long run return on stocks (only price changes) for the most actively traded stocks from 1901 to 1922, and showed that an equity investor (even without market timing or stock selection ability) outperformed a bond investor over this period[6].

Cowles (1939) published the first empirical study carefully done on the performance of the stock market. Cowles calculated the total return to equity from 1872 to 1937 for the NYSE, documenting a positive long term equity performance.

Fisher and Lorie (1964), using for the first time the database of stock prices completed at the University of Chicago's Center for Research in Security Prices (CRSP), showed that the average return from a random investment in NYSE stocks from 1926 to 1964 was 9.1% a year[7].

2.2. Estimates of the historical equity premium of the US

Table 1 contains the 1926-2005 average returns and HEP for the US according to Ibbotson Associates (2006). The HEP in table 1 is the difference between the average return on the S&P 500 and the return of Gov. Bonds or T-Bills. However, Ibbotson Associates (2006, page 73), use the income return (the portion of the total return that results from a periodic bond coupon payment) of the Gov. Bonds (5.2%) and consider that the relevant HEP during the period 1926-2005 is 7.1% (12.3-5.2).

Schwert (1990) and Siegel (1994, 1999, 2002, 2005a) studied the relationship between U.S. equity and bonds before 1926. The data on which they base their studies is less reliable than recent data, but the results are interesting, nevertheless. Table 2 shows their conclusions: the HEP and the inflation in the period 1802-1925 were substantially smaller than in subsequent years[8]. Note that table 1 provides a higher HEP than table 2 for the period after 1926 because Ibbotson do not consider the income return of the bonds.

[5] This average differential return may be arithmetic or geometric. Different stock market indexes are used as the market portfolio, and Government bonds of different maturities are used as risk-free debt. A good discussion of the geometric and arithmetic average is Jacquier, Kane, and Marcus (2003).
[6] Three years after publication, the market crash happened. Benjamin Graham blamed Smith's book for inspiring an *“orgy of uncontrolled speculation”*.
[7] For a more detailed history see Goetzmann and Ibbotson (2006).
[8] Siegel (1999) argues that this is because bond returns were exceptionally low after 1926, while total equity returns were relatively stable over the whole time period.

Wilson and Jones (2002) provide a monthly stock price index from 1871 through 1999. They note that the S&P Index returns have often been misrepresented[9] and reconstruct the weekly S&P Composite for the period 1926-56 containing more than 400 stocks (instead of 90 as the daily S&P Composite). They get some differences versus other used indexes that are summarized on table 3.

Ibbotson and Chen (2003) use 1926-2000 historical equity returns and conclude that the expected long-term equity premium (relative to the long-term government bond yield) is 5.9% arithmetically, and 3.97% geometrically.

Goetzmann and Ibbotson (2006) employ a new NYSE database for 1815–1925[10] to estimate the U.S. equity returns and the HEP since 1792 (but they mention that dividend data is absent pre-1825, and is incomplete in the period 1825–71). Their main results are in table 4.

Table 1. Returns and HEP according to Ibbotson Associates (2006). 1926-2005

	Average return		Standard	Serial
Nominal Returns 1926-2005	Arithmetic	Geometric	deviation	correlation
S&P 500	12,3%	10,4%	20,2%	3%
Income	*4,2%*	*4,2%*	*1,6%*	*89%*
Capital appreciation	*7,8%*	*5,9%*	*19,5%*	*3%*
Long-Term Gov. Bonds	5,8%	5,5%	9,2%	-8%
Income	*5,2%*	*5,2%*	*2,7%*	*96%*
Capital appreciation	*0,5%*	*0,4%*	*4,4%*	*-19%*
T-Bills	3,8%	3,7%	3,1%	91%
Inflation	3,1%	3,0%	4,3%	65%

HEP over Gov. Bonds	6,5%	4,9%
HEP over T-Bills	8,5%	6,7%

Table 2 - Real returns and HEP from Siegel (2005a)

arith. = arithmetic average. geom. = geometric average

	Average real returns (%)						
	Stocks		Bonds		HEP (%)		
	arith.	geom.	arith.	geom.	arith.	geom.	Inflation (%)
1802-1870	8.28	7.02	5.11	4.78	3.17	2.24	0.1
1871-1925	7.92	6.62	3.93	3.73	3.99	2.89	0.6
1926-2004	8.78	6.78	2.77	2.25	6.01	4.53	3.1
1802-2004	8.38	6.82	3.88	3.51	4.50	3.31	1.4

Table 3. Geometric average of the returns of different indexes in selected periods

(%)	Cowles	S&P	Wilson and Jones	Ibbotson	CRSP NYSE
1871-1925	7,24	7,28	7,28		
1926-1940	3,27	4,20	3,23	4,04	3,01
1941-1956		15,60	15,20	16,11	15,36
1957-1999		12,10	12,28	12,24	11,79
1926-1999		11.08	11.00	11.35	10.70
1871-1999		9,51	9,40		

Table 4. Average return of the US according to Goetzmann and Ibbotson (2006)

	1792-1925		
	Arithmetic return	Geometric return	Standard deviation
Stocks	7.93%	6.99%	14.64%
Bonds	4.17%	4.16%	4.17%
Comm. Paper	7.62%	7.57%	3.22%
Inflation	0.85%	0.61%	7.11%
HEP (Bonds)	3.76%	2.83%	

	1926-2004		
	Arithmetic return	Geometric return	Standard deviation
Stocks	12.39%	10.43%	20.32%
Gov. Bonds	5.82%	5.44%	9.30%
T-Bills	3.76%	3.72%	3.14%
Inflation	3.12%	3.04%	4.32%
HEP (Bonds)	6.57%	4.99%	
HEP (Bills)	8.63%	6.71%	

Total returns from 1871 to 1925 are constructed from the Price-Weighted NYSE and the Cowles Income Return Series.

[9] Standard & Poor's first developed stock price indices in 1923 and in 1927 created the Composite Index (90 stocks). On 1 March 1957, the Composite was expanded to 500 stocks and renamed S&P 500 Index (its market value was $173 billion, 85% of the value of all NYSE listed stocks). From 1926 to 1957 there were 2 different S&P Composite indexes: one was weekly and the other was daily. The S&P Composite daily covered 90 stocks until 1957; The S&P Composite weekly covered more than 400.

[10] See Goetzmann, Ibbotson, and Peng (2001), who collected U.S. stock market data by hand from 1815.

In a very interesting article, Siegel and Schwartz (2006) calculate the return of the original S&P 500 companies since 1957 until 2003 and find that their return has been higher than the return of the S&P 500[11]. The average geometric return of the S&P 500 was 10.85% (standard deviation of 17%), while the return of the original 500 companies was 11.31% (standard deviation of 15.7%).

Table 5. Different Historical Equity Premiums (HEP) in the US according to different authors

			Ibbotson	Shiller	WJ	Damodaran	Siegel	Max-min
HEP vs. LT Gov. Bonds	Geometric	1926-2005	4,9%	5,5%	4,4%	5,1%	4,6%	1,0%
		1926-1957	6,0%	7,3%	5,1%	5,8%		2,2%
		1958-2005	4,1%	4,2%	4,0%	4,5%		0,6%
	Arithmetic	1926-2005	6,5%	7,0%	5,8%	6,7%	6,1%	1,2%
		1926-1957	8,8%	10,1%	7,6%	8,7%		2,5%
		1958-2005	4,9%	5,0%	4,7%	5,4%		0,7%
HEP vs. T-Bills	Geometric	1926-2005	6,7%	6,0%	6,2%	6,3%	6,2%	0,7%
		1926-1957	8,2%	8,4%	7,3%	7,6%		1,1%
		1958-2005	5,6%	4,3%	5,4%	5,4%		1,3%
	Arithmetic	1926-2005	8,5%	7,7%	7,9%	8,2%	8,2%	0,8%
		1926-1957	11,1%	11,2%	9,9%	10,5%		1,4%
		1958-2005	6,8%	5,4%	6,6%	6,6%		1,5%

Ibbotson figures come from Ibbotson Associates (2006). Shiller figures come from http://aida.econ.yale.edu/~shiller/data.htm. WJ figures have been updated from Wilson and Jones (2002). Damodaran figures come from http://pages.stern.nyu.edu/~adamodar/. Siegel figures have been updated from Siegel (2005a).

Note that not all the authors get the same result, even for the HEP. Table 5 is a comparison of the HEP in the US according to different authors. The differences are substantial, especially for the period 1926-1957. The differences are mainly due to the stock indexes chosen. It is also important to keep in mind that the data from the 19th century and from the first part of the 20th century is quite poor and questionable. Table 6 shows the differences among the different indexes commonly used.

Table 6. Number of securities in the US indexes commonly used

	S&P composite weekly	Ibbotson	CRSP NYSE
1926-1957	228 stocks in 1927, 410 in 1928, 480 in 1956	S&P Composite daily: 90 stocks	Growing number of stocks: 592 in 1927; 1059 in 1957
1957-2006	abandoned	S&P Composite daily: 500 stocks	Growing number of stocks: 1500 in 1975; 2813 in 1999

2.3. A closer look at the historical data

Figure 1 shows that interest rates were lower than dividend yields until 1958 and than the earnings to price ratio until the 1980s. It suggests that many things have changed in the capital markets and that the last 40 years have been different than the previous ones. It is quite sensible to assume that the portfolio theory, the CAPM, the APT, the VAR analysis, the futures and options markets, the appearance of many mutual and hedge funds, the increase of investors, the legislation to protect investors, financial innovation, electronic trading, portfolio insurance, market participation,… have changed the behaviour and the risk attitudes of today's investors vs. past investors. In fact, financial markets are so different that the relative magnitude of dividend yields to interest rates has been reversed.

It is interesting to look at historical data to know what happened to our grandparents (or to our great grandparents), but it is not sensible to assume that their markets and their investment behaviour were similar to ours[12].

Figure 2 shows the evolution of the 20-year rolling correlation of (dividend yield – R_F) versus R_F (the yield on Government long-term bonds). Again, we may see that something has changed in the

[11] The market value of the S&P 500 companies that have survived from the original 1957 list was only 31% of the 2003 year-end S&P 500's market value. Since the S&P 500 was formulated, more than 900 new companies have been added to the index (and an equal number deleted from).

[12] Neither the exam of Ec1010 in 1932 is very useful for a student today.

markets because that correlation after 1960 has been lower than ever before. Figure 3 shows the raw data used to calculate the correlations of Figure 2 and permits to contrast the different behavior of the markets in the periods 1871-1959 and 1960-2005. In section 7 we analyze this data and derive implications.

Figure 4 shows the evolution of the 20-year rolling HEP (arithmetic and geometric) relative to the T-Bills. It may be seen that the periods with equity returns much higher than the T-Bill rates were the 50s and the 90s.

Figure 5 compares the 20-year rolling HEP with the current T-Bond yield. From 1960 to 2000 the HEP increased when the yield decreased and vice versa. It did not happen so clearly in previous years.

Figure 1. 10-year T-Bond yields, Earnings to Price ratio (E/P) and Dividend yield of the US
Source: Robert Shiller's Website: http://aida.econ.yale.edu/~shiller/data.htm

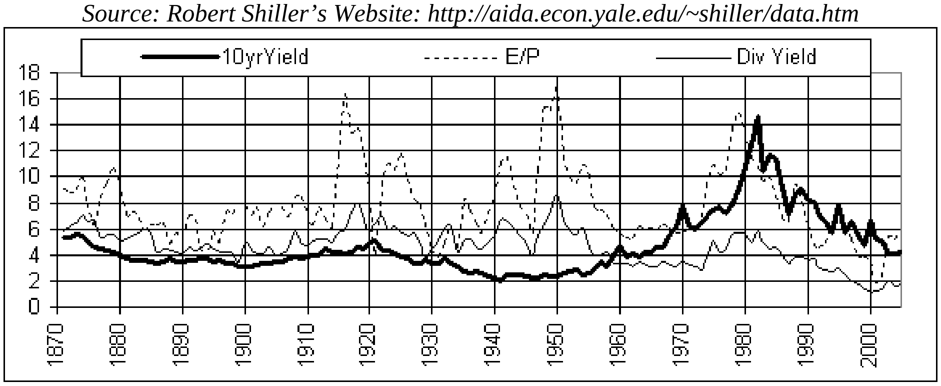


Figure 2. 20-year rolling correlation of (dividend yield – R_F) versus R_F (yield on T-Bonds). Monthly data.
Source of the raw data: Robert Shiller's Website: http://aida.econ.yale.edu/~shiller/data.htm

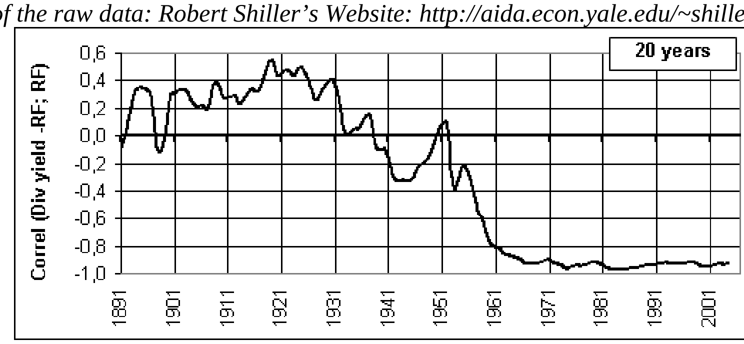


Figure 3. (Dividend yield – R_F) versus R_F (yield on Government long-term bonds)
Source of the raw data: Robert Shiller's Website: http://aida.econ.yale.edu/~shiller/data.htm

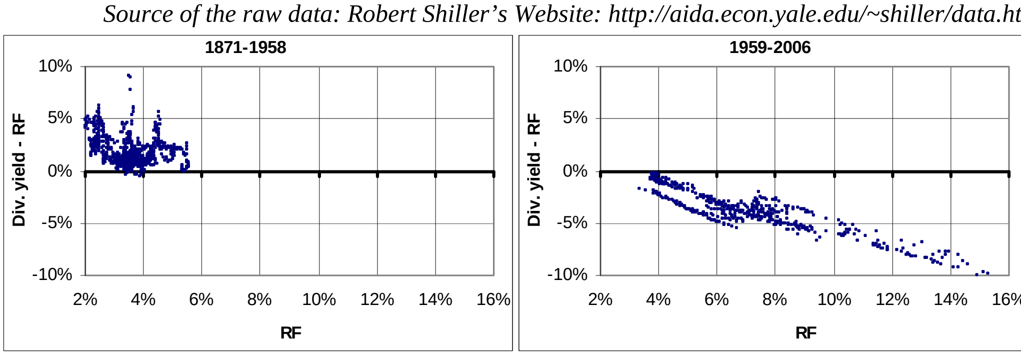


Figure 4. 20-year rolling HEP versus the T-Bills.
Source of the raw data: Robert Shiller's Website: http://aida.econ.yale.edu/~shiller/data.htm

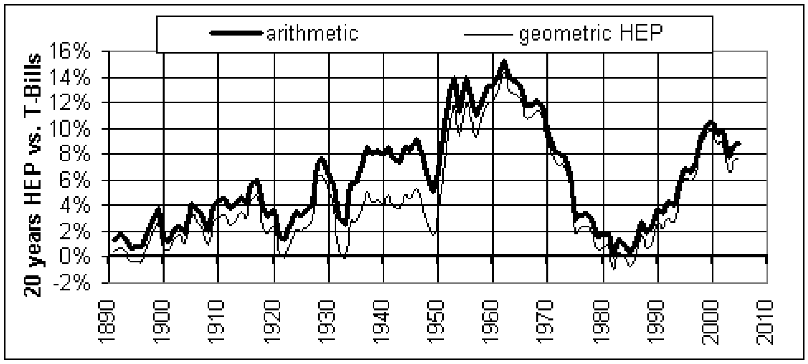
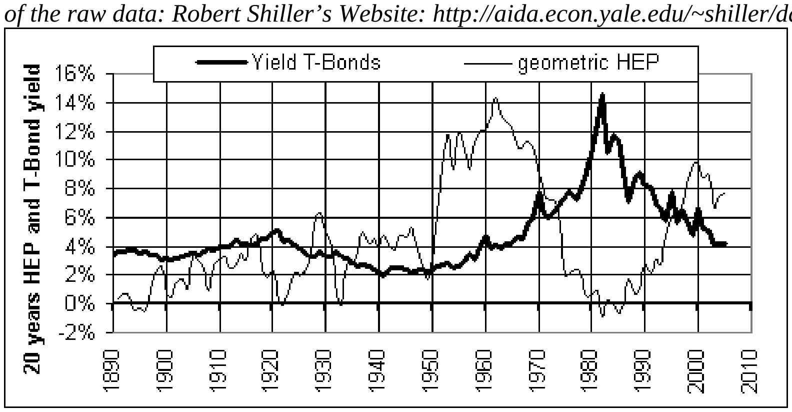


Figure 5. 20-year rolling geometric HEP versus the T-Bills, and T-Bond yield
Source of the raw data: Robert Shiller's Website: http://aida.econ.yale.edu/~shiller/data.htm



2.4. Estimates of the Historical Equity Premium (HEP) in other countries

Blanchard (1993) examined the evolution of stock and bonds rates over the period 1978 to 1992 for the US, Japan, Germany, France, Italy and the UK. He constructed 'world' rates of return (using relative GDP weights for the countries) and documented a postwar decline in the dividend yield and in various measures of the HEP.

Table 7. Equity return of selected countries, according to Jorion and Goetzmann (1999)

Country	Period	Nominal Return	Real Return	Dollar Return	Inflation
U.S.	21-96	6.95%	4.32%	6.95%	2.52%
Sweden	21-96	7.42%	4.29%	7.00%	3.00%
Germany	21-96	4.43%	1.91%	5.81%	2.47%
Canada	21-96	5.78%	3.19%	5.35%	2.51%
U.K.	21-96	6.30%	2.35%	5.20%	3.86%
France	21-96	9.09%	0.75%	4.29%	8.28%
Belgium	21-96	4.45%	-0.26%	3.51%	4.73%
Italy	28-96	10.10%	0.15%	3.22%	9.94%
Japan	21-96	7.33%	-0.81%	1.80%	8.21%
Spain	21-96	4.66%	-1.82%	1.53%	6.61%
Median 39 countries			0.75%	4.68%	
11 countries with continuous histories into the 1920s:		Mean	1.88%	5.09%	
		Median	2.35%	5.20%	

Jorion and Goetzmann (1999) constructed a database of capital gain indexes for 39 markets, with 11 of them starting in 1921(see table 7). However, they obtained pre-1970 dividend information only for 6 markets. They concluded that *"for 1921 to 1996, US equities had the highest real return for all countries, at 4.3%, versus a median of 0.8% for other countries. The high equity premium obtained for U.S. equities appears to be the exception rather than the rule"*. According to the authors, *"there are reasons to suspect that [the US] estimates are subject to survivorship"*.

However, Dimson and Marsh (2001) do not find survivorship bias for the US. They calculate the geometric HEP for 1955-1999 of US, UK, Germany and Japan and get 6.2%, 6.2%, 6.3% and 7.0%.

Table 8. HEP vs. short (30 days) and long term (10 or 30 years) fixed income in 17 countries. 1900-2005. Annualized returns. *Source: Table 3 of Dimson, Marsh and Staunton (2006c)*

% p.a.	HEP relative to					
	Bills			Bonds		
Country	Geometric Mean	Arithmetic Mean	Standard Error	Geometric Mean	Arithmetic Mean	Standard Error
Australia	7,08	8,49	1,65	6,22	7,81	1,83
Japan	6,67	9,84	2,70	5,91	9,98	3,21
South Africa	6,20	8,25	2,15	5,35	7,03	1,88
Germany	3,83	9,07	3,28	5,28	8,35	2,69
Sweden	5,73	7,98	2,15	5,21	7,51	2,17
U.S.	5,51	7,41	1,91	4,52	6,49	1,96
U.K.	4,43	6,14	1,93	4,06	5,29	1,61
Italy	6,55	10,46	3,12	4,30	7,68	2,89
Canada	4,54	5,88	1,62	4,15	5,67	1,74
France	6,79	9,27	2,35	3,86	6,03	2,16
Netherlands	4,55	6,61	2,17	3,86	5,95	2,10
Ireland	4,09	5,98	1,97	3,62	5,18	1,78
Belgium	2,80	4,99	2,24	2,57	4,37	1,95
Norway	3,07	5,70	2,52	2,55	5,26	2,66
Spain	3,40	5,46	2,08	2,32	4,21	1,96
Denmark	2,87	4,51	1,93	2,07	3,27	1,57
Switzerland	3,63	5,29	1,82	1,80	3,28	1,70
Average	4,81	7,14	2,21	3,98	6,08	2,11
World-ex U.S.	4,23	5,93	1,88	4,10	5,18	1,48

Dimson *et al* (2006c) use a unique database to calculate the historical equity premium for 17 countries over 106 years (1900-2005). Their estimates (see Table 8) are lower than frequently quoted HEPs mainly due to the incorporation of the earlier part of the 20th century as well as the opening years of the 21st century[13].

But, apart from the historical interest, how useful and accurate is that data? As Dimson *et al* (2006c) point out, *"virtually all of the 16 countries experienced trading breaks … often in wartime. The U.K. and European exchanges, and even the NYSE, closed at the start of World War I…Similarly, the Danish, Norwegian, Belgian, Dutch and French markets …when Germany invaded in 1940, and even the Swiss market closed from May to July 1940 for mobilization. … Japan after the Great Tokyo Earthquake of 1923. …Germany and Japan from towards the end of World War II, and Spain during the Civil War"*. They claim that *"we were able to bridge these gaps"*, but this assertion is questionable. They admit that *"the end-year index levels recorded for Germany for 1943–47, Japan for 1945, and Spain for 1936–38 cannot be regarded as market-determined values"*. Dimson *et al* (2006c) explain in their footnote 7 that *"In Spain, trading was suspended during the Civil War from July 1936 to April 1939, and the Madrid exchange remained closed through February 1940; over the closure we assume a zero change in nominal stock prices and zero dividends"*. It is not clear why this assumption is a reasonable one. They also mention one *"unbridgeable discontinuity, namely, bond and bill (but not equity) returns in Germany during the hyperinflation of 1922–23, when German bond and bill investors suffered a total loss of –100%. …bonds and bills can become riskier than equities. When reporting equity premiums for Germany … we thus have no alternative but to exclude the years 1922–23"*.

In a previous work Dimson, Marsh and Staunton (2002) show that the HEP was generally higher for the second half century: the World had 4.7% in the first half, compared to 6.2% in the second half.

Table 9 contains some of the HEPs reported by different authors for the US.

Table 9. Historical Equity Premium (HEP) for the US according to different authors

[13] Their database contains annual returns on stocks, bonds, bills, inflation, and currencies for 17 countries from 1900–2005, and is described in Dimson *et al* (2006a and 2006b). They construct a World equity index (U.S. dollars index of 17 countries weighted by its starting-year market capitalization or by its GDP, before capitalizations were available) and a World bond index, constructed with each country weighted by its GDP. The series were compiled to avoid the survivorship bias that can arise from backfilling. Their choice of international markets was limited by their requirement to have data for the whole century.

Author(s)	Reference/average	Period for HEP	Value
Siegel (2002)	T-Bonds, geo.	1926-2001	4.9%
Ibbotson and Chen (2003)	T-Bonds, geo.	1926-2000	3.97%
Siegel (2005a)	T-Bonds, geo.	1926-2004	4.53%
Ibbotson Associates (2006)	T-Bonds arith. capital aprec. only	1926-2005	7.1%
Goetzmann and Ibbotson (2006)	T-Bonds, geo.	1792-1925	2.83%
Goetzmann and Ibbotson (2006)	T-Bonds, geo.	1926-2004	4.99%
Goyal and Welch (2007)		1872-2004	4.77%
Goyal and Welch (2007)		1927-2004	6.35%
Dimson & al.(2006c)	T-Bonds, geo. US	1900-2005	4.52%
Dimson & al.(2006c)	T-Bonds, geo. World	1900-2005	4.04%

This section has revised different estimates of the Historical Equity Premium (HEP) and permits to note that not all the authors get the same result for the HEP. We highlight the change in the market around 1960. Before that date, the dividend yield was higher than the risk-free rate, but after that date has been always smaller. We question the usefulness of historical data to predict the future.

3. Expected Equity Premium (EEP)

The **Expected** Equity Premium (EEP) is the answer to a question we would all (especially analysts and fund managers) like to answer accurately in the short term, namely: what incremental return do I expect from the market portfolio over the risk-free rate over the next years? Campbell (2007, pg. 1) identifies the EEP with the REP: *"What return should investors expect the stock market to deliver, above the interest rate on a safe short-term investment? In other words, what is a reasonable estimate of the equity premium?"*

Estimates of the EEP based on historical analysis presume that the historical record provides an adequate guide for future expected long-term behaviour. However, the HEP changes over time, and it is not clear why capital market data from the 19th century or from the first half of the 20th century may be useful in estimating expected returns in the 21st century.

Numerous papers assert that there must be **an** EEP common to all investors (to the representative investor). But it is obvious that investors do not share "homogeneous expectations"[14] and, also, that many investors do not hold the market portfolio but, rather, a subgroup of stocks and bonds[15]. Heterogeneous investors do not hold the same portfolio of risky assets; in fact, no investor must hold the market portfolio to clear the market.

We claim in section 7 that without "homogeneous expectations" there is **not one** EEP (but several), and there is **not one** REP (but several).

3.1. The Historical Equity Premium (HEP) is not a good estimator of the EEP

Although many authors consider that the equity premium is a stationary process, and then the HEP is an unbiased estimate of the EEP (*unconditional* mean equity premium), we do not agree with that statement: the HEP is not a good estimator of the EEP. For example, Mehra and Prescott (2003) state that *"…over the long horizon the equity premium is likely to be similar to what it has been in the past"*.

The magnitude of the error associated with using the HEP as an estimate of the EEP is substantial. Shiller (2000) points out that *"the future will not necessarily be like the past"*. Booth (1999) concludes that the HEP is not a good estimator of the EEP and estimates the later in 200 basis points smaller than the HEP[16]. Mayfield (2004) suggest that a structural shift in the process governing the volatility of market returns after the 1930s resulted in a decrease in the expected level of market risk, and concluded that EEP = HEP – 2.4% = 5.9% over the yield on T-bills (4.1% over yields on T-bonds).

[14] Brennan (2004) also admits that "*different classes of investor may have different expectations about the prospective returns on equities which imply different assessments of the risk premium*".
[15] But, even with "homogeneous expectations" (all investors have equal EEP), the REP would not be equal for all investors. In that situation, the investors with lower REP would clear the market.
[16] He also points out that the nominal equity return did not follow a random walk and that the volatility of the bonds increased significantly over the last 20 years.

Survivorship bias[17] was identified by Brown, Goetzmann and Ross (1995) as one of the main reasons why the results based on historical analyses can be too optimistic. They pointed out that the observed return, *conditioned on survival* (HEP), can overstate the unconditional expected return (EEP). However, Li and Xu (2002) show that the survival bias fails to explain the equity premium puzzle: *"To have high survival bias, the probability of market survival over the long run has to be extremely small, which seems to be inconsistent with existing historical evidence"*. Siegel (1999, p. 13) mentions that *"Although stock returns may be lower in foreign countries than in the U.S., the real returns on foreign bonds are substantially lower"*.

Pastor and Stambaugh (2001) present a framework allowing for structural breaks in the risk premium over time and estimate that the EEP fluctuated between 4% and 6% over the period from 1834 to 1999, declined steadily since the 1930s (except for a brief period in the mid-1970s) and had the sharpest drop in the last decade of the 20th century. Using extra information from return volatility and prices, they narrow the confidence interval of their estimation (two standard deviations) to plus or minus 280 basis points around 4.8%.

Constantinides (2002) addresses different ways in which we may account for biases in the sample mean premium in order to estimate the expected premium and draws a sharp distinction between conditional, short-term forecasts of the mean equity premium and estimates of the unconditional mean. He says that the conditional EEPs at the end of the 20th century and the beginning of the 21st are substantially lower than the estimates of the unconditional EEP (7%) *"by at least three measures"*. But he concludes that *"the currently low conditional, short-term forecasts of the equity premium do not necessarily imply that the unconditional estimate of the mean premium is lower than the sample average. Therefore, the low conditional forecasts do not necessarily lessen the burden on economic theory to explain the large sample average of the equity return and premium over the past 130 years"*.

Dimson *et al* (2003) highlight the survivorship bias relative to the market, *"even if we have been successful in avoiding survivor bias within each index, we still focus on markets that survived"* and concluded that the geometric EEP for the world's major markets should be 3% (5% arithmetic). Dimson *et al* (2006c) admit that *"we cannot know today's consensus expectation for the equity premium"*, but they conclude that *"investors expect an equity premium (relative to bills) of around 3-3½% on a geometric mean basis"*, substantially lower than the HEP found in their own study.

3.2. Surveys

A direct way to obtain an expectation of the equity premium is to carry out a survey of analysts or investors although Ilmanen (2003) argues that surveys tend to be optimistic: *"because of behavioural biases, survey-based expected returns may tell us more about hoped-for returns than about required returns"*.

Welch (2000) performed two surveys with finance professors in 1997 and 1998, asking them what they thought the EEP was over the next 30 years. He obtained 226 replies, ranging from 1% to 15%, with an average arithmetic EEP of 7% above T-Bonds.[18] Welch (2001) presented the results of a survey of 510 finance and economics professors performed in August 2001 and the consensus for the 30-year arithmetic EEP was 5.5%, much lower just 3 years earlier.

Graham and Harvey (2005) indicate that U.S. CFOs reduced their average EEP from 4.65% in September 2000 to 2.93% by September 2005. Over this period, the HEP had fallen only 0.4%.

Goldman Sachs (O'Neill, Wilson and Masih, 2002) conducted a survey of its global clients in July 2002 and the average long-run EEP was 3.9%, with most responses between 3.5% and 4.5%. The magazine *Pensions and Investments* (12/1/1998) carried out a survey among professionals working for institutional investors and the average EEP was 3%.

3.3. Regressions

Attempts to predict the equity premium typically look for some independent lagged predictors (X) on the equity premium: $\text{Equity Premium}_t = a + b \cdot X_{t-1} + \varepsilon_t$

---

[17] "Survivorship" or "survival" bias applies not only to the stocks within the market (the fact that databases contain data on companies listed today, but they tend not to have data on companies that went bankrupt or filed for bankruptcy protection in the past), but also for the markets themselves ("US market's remarkable success over the last century is typical neither of other countries nor of the future for US stocks" (Dimson *et al* 2004)).

[18] The interest rate paid by long-term T-bonds in April 1998 was approximately 6%. At that time, the most recent Ibbotson Associates Yearbook was the 1998 edition, with an arithmetic HEP versus T-bills of 8.9% (1926–1997).

Many predictors have been explored in the literature. Some examples are:

- Dividend yield: Ball (1978), Rozeff (1984), Campbell (1987), Campbell and Shiller (1988), Fama and French (1988), Hodrick (1992), Campbell and Viceira (2002), Campbell and Yogo (2003), Lewellen (2004), and Menzly, Santos, and Veronesi (2004). Cochrane (1997) has a good survey of the dividend yield prediction literature.
- The short term interest rate: Hodrick (1992).
- Earnings price and payout ratio: Campbell and Shiller (1988), Lamont (1998) and Ritter (2005).
- The term spread and the default spread: Avramov (2002), Campbell (1987), Fama and French (1989), and Keim and Stambaugh (1986).
- The inflation rate (money illusion): Fama and Schwert (1977), Fama (1981), and Campbell and Vuolteenaho (2004a,b), and Cohen, Polk and Vuolteenaho (2005).
- Interest rate and dividend related variables: Ang and Bekaert (2003).
- Book-to-market ratio: Kothari and Shanken (1997).
- Value of high and low-beta stocks: Polk, Thompson and Vuolteenaho (2006)[19].
- Consumption and wealth: Lettau and Ludvigson (2001).
- Aggregate financing activity: Baker and Wurgler (2000) and Boudoukh *et al* (2006).

Goyal and Welch (2007) used most of the mentioned predictors and could not identify one that would have been robust for forecasting the equity premium and, after all their analysis, they recommended *"assuming that the equity premium is 'like it always has been'"*. They also show that most of these models have not performed well for the last thirty years, that are not stable, and that are not useful for market-timing purposes.

However, Campbell and Thompson (2007) claim that some variables (ratios, patterns, levels of sort and long term interest rates) are correlated with subsequent market returns and that *"forecasting variables with significant forecasting power insample generally have a better out-of-sample performance than a forecast based on the historical average return"*. They explore the mapping from R^2 statistics in predictive regressions to profits and welfare gains for market timers. *"The basic lesson is that investors should be suspicious of predictive regressions with high R^2 statistics, asking the old question 'If you're so smart, why aren't you rich?'"*

3.4. Other estimates of the expected equity premium

Siegel (2002, page 124) concluded that "the future equity premium is likely to be in the range of 2 to 3%, about one-half the level that has prevailed over the past 20 years" [20]. Siegel (2005a, page 172) affirms that *"over the past 200 years, the equity risk premium has averaged about 3%"*. Siegel (2005b) maintains that *"although the future equity risk premium is apt to be lower than it has been historically, U.S. equity returns of 2-3% over bonds will still amply reward those who will tolerate the short-term risk of stocks"*. However, in a presentation at the SIA annual meeting (November 10, 2005) Siegel maintained that *"equity premium is 4% to 5% now"*.

In the *TIAA-CREF Investment Forum* of June 2002, Ibbotson forecasted *"less than 4% in excess of long-term bond yields"*, and Campbell *"1.5% to 2%"*.

McGrattan and Prescott (2001) did not find corporate equity overvalued in 2000 and forecasted that the real returns on debt and equity should both be near 4%: *"Therefore, barring any institutional changes, we predict a small equity premium in the future"*.

Arnott and Ryan (2001) claim that the expected equity premium is near zero. They base their conclusion on the low dividend yield and their low expectation of dividend growth. Arnott and Bernstein (2002) also conclude that *"the current risk premium is approximately zero"*.

[19] Polk, Thompson, and Vuolteenaho (2006) argue that if the CAPM holds, then a high equity premium implies low prices for stocks that have high betas. Therefore, value stocks should tend to have high betas. This was true from the 1930's through the 1950's, but in recent decades growth stocks have had higher betas than value stocks. Polk, Thompson, and Vuolteenaho argue that this change in cross-sectional stock pricing reflects a decline in the equity premium.

[20] Siegel also affirms that: "Although it may seem that stocks are riskier than long-term government bonds, this is not true. The safest investment in the long run (from the point of view of preserving the investor's purchasing power) has been stocks, not Treasury bonds".

Bostock (2004) concludes that according to historical average data, equities should offer a risk premium over government bonds between 0.6% and 1.8%.

Grabowski (2006) concludes that *"after considering the evidence, any reasonable long-term estimate of the normal EEP as of 2006 should be in the range of 3.5% to 6%"*.

Maheu and McCurdy (2006) claim that the US Market had *"three major structural breaks (1929, 1940 and 1969), and possibly a more recent structural break in the late 1990s"*, and suggest an EEP in 2004 between 4.02% and 5.1%.

Table 10. Estimates of the EEP (Expected Equity Premium) according to different authors

Authors	Conclusion about EEP	Note
Surveys		
Pensions and Investments (1998)	3%	Institutional investors
Graham and Harvey (2000)	4.65%	CFOs
Welch (2000)	7% arithmetically, 5.2% geometrically	Finance professors
Welch (2001)	5.5% arithmetically, 4.7% geometrically	Finance professors
O'Neill, Wilson and Masih (2002)	3.9%	Global clients Goldman
Graham and Harvey (2005)	2.93%	CFOs
Other publications		
Booth (1999)	EEP = HEP - 2%	
Pastor and Stambaugh (2001)	4 -6%	
McGrattan and Prescott (2001)	near zero	
Arnott and Ryan (2001)	near zero	
Arnott and Bernstein (2002)	near zero	
Siegel (2002, 2005b)	2 - 3%	
Ibbotson (2002)	< 4%	
Campbel (2002)	1.5 - 2%	
Mayfield (2004)	EEP = HEP - 2.4%= 5.9% + T-Bill	
Bostock (2004)	0.6 – 1.8%	
Goyal and Welch (2007)	EEP = HEP	
Dimson, Marsh and Stauton (2006c)	3 - 3.5%	
Grabowski (2006)	3.5 – 6%	
Maheu and McCurdy (2006)	4.02% and 5.1%.	
Ibbotson Associates (2006)	EEP = HEP = 7.1%	

4. Required and implied equity premium

The Required Equity Premium (REP) of an investor is the incremental return that she requires, over the risk-free rate, for investing in a diversified portfolio of shares. It is a crucial parameter in valuation and capital budgeting because the REP is the key to determining the company's required return to equity and the required return to any investment project. The HEP is misleading for predicting the REP. If there was a reduction in the REP, this fall in the discount rate led to re-pricing of stocks, thus adding to the magnitude of HEP. The HEP, then, overstates the REP.

The IEP is the implicit REP used in the valuation of a stock (or a market index) that matches the current market value with an estimate of the future cash flows to equity. The IEP is also called the *ex ante* equity premium. However, the existence of *a unique* IEP implies to consider that the equity market can be explained with a representative consumer, or to consider that all investors have at any moment the same expectations about future cash flows and use the same discount rate to value each company.

Two models are widely used to calculate the IEP: the Gordon (1962) model (constant dividend growth model) and the residual income (or abnormal return) model.

According to the Gordon (1962) model, the current price per share (P_0) is the present value of expected dividends discounted at the required rate of return (k). If d_1 is the dividend per share expected to be received at time 1, and g the expected long term growth rate in dividends per share[21],

$P_0 = d_1 / (k - g)$, which implies: $k = d_1/P_0 + g$. $IEP = d_1/P_0 + g - R_F$ (1)

The abnormal return method is another version of the Gordon (1962) model when the "clean surplus" relation holds ($d_t = e_t – (BV_t – BV_{t-1})$, being d the dividends per share, e the earnings per share and bv the book value per share):

$P_0 = bv_0 + (e_1 – k\, bv_0) / (k - g)$, which implies: $k = e_1/P_0 + g\,(1 - bv_0/ P_0)$[22] (2)

[21] Although we say "dividends per share", we refer to equity cash flow per share: dividends, repurchases and all expected cash for the shareholders.

[22] Comparing the two models, it is clear than in a growing perpetuity, $D_1 = E_1 – g\, BV_0$. The equivalence of the two models may be seen in Fernandez (2005)

Jagannathan, McGrattan and Scherbina (2000) use the Gordon model, assume that dividends will growth as fast as GNP, and come with an estimate of 3.04%. They mention that *"to get the estimate up to Brealey and Myer's 9.2%, we would need to assume nominal dividend growth of 13.2%. This is an unreasonable assumption"*. They also revise Welch (2000) and point out that *"apparently, finance professors do not expect the equity premium to shrink"*.

Claus and Thomas (2001) calculate the equity premium using the Gordon model and the residual income model, assuming that g is the consensus of the analysts' earnings growth forecasts for the next five years and that the dividend payout will be 50%. They also assume that the residual earnings growth after year 5 will be the current 10-year risk-free rate less 3%. With data from 1985 to 1998, they find that the IEP is smaller than the HEP, and they recommend using a REP of about 3% for the US, Canada, France, Germany, Japan and UK.

Harris and Marston (2001), using the dividend discount model and estimations of the financial analysts about long-run growth in earnings, estimate an IEP of 7.14% for the S&P 500 above T-Bonds over the period 1982-1998. They also claim that the IEP move inversely with government interest rates, which is hard to believe.

Easton, Taylor, Shroff and Sougiannis (2002) used the residual income model with IBES data for expected growth[23], and estimated an average IEP of 5.3% over the years 1981-1998.

Goedhart, Koller and Wessels (2002) used the dividend discount model (considering also share repurchases), with GDP growth as a proxy for expected earnings growth and with the average inflation rate of the last 5 years as a proxy for expected inflation. Table 11 contains their results that they report. They conclude that *"we estimate that the real cost of equity has been remarkably stable at about 7% in the US and 6% in the UK since the 1960s. Given current, real long-term bond yields of 3% in the US and 2.5% in the UK, the implied equity risk premium is around 3.5% to 4% for both markets"*.

Table 11. IEP and real cost of equity in the US and the UK according to Goedhart *et al* (2002)

	US		UK	
	1962-1979	1990-2000	1962-1979	1995-2000
Market risk premium	**5.0%**	**3.6%**	**4.3%**	**3.0%**
Real risk-free rate	2.2%	3.1%	1.4%	2.8%
Real cost of equity	7.2%	6.7%	5.7%	5.8%

Fama and French (2002), using the discounted dividend model, estimated the IEP for the period 1951-2000 between 2.55% and 4.32%, far below the HEP (7.43%). For the period 1872-1950, they estimated an IEP (4.17%) similar to the HEP (4.4%). They claimed that in the period 1951-2000 *"a decline in the expected stock return is the prime source of the unexpected capital gain"*, and that *"the unconditional EEP of the last 50 years is probably far below the realized premium"*[24].

Ritter and Warr (2002) claim that in 1979-1997, the IEP declined from +12% to -4%. However, Ritter estimate of the IEP in 2006 is a little over 2% on a geometric basis.

Harris, Marston, Mishra and O'Brien (2003) estimated discount rates for several companies using the dividend discount model and assuming that g was equal to the consensus of the analysts' growth of dividends per share forecasts. They found an IEP of 7.3% (if betas calculated with a domestic index) and 9.7% (when betas calculated with a world index).

Many authors use an expected growth of dividends per share (g) equal to the consensus of the analysts' forecasts, but Doukas, Kim and Pantzalis (2006) find that stock returns are positively associated with analyst's divergence of opinion, and consider the divergence of opinion as risk.

Vivian (2005) replicated Fama and French (2002) to the UK, obtained similar results (see table 12), and concluded that the discount rate (REP) declined in the later part of the 20th Century.

Table 12. REP and HEP in the US and in the UK according to Fama and French (2002) and Vivian (2005)

Table I of Fama and French (2002)

US	REP	HEP
1872-2000	3.54%	5.57%
1872-1950	4.17%	4.40%
1951-2000	2.55%	7.43%

Table 1 of Vivian (2005)

UK	REP	HEP
1901-2002	4.41%	5.68%
1901-1950	4.22%	3.49%
1951-2002	4.60%	7.79%
1966-2002	3.00%	6.79%

[23] Although Chan, Karceski and Lakonishok (2001) report that *"IBES forecasts are too optimistic and have low predictive power for long-term growth"*.
[24] Fama and French (1992) report that in the period 1941-1990 an equally weighted index outperformed the value weighted (average monthly returns of 1.12% and 0.93%) in the whole period and in most sub sample periods.

O'Hanlon and Steele (2000) proposed calculating the REP using accounting figures and got a variety of estimates between 4 and 6%.

Glassman and Hassett (2000) calculated in their book Dow 36,000 that the REP for the U.S. in 1999 was 3%, arguing that stocks should not carry any risk premium at all, and that stock prices will rise dramatically further once investors come to realize this fact[25].

Faugere and Erlach (2006) claimed that the equity premium tracks the value of a put option on the S&P 500. However, their conclusion is not very helpful: *"using an 8.1% premium in valuation formulas and capital budgeting problems may be appropriate, since the observed level of the long-run equity premium is fully consistent with the observed steady-state GDP growth and consistent with risk explanations as well. However, if one believes that the recent 1990's trends in dividend yields, interest rates, taxes and inflation represent permanent regime shifts, our model can be parameterized to yield a 3.5% equity premium"*.

Donaldson, Kamstra and Kramer (2006) simulate the distribution from which interest rates, dividend growth rates, and equity premia are drawn and claim that *"the true ex ante equity premium is 3.5% pus or minus 50 basis points"*. They say that previous studies *"estimate the equity premium with great imprecision: often a 5% to 6% ex post estimate can not be statistically distinguished from an ex ante value as low as 1% or as high as 10%"*.

One problem of all these estimates is that they depend on the particular assumption made for the expected growth.

Table 13. Implied Equity Premium **(IEP)** and Required Equity Premium **(REP)** according to different authors

Author(s)	Method		IEP = REP
O'Hanlon and Steele (2000)	accounting		4 to 6%
Jagannathan & al (2000)	DDM		3.04%
Glassman and Hasset (2000)			3%
Harris and Marston (2001)	DDM		7.14%
Claus and Thomas (2001)	RIM	1985-1998	3%
Fama and French (2002)	DDM	1951-2000	2.55%
Fama and French (2002)	DDM	1872-1950	4.17%
Goedhart, Koller and Wessels (2002)	DDM	1990-2000	3.5 to 4%
Ritter (2002)	DDM	2001	0.7%
Ritter and Warr (2002)	RIM	1979-1997	+12% to -4%.
Harris & al (2003)	DDM		7.3%
Vivian (2005)	DDM & RIM	1951-2002 UK	4.6%
Ibbotson Associates (2006)	REP=EEP=HEP	1926-2005	7.1%
Donaldson, Kamstra and Kramer (2006)	DDM	1952-2004	3.5%

DDM = dividend discount model. RIM = residual income model

5. The equity premium puzzle

The **equity premium puzzle**, a term coined by Mehra and Prescott (1985), is the inability of a *standard representative consumer asset pricing model*, using aggregate data, to reconcile the HEP. To reconcile the model with the HEP, individuals must have implausibly high risk aversion according to standard economics models[26]. Mehra and Prescott (1985) argued that stocks should provide at most a 0.35% premium over bills. Even by stretching the parameter estimates, Mehra and Prescott (2003) concluded that the premium should be no more than 1%. This contrasted starkly with their HEP estimate of 6.2%.

5.1. Attempts to solve the equity premium puzzle

This puzzle has lead to an extensive research effort in both macroeconomics and finance. Over the last 20 years, researchers have tried to resolve the puzzle by generalizing and adapting (weakening one or more of the assumptions) the Mehra-Prescott (1985) model, but still there is not a solution generally accepted by the economics profession. Some of the adapted assumptions include:

- alternative assumptions about preferences (state separability, leisure, precautionary savings) or generalizations to state-dependent utility functions: Abel (1990); Constantinides (1990); Epstein

[25] Not to be outdone, Kadlec and Acampora (1999) gave their book the title, *Dow 100,000: Fact or Fiction?*

[26] Kocherlakota (1996) reduces the models to just 3 assumptions: individuals have preferences associated with the standard utility function, asset markets are complete (individuals can write insurance contracts against any contingency), and asset trading is costless.

and Zin (1991); Benartzi and Thaler (1995); Bakshi and Chen (1996); Campbell and Cochrane (1999); and Barberis, Huang, and Santos (2001),
- narrow framing[27]: Barberis and Huang (2006),
- probability distributions that admit disastrous events such as fear of catastrophic consumption drops: Rietz (1988); Mehra and Prescott (1988), Barro (2005),
- survivorship bias: Brown, Goetzmann, and Ross (1995),
- liquidity premium: Bansal and Coleman (1996),
- taxes and regulation: McGrattan and Prescott (2005),
- the presence of uninsurable income shocks or incomplete markets: Mankiw (1986); Constantinides and Duffie (1996); Heaton and Lucas (1996) and (1997); Storesletten, Telmer, and Yaron (1999),
- relative volatility of stocks and bonds: Asness (2000)
- limited stock market participation and limited diversification: Saito (1995), Basak and Cuocco (1998), Heaton and Lucas (2000), Vissing-Jorgensen (2002), Gomes and Michaelides (2005),
- distinguishing between the cash flows to equity and aggregate consumption: Brennan and Xia (2001), who claim to be able to justify an equity premium of 6%.
- borrowing constraints: Constantinides, Donaldson, and Mehra (2002),
- other market imperfections: Aiyagari and Gertler (1991); Alvarez and Jermann (2000),
- disentangling the equity premium into its cash flow and discounting components: Bakshi and Chen (2006);
- measurement errors and poor consumption growth proxies: Breeden, Gibbons, and Litzenberger (1989), Mankiw and Zeldes (1991), Ferson and Harvey (1992), Ait-Sahalia, Parker, and Yogo (2004).

There are several excellent surveys of this work, including Kocherlakota (1996), Cochrane (1997) and Mehra and Prescott (2003 and 2006). Kocherlakota (1996) says that *"while there are several plausible explanations for the low level of Treasury returns, the large equity premium is still largely a mystery to economists"*.

Rietz (1988) and Barro (2005) suggest that low-probability disasters, such as a small a large "crash" in consumption, may justify a large equity premium. However, Mehra and Prescott (1988) challenge Rietz to identify such catastrophic events and estimate their probabilities.

McGrattan and Prescott (2005) argue that the 1960-2001 HEP is mainly due to changes in taxes and regulatory policy during this period. They also say that *"Allowing for heterogeneous individuals will also help quantify the effects of increased market participation and diversification that has occurred in the past two decades. Until very recently, mutual funds were a very expensive method of creating a diversified equity portfolio"*.

Limited stock market participation can increase the REP by concentrating stock market risk on a subset of the population. To understand why limited participation may have quantitative significance for the REP, it is useful to review basic facts about the distribution of wealth, and its dynamics over time. Mishel, Bernstein and Allegretto (2006) document that wealth and stock holdings in the U.S. remain highly concentrated in dollar terms: in 2004, the wealthiest 10% held 78.8% of the stocks (84% in 1989 and 76.9% in 2001), and the wealthiest 20% held over 90% of all stocks. Only 48.6% of U.S. households held stocks in 2004 (51.9% in 2001 and 31.7% in 1989) and only 34.9% (40.1% in 2001 and 22.6% in 1989) held stock worth more than $5,000. Of this 34.9%, only 13.5% had direct holdings. Mankiw and Zeldes (1991) reported that 72.4% of the 2998 families in their survey held no stocks at all. Among families that held more than $100,000 in other liquid assets, only 48% held stock. The covariance of stock returns and consumption of the families that hold stocks is triple than that of no stockholders ant it may explain part of the puzzle.

Brennan (2004) highlights the "*democratization of Equity Investment*": *"The increase in the number of participants in equity markets was accompanied by a massive increase in the scale of the equity mutual fund industry: the assets under management rose from $870 per capita in 1989 to over $14,000 per capita in 1999, before declining to a little over $12,000 per capita in 2001. On the other hand, holdings of bond mutual funds grew only from $966 per capita in 1989 to $2887 in 1989. In other*

[27] Narrow framing is the phenomenon documented in experimental settings whereby, when people are offered a new gamble, they sometimes evaluate it in isolation, separately from their other risks.

words, while bond funds roughly tripled, equity funds went up by a factor of over 14!" and *"the share of corporate equity held by mutual funds rose from 6.6% in 1990 to 18.3% in 2000"*.

Heaton and Lucas (2000) introduced Limited Participation and Limited Diversification in an overlapping generations model and concluded that the increases in participation of the past two decades are unlikely to cause a significant reduction in the EEP, but that improved portfolio diversification might explain a fall in the EEP of several percentage points.

There is some promising research on heterogeneity. Abel (1991) hoped that *"incorporating differences among investors or more general attitudes toward risk can explain the various statistical properties of asset returns"*. Levy and Levy (1996) mentioned that the introduction of a small degree of diversity in expectations changed the dynamics of their model and produced more realistic results. Constantinides and Duffie (1996) introduced heterogeneity in the form of uninsurable, persistent and heteroscedastic labor income shocks. Bonaparte (2006) used micro data on households' consumption and provides a new method on estimating asset pricing models, considering each household as living on an island and taking into account its lifetime consumption path. Due to the great deal of heterogeneity across households, he replaced the representative agent with an average agent.

Bakshi and Chen (2006) claim that *"disentangling the equity premium into its cash flow and discounting components produces an economic meaningful equity premium of 7.31%"*.

Shalit and Yitzhaki (2006) show that at equilibrium, heterogeneous investors hold different risky assets in portfolios, and no one must hold the market portfolio.

It is interesting the quotation in Siegel and Thaler (1997): *"no economic theorist has been completely successful in resolving the [equity premium] puzzle"* ... but ... *"most economists we know have a very high proportion of their retirement wealth invested in equities (as we do)"*.

6. The equity premium in the textbooks

This section contains the main messages about the equity premium conveyed in the finance textbooks and valuation books. More details may be found in Fernandez (2006). Figure 6 collects the evolution of the Required Equity Premium (REP) used or recommended by the textbooks and by the academic papers mentioned on previous sections. Table 14 contains the equity premium recommended and used in different editions of several textbooks. Ritter (2002) mentions the use of the historical equity risk premium in textbooks as an estimate of the future as one of the "*The Biggest Mistakes We Teach*". Looking at Figure 6 and at Table 14, it is quite obvious that there is not much consensus, creating a lot of confusion among students and practitioners (and finance authors, also) about the Equity Premium.

Brealey and Myers considered REP = EEP = HEP in the 2nd, 3rd, 4th and 5th editions (1984, 1988, 1991 and 1996), using Ibbotson data that ranged from 8.2 to 8.5% (arithmetic HEPs over T-Bills in periods starting in 1926). In the 6th, 7th and 8th editions (2000, 2003 and 2005 with Allen), they said that *"Brealey, Myers and Allen have no official position on the exact market risk premium, but we believe that a range of 5 to 8.5 percent is reasonable for the risk premium in the United States."* (In the previous editions the ranges was 6 to 8.5%).

Copeland, Koller and Murrin (McKinsey) used a REP = geometric HEP versus Government T-Bonds in the two first editions (1990 and 1995). However, they changed criteria in the 3rd and 4th editions: they advised to use the arithmetic HEP of 2-year returns versus Government T-Bonds reduced by a survivorship bias. In the 1st edition (1990), they recommended 5-6%, in the 2nd edition (1995) they recommended 5-6%, in the 3rd edition (2000) they recommended 4.5-5% ("*we substract a 1.5 to 2% survivorship bias from the long-term arithmetic average of 6.5%"*) and in the 4th edition (Koller, Goedhart and Wessels, 2005) they recommended 3.5-4.5% ("*we subtract a 1% to 2% survivorship bias from the long-term arithmetic average of 5.5%"*).

Ross, Westerfield and Jaffe recommended in all editions they REP = EEP = arithmetic HEP vs. T-Bills, using Ibbotson data. In (1988, 2nd edition), (1993, 3rd edition) and (1996, 4th edition) they recommended 8.5%. In (1999, 5th edition) 9.2%; in (2002, 6th edition) 9.5%; and in (2005, 7th edition) 8.4%.

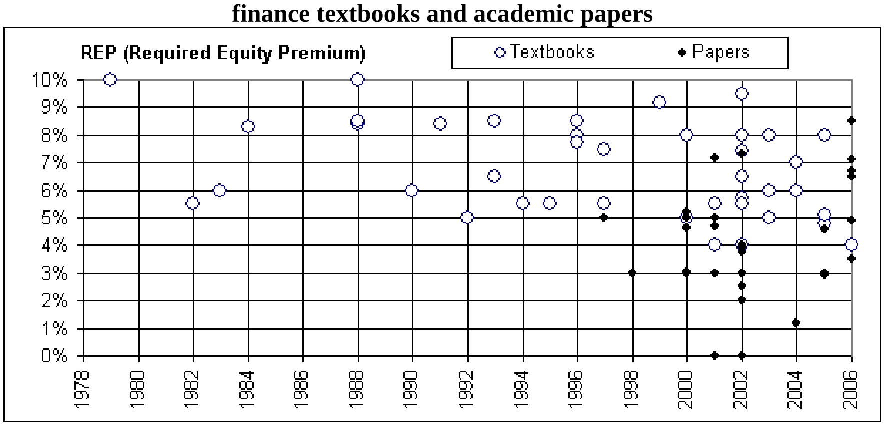
Bodie, Kane and Marcus (1993, 2nd edition) used a REP = EEP = 6.5% to value Hewlett-Packard. In the 3rd edition (1996, page 535), they used a REP = EEP = HEP – 1% = 7.75% to value Motorola. In the 5th edition (2002, page 575), they valued Motorola using a REP = 6.5%. In the 6th edition (2003), they used in the examples different REPs: 8% (pages 426, 431) and 5% (page 415).

Damodaran (1994, 2002) recommended REP = EEP = geometric HEP versus T-bonds. In 1997 he used a REP = arithmetic HEP versus T-Bills. In 2001a and 2006 he recommended REP = EEP = IEP. Damodaran *on Valuation* (1994), recommended an EEP of 5.5%, the geometric HEP using T-bonds for the period 1926-1990. Damodaran (2001a, 2006, 2nd edition) used a REP =IEP of 4% for the US. because *"the implied premium for the US and the average implied equity risk premium has been about 4% over the past 40 years"*. Damodaran (1996, 1997, 2001b, 2001c and 2002), however, used a REP of 5.5%. In (1996. page 48) he shows that 5.5% is the geometric HEP versus T-bonds in the period 1926-90.

Copeland and Weston (1979, 1988) used a REP = 10%. However, Weston and Copeland (1992), used a REP = 5%.

Van Horne (1968, 1st ed.) still did not mention the CAPM or the equity premium. In (1983, 6th ed.), he used a REP = 6% He justified it: *"Suppose, for easy illustration, that the expected risk-free rate is an average of the risk-free rates that prevailed over the ten-year period and that the expected market return is average of market returns over that period"*. In (1992, Fundamentals, 8th ed.), he used a REP = 5% and justified it: *"Assume that a rate of return of about 13% on stocks in general is expected to prevail and that a risk-free rate of 8% is expected"*.

Figure 6. Evolution of the Required Equity Premium (REP) used or recommended in the most important finance textbooks and academic papers



Penman (2001, 1st ed.) said that *"the market risk premium is a big guess. Research papers and textbooks estimate it in the range of 4.5% to 9.2%. … No one knows what the market risk premium is"*. In (2003, 2nd ed.), he admitted that *"we really do not have a sound method to estimate the cost of capital… Estimates [of the equity premium] range, in texts and academic research, from 3.0% to 9.2%"*, and he used 6%.

Weston and Brigham (1968) still did not defined equity premium. In (1982, 6th edition) they said that *"the market risk premium can be considered relatively stable at 5 to 6% for practical application"*. Weston, Chung and Siu (1997) recommended 7.5%. Bodie and Merton (2000) used 8% for USA.

Stowe, Robinson, Pinto and McLeavey (2002), in their book for the CFA (Chartered Financial Analysts) Program use (page 49) a REP = Geometric HEP using T-Bonds during 1926-2000, according to Ibbotson = 5.7%. Pratt (2002) assumes that REP=EEP=HEP and uses 7.4% (page 68) and 8% (page 74). Hawawini and Viallet (2002) use a REP = 6.2% = geometric HEP over T-bonds in the period 1926-1999 according to Ibbotson.

Fernandez (2002) is the only finance textbook claiming that *"it is impossible to determine the premium for the market as a whole, because it does not exist"*. He also mentions that we *"could only talk of a market risk premium if all investors had the same cash flow expectations… However, expectations are not homogeneous"*. Fernandez (2004, 2001) also mentioned that "*the HEP, the EEP and the REP are different concepts*" and that "*different investors have different REPs*". In the examples he uses REP = 4%.

Table 14. Equity premiums recommended and used in textbooks

Author(s) of the Textbook	Assumption	Period for HEP	REP recommended	REP used
Brealey and Myers				

2nd edition. 1984	REP=EEP= arith HEP vs. T-Bills	1926-81	8.3%	8.3%
3rd edition. 1988	REP=EEP= arith HEP vs. T-Bills	1926-85	8.4%	8.4%
4th edition. 1991	REP=EEP= arith HEP vs. T-Bills	1926-88	8.4%	8.4%
5th edition. 1996	REP=EEP= arith HEP vs. T-Bills		8.2 - 8.5%	
6th and 7th edition. 2000 and 2003	No official position		6.0 - 8.5%	8.0%
8th edition. 2005 (with Allen)	No official position		5.0 - 8.5%	
Copeland, Koller and Murrin (McKinsey)				
1st edition. 1990	REP=EEP= geo HEP vs. T-Bonds	1926-88	5 - 6%	6%
2nd ed. 1995	REP=EEP= geo HEP vs. T-Bonds	1926-92	5 - 6%	5.5%
3rd ed. 2000	REP=EEP= arith HEP – 1.5-2%	1926-98	4.5 - 5%	5%
4th ed. 2005. Goedhart, Koller & Wessels	REP=EEP= arith HEP – 1-2%	1903-2002	3.5 – 4.5%	4.8%
Ross, Westerfield and Jaffe				
2nd edition. 1988	REP=EEP= arith HEP vs. T-Bills	1926-88	8.5%	8.5%
3rd edition. 1993	REP=EEP= arith HEP vs. T-Bills	1926-93	8.5%	8.5%
4th edition. 1996	REP=EEP= arith HEP vs. T-Bills	1926-94	8.5%	8.5%
5th edition. 1999	REP=EEP= arith HEP vs. T-Bills	1926-97	9.2%	9.2%
6th edition. 2002	REP=EEP= arith HEP vs. T-Bills	1926-99	9.5%	9.5%
7th edition. 2005	REP=EEP= arith HEP vs. T-Bills	1926-02	8.4%	8%
Van Horne, 6th edition. 1983				6.0%
8th edition. 1992			3 - 7%	5.0%
Copeland and Weston (1979 and 1988)				10%
Weston and Copeland (1992)				5%
Bodie, Kane and Marcus				
2nd edition. 1993	REP=EEP		6.5%	6.5%
3rd edition. 1996	REP=EEP=arith HEP vs. T-Bills - 1%		7.75%	7.75%
5th edition. 2002			6.5%	6.5%
2003	REP=EEP= arith HEP vs. T-Bills	1926-2001		5%; 8%
Damodaran 1994 Valuation. 1st ed.	REP=EEP= geo HEP vs.T-Bonds	1926-90	5.5%	5.5%
1996, 1997, 2001b, 2001c	REP=EEP= geo HEP vs.T-Bonds		5.5%	5.5%
2001a	average IEP	1970-2000	4%	4%
2002	REP=EEP= geo HEP vs.T-Bonds	1928-00	5.51%	5.51%
2006 Valuation. 2nd ed.	REP=EEP= geo HEP vs.T-Bonds	1928-2004	4.84%	4%
Weston & Brigham (1982)			5-6%	
Weston, Chung and Siu (1997)			7.5%	
Bodie and Merton (2000)				8%
Stowe *et al* (2002)	REP=EEP= geo HEP vs.T-Bonds	1926-00	5.7%	5.7%
Hawawini and Viallet (2002)	REP=EEP= geo HEP vs.T-Bonds	1926-99		6.2%
Pratt (2002)	REP=EEP=HEP			7.4%, 8%
Fernandez (2002)	"is impossible to determine the premium for the market as a whole"			
Penman (2003)	"No one knows what the REP is"			6%
Fernandez (2001, 2004)	"different investors have different REPs"			4%
Bruner (2004)	REP=EEP= geo HEP vs.T-Bonds	1926-2000	6%	6%
Palepu, Healy and Bernard (2004)	REP=EEP= arith HEP vs.T-Bonds	1926-2002	7%	7%
Weston, Mitchel & Mulherin (2004)	REP=EEP= arith HEP vs.T-Bonds	1926-2000	7.3%	7%
Arzac (2005)	REP=IEP		5.08%	5.08%

Palepu, Healy and Bernard (2004, page 8-3) mention that the HEP "*constitutes an estimate of the REP*" and use REP = 7% in the examples (page 8-5).

Weston, Mitchel and Mulherin (2004) mention that the arithmetic HEP over T-bonds in the period 1926-2000 according to Ibbotson was 7.3% and (page 260) they use REP = EEP = 7%.

Bruner (2004) used a REP of 6% because *"from 1926 to 2000, the risk premium for common stocks has averaged about 6% when measured geometrically"*.

Arzac (2005) uses a REP = IEP = 5.08% for a valuation done in December 2002 (the IEP equity premium as of that date calculated using the Gordon equation).

In the following section we claim that the confusion comes from the fact that **there is not a REP** for the market as a whole: different investors use different **REPs**. Last sentence may me rewritten as: **there is not an IEP** for the market as a whole: different investors use different **IEPs**. A unique IEP requires assuming homogeneous expectations for the expected growth (g), but there are several pairs (IEP, g) that satisfy current prices.

7. There is not an IEP, but many pairs (IEP, g) which are consistent with market prices

Even if market prices are correct for all investors, there is not a unique REP common for all investors. In a simple Gordon model, there are many pairs (Ke, g) that satisfy equation (1). As Ke is the sum of the Implied Equity Premium (IEP) plus the risk-free rate (R_F), there are many pairs (IEP, g) that satisfy equation (1). A unique IEP requires assuming homogeneous expectations for the expected growth (g). If equation (1) holds, the *expected* return for the shareholders is equal to the *required* return for the shareholders (Ke), but there are many *required* returns (as many as expected growths, g) in the market. On top of that, IEP and g change over time.

If investors' expectations were homogenous, it would make sense to calculate a unique IEP, as all investors would have the market portfolio and the same expectations regarding the portfolio[28]. However, as expectations are not homogenous[29], different investors use different **REPs**: investors who expect higher growth will have a higher REP. Heterogeneous investors do not hold the same portfolio of risky assets; in fact, no investor must hold the market portfolio to clear the market: it does not make sense to search for a common REP because it does not exists.

We can find out an investor's REP by asking him, although for many investors the REP is not an explicit parameter but, rather, an implicit one that manifests in the price they are prepared to pay for shares[30]. However, it is impossible to determine the REP for the market as a whole, because it does not exist. Even if we knew the market premiums of all the investors who operated on the market, it would be meaningless to talk of a premium for the market as a whole.

A rationale for this may be found in the aggregation theorems of microeconomics, which in actual fact are non-aggregation theorems. One model that works well individually for a number of people may not work for all of the people together[31]. For the CAPM, this means that although the CAPM may be a valid model for each investor, it is not valid for the market as a whole, because investors do not have the same return and risk expectations for all shares. Prices are a statement of expected cash flows discounted at a rate that includes the risk premium. Different investors have different cash flow expectations and different future risk expectations. One could only talk of an equity premium if all investors had the same cash flow expectations.

Reallocating terms in equation (1), we get:

$$IEP - g = d_1/P_0 - R_F \tag{3}$$

There are many pairs (IEP, g) that satisfy the Gordon equation at any moment. All the papers that we revised on section 5 assume that there is an "expected growth rate for the market" and get an "IEP for the market". But without homogeneous expectations, there is not an "expected growth rate for the market".

Similarly, for having an EEP common for all investors we need to assume homogeneous expectations (or a representative investor) and, with our knowledge of financial markets, this assumption is not reasonable. A theory with a representative investor cannot explain either why the annual trading volume of most exchanges more than double the market capitalization.

We also find that the difference (IEP – g),[32] is related to the risk free rate in the period after 1960. Figure 7 shows the relationship for the period after 1980 for the US, Spain and the UK. It may be seen the high negative correlation between (IEP – g) and the risk free rate in the three markets. Table 15 presents the regressions for more countries.

Figure 7. Correlations (d_1/P_0 - R_F) – (R_F) for the US, Spain and the UK. Monthly data.
(d_1/P_0 - R_F) = IEP – g. *Source of the data: Datastream*

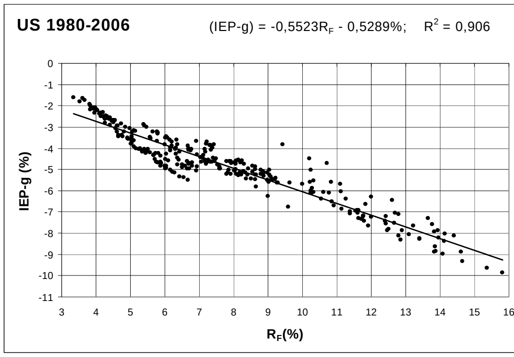
[28] Even then, this method requires knowing the expected growth of dividends. A higher growth estimate implies a higher premium.

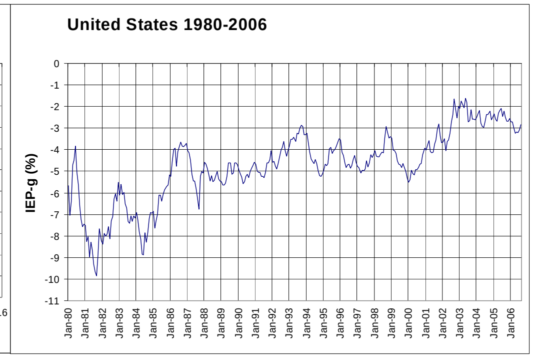
[29] Doukas, Kim and Pantzalis (2006) document analysts' divergence of opinion.

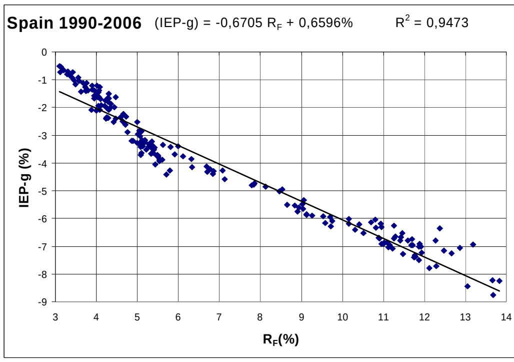
[30] An example: An investor is prepared to pay 80 euros for a perpetual annual cash flow of 6 euros in year 1 and growing at an annual rate of 3%, which he expects to obtain from a diversified equity portfolio. This means that his required market return is 10.5% ([6/80] + 0.03).

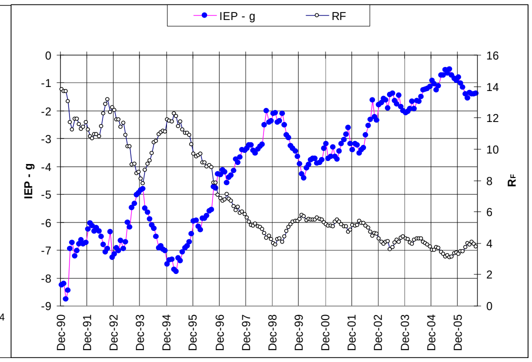
[31] As Mas-Colell *et al.* (1995, page 120) say, "it is not true that whenever aggregate demand can be generated by a representative consumer, this representative consumer's preferences have normative contents. It may even be the case that a positive representative consumer exists but that there is no social welfare function that leads to a normative representative consumer."

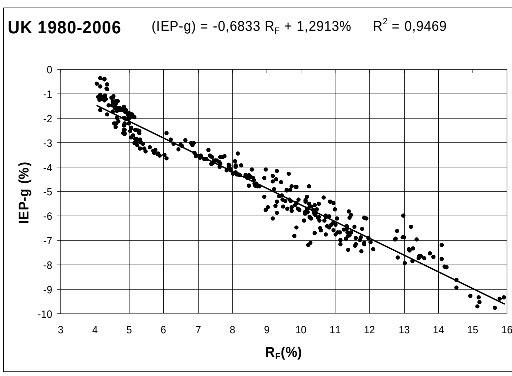
[32] (d_1/P_0 - R_F) is equal to (IEP – g)











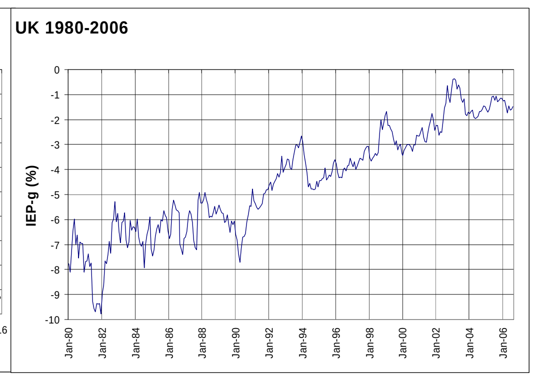


Table 15. Regressions with monthly data of Y (IEP – g) on R_F (10 year Gov. Bond Yield)
Monthly data. (d_1/P_0 - R_F) = IEP – g. *Source of the data: Datastream*

	Full period	(R squared)		Without 1997-02	(R squared)
USA 1980-2006	Y = -0.5523 R_F - 0.5289%	0.9060		Y = -0.5864 R_F - 0.1278%	0.9417
Germany 1980-2006	Y = -0.7192 R_F + 0.5907%	0.8205		Y = -0.7569 R_F + 0.9362%	0.8427
UK 1980-2006	Y = -0.6833 R_F + 1.2913%	0.9469		Y = -0.7195 R_F + 1.7119%	0.9551
France 1988-2006	Y = -0.9587 R_F + 2.5862%	0.9245		Y = -1.0273 R_F + 3.2364%	0.9625
Italy 1991-2006	Y = -1.0693 R_F + 3.0398%	0.9563		Y = -1.1223 R_F + 3.7155%	0.9730
Spain 1991-2006	Y = -0.6705 R_F + 0.6596%	0.9473		Y = -0.7135 R_F + 1.1954%	0.9747

8. How do I calculate the REP?

For calculating the cost of equity (required return to equity cash flows) of a company, a valuator has to answer the following question: which differential rate over current T-Bond yields do I think compensates the risk of holding the shares? If there is only an owner of the shares, we can directly ask him the question. But if it is a traded company, the valuator has to make a prudential judgment. As Grabowski (2006), points out, *"the entire appraisal process is based on applying reasoned judgment to the evidence derived from economic, financial and other information and arriving at a well reasoned opinion of value"*.

We need the cost of equity to discount the expected equity cash flows of the company. Note that there is a kind of schizophrenic approach to valuation: while all authors admit that different valuators and investors may have different expectations of equity cash flows, most authors look for a unique discount rate. It seems as if the expectations of equity cash flows are formed in a democratic regime, while the discount rate is determined in a dictatorship. In any market, different investors may have different expectations of equity cash flows and different evaluations of its risk (that translate into different discount rates). Then, in the case of a traded company, there are investors that think that the company is undervalued (and buy or hold shares), investors that think that the company is overvalued (and sell or not buy shares), and investors that think that the company is fairly valued (and sell or hold shares). The investors that did the last trade, or the rest of the investors that held or did not have shares do not have a common REP (nor common expectations of equity cash flows).

For calculating the REP, we must answer the same question, but thinking in a diversified portfolio of shares, instead in just the shares of a company. In the valuations that I have done in the 21st century I have used REPs between 3.8 and 4% for Europe and for the U.S. Given the yields of the T-Bonds, I think[33] that an additional 4% compensates the additional risk of a diversified portfolio.

9. Conclusion

The equity premium (also called *market risk premium*, *equity risk premium*, *market premium* and *risk premium*), is one of the most important, discussed but elusive parameters in finance. Much of the confusion arises from the fact that the term equity premium is used to designate four different concepts (although many times they are mixed): Historical Equity Premium (HEP), Expected Equity Premium (EEP); Required Equity Premium (REP) and Implied Equity Premium (IEP).

In the finance literature and in valuation textbooks, there are authors that claim different identities among the four equity premiums defined above: some claim that **HEP = EEP = REP**; others claim that **EEP is smaller than HEP;** others claim that there is **a unique IEP and that REP = IEP**; others ***"have no official position"***; others claim that **EEP is near zero;** others try to find the **EEP doing surveys;** others affirm ***"that no one knows what the REP is"***.

The **HEP** is equal for all investors, but the **REP,** the **EEP and** the **IEP** are different for different investors. There is no an **IEP** for the market as a whole: different investors have different **IEPs** and use different **REPs**. A unique IEP requires assuming homogeneous expectations for the expected growth (g), but there several pairs (IEP, g) that satisfy current prices.

We claim that different investors have different REPs and that it is impossible to determine the REP for the market as a whole, because it does not exist. Heterogeneous investors do not hold the same portfolio of risky assets; in fact, no investor must hold the market portfolio to reach equilibrium.

There is a kind of schizophrenic approach to valuation: while all authors admit that different valuators and investors may have different expectations of equity cash flows, most authors look for a unique discount rate. It seems as if the expectations of equity cash flows are formed in a democratic regime, while the discount rate is determined in a dictatorship. In any market, different investors may have different expectations of equity cash flows and different evaluations of its risk (that translate into different discount rates).

It has been argued that, from an economic standpoint, we need to establish the primacy of the EEP, since it is what guides investors' decisions. However, the REP is more important for many important decisions, among others, valuations of projects and companies, acquisitions, and corporate investment decisions. On the other hand, EEP is important only for the investors that hold the market portfolio.

For calculating the cost of equity (required return to equity cash flows) of a company, a valuator has to answer the following question: which differential rate over current T-Bond yields do I think compensates the risk of holding the shares? If there is only an owner of the shares, we can directly ask him the question. But if it is a traded company, the valuator has to make a prudential judgment. There are investors that think that the company is undervalued (and buy or hold shares), investors that

[33] And also my clients that are able to answer to that question.

think that the company is overvalued (and sell or not buy shares), and investors that think that the company is fairly valued (and sell or hold shares). For calculating the REP, we must answer the same question, but thinking in a diversified portfolio of shares, instead in just the shares of a company. Recently, I have used REPs between 3.8 and 4% for Europe and for the U.S. Given the yields of the T-Bonds, I think that an additional 4% compensates the additional risk of a diversified portfolio.

References

Abel, A.B. (1990), "Asset Prices under Habit Formation and Catching Up with the Joneses", *American Economic Review*, vol. 80, no. 2 (May), pp. 38–42.

Abel, A.B. (1991), "The Equity Premium Puzzle", *Business Review* (Federal Reserve Bank of Philadelphia), pp. 3-14.

Ait-Sahalia, Y., J. A. Parker and M. Yogo (2004), "Luxury Goods and the Equity Premium", *Journal of Finance* 59, pp. 2959-3004.

Aiyagari, S.R., and M. Gertler (1991), "Asset Returns with Transactions Costs and Uninsured Individual Risk", *Journal of Monetary Economics*, vol. 27, no. 3 (June), pp. 311–331.

Alvarez, F., and U. Jermann (2000), "Efficiency, Equilibrium, and Asset Pricing with Risk of Default", *Econometrica*, vol. 68, no. 4 (July), pp.775–797.

Ang, A. and G. Bekaert (2001), "Stock Return Predictability: Is It There?", Unpublished paper, Graduate School of Business, Columbia University.

Arnott, R. D. and R. J. Ryan (2001), "The Death of the Risk Premium", *Journal of Portfolio Management*, Spring, pp. 61-74.

Arnott, Robert D. and Peter L. Bernstein (2002), "What Risk Premium is 'Normal'?", *Financial Analysts Journal*, Vol. 58, No. 2, pp. 64-84.

Arzac, Enrique R. (2005), *Valuation for Mergers, Buyouts, and Restructuring*, John Wiley & Sons Inc.

Asness, C.S. (2000), "Stocks versus Bonds: Explaining the Equity Risk Premium", *Financial Analysts Journal*, March/April, Vol. 56, No. 2, pp.96-113.

Avramov, D. (2002), "Stock Return Predictability and Model Uncertainty", *Journal of Financial Economics* (3), pp. 423–458.

Baker, M and J. Wurgler (2000), "The Equity Share in New Issues and Aggregate Stock Returns", *Journal of Finance,* 55, 5, pp. 2219-57.

Bakshi, G. and Z. Chen (1996), "The Spirit of Capitalism and Stock Market Prices", *American Economic Review* 86, pp.133-157.

Bakshi, G. and Z. Chen (2006), "Cash Flow Risk, Discounting Risk, and the Equity Premium Puzzle", *Handbook of Investments: Equity Premium*, edited by Rajnish Mehra.

Ball, R. (1978), "Anomalies in Relationships Between Securities' Yields and Yield Surrogates", *Journal of Financial Economics* 6, pp. 103—126.

Bansal, R., and J.W. Coleman (1996), "A Monetary Explanation of the Equity Premium, Term Premium, and Risk-Free Rate Puzzles", *Journal of Political Economy*, vol. 104, no. 6 (December), pp.1135–71.

Barberis, N. and M. Huang (2006), "The Loss Aversion/Narrow Framing Approach to the Equity Premium Puzzle", *Handbook of Investments: Equity Premium*, edited by Rajnish Mehra.

Barberis, N., M. Huang and J. Santos (2001), "Prospect Theory and Asset Prices", *Quarterly Journal of Economics*, vol. 116, pp.1–53.

Barro, R. J. (2005), "Rare Events and the Equity Premium", NBER Working Paper No. W11310.

Basak, S., and D. Cuoco (1998), "An Equilibrium Model with Restricted Stock Market Participation", *Review of Financial Studies* 11, pp. 309-341.

Benartzi, S., and R.H. Thaler (1995), "Myopic Loss Aversion and the Equity Premium Puzzle", *Quarterly Journal of Economics*, vol. 110, no. 1 (February), pp.73–92.

Blanchard, O. (1993), "The Vanishing Equity Premium", *Finance and the International Economy* 7, pp. 23-39.

Bodie, Z. and R. Merton (2000), *Finance*, New Jersey: Prentice Hall.

Bodie, Z., A. Kane, and A. J. Marcus (2004), *Investments*, 6th edition. NY: McGraw Hill. (6th edition in 2004). Previous editions: 1989, 1993, 1996, 1999, 2002.

Bodie, Z., A. Kane, and A. J. Marcus (2003), *Essentials of Investments*, 5th edition. NY: McGraw Hill.

Bonaparte, Y. (2006), "The Equity Premium Puzzle: A Reconciliation", The University of Texas at Austin, Available at SSRN: paper n. 889410.

Booth, L. (1999), "Estimating the Equity Risk Premium and Equity Costs: New Ways of Looking at Old Data", *Journal of Applied Corporate Finance*, Vol. 12 No. 1, pp. 100-112.

Bostock, P. (2004), "The Equity Premium", *Journal of Portfolio Management* 30(2), pp. 104-111.

Boudoukh, J., R. Michaely, M.P. Richardson and M.R. Roberts (2006),"On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing", Forthcoming, *Journal of Finance*

Brealey, R.A. and S.C. Myers (2003), *Principles of Corporate Finance*, 7th edition, New York: McGraw-Hill. Previous editions: 1981, 1984, 1988, 1991, 1996 and 2000.

Brealey, R.A., S.C. Myers and F. Allen (2005), *Principles of Corporate Finance*, 8th edition, McGraw-Hill/Irwin.

Breeden, D., M. Gibbons and R. Litzenberger (1989), "Empirical Test of the Consumption-Oriented CAPM", *Journal of Finance* 44 (2), pp.231-262.

Brennan, M. (2004), "How Did It Happen?", *Economic Notes* by Banca Monte dei Paschi di Siena, vol. 33, no. 1, pp. 3–22.

Brennan, M. and Y. Xia (2001), "Stock Price Volatility and the Equity Premium", *Journal of Monetary Economics* 47, pp. 249-283

Brown, S. J., W. N. Goetzmann and S. A. Ross (1995), "Survival", *Journal of Finance*, July, pp. 853-873.

Bruner, Robert F. (2004), *Applied Mergers and Acquisitions*, NY: John Wiley & Sons.

Campbell, J. Y. (1987), "Stock Returns and the Term Structure", *Journal of Financial Economics* 18, pp. 373—399.

Campbell, J. Y. (2002), *TIAA-CREF Investment Forum,* June.

Campbell, J. Y. (2007), "Estimating the Equity Premium", NBER Working Papers from National Bureau of Economic Research (NBER) n. 13423

Campbell, J.Y. and J.H. Cochrane (1999), "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior", *Journal of Political Economy*, vol. 107, no. 2 (April), pp. 205– 251.

Campbell, J. Y., and R. J. Shiller (1988), "Stock Prices, Earnings, and Expected Dividends", *Journal of Finance*, 43(3), pp. 661-676.

Campbell, J. Y. and S.B. Thompson (2005), "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?", forthcoming *Review of Financial Studies*

Campbell, J. Y. and L. M. Viceira (2002), *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors,* New York: Oxford University Press.

Campbell, J. Y. and T. Vuolteenaho (2004), "Inflation Illusion and Stock Prices", *American Economic Review*, vol. 94(2), pp. 19-23.

Campbell, J. Y and M. Yogo (2003), "Efficient Tests of Stock Return Predictability", Working Paper, Harvard University and NBER.

Chan, L., J. Karceski and J. Lakonishok (2001), "The Level and Persistence of Growth Rates", NBER Working Paper 8282.

Claus, J.J. and J.K. Thomas (2001), "Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets", *Journal of Finance*. 55, (5), pp. 1629-66.

Cochrane, J.H. (1997), "Where is the Market Going? Uncertain Facts and Novel Theories", *Economic Perspectives*, Vol. 21, pp. 3–37.

Cohen, R.B, C. Polk, and T. Vuolteenaho (2005), "Money Illusion in the Stock Market: The Modigliani-Cohn Hypothesis," *Quarterly Journal of Economics* 120(2), pp. 639–68.

Constantinides, G.M (1990) "Habit Formation: A Resolution of the Equity Premium Puzzle", *Journal of Political Economy*, vol. 98, no. 3 (June), pp.519-543.

Constantinides, G.M. (2002), "Rational Asset Prices", *Journal of Finance* 57(4), pp. 1567-1591.

Constantinides, G. M. and D. Duffie (1996), "Asset Pricing with Heterogeneous Consumers", *Journal of Political Economy* 104, 2, pp. 219-240.

Constantinides, G.M., J.B. Donaldson, and R. Mehra (2002), "Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle", *Quarterly Journal of Economics*, vol. 117, no. 1 (February), pp. 269-296.

Copeland, T. E., T. Koller, and J. Murrin (2000), *Valuation: Measuring and Managing the Value of Companies.* 3rd edition. New York: Wiley. Previous editions: 1990 and 1995.

Copeland, T. E., and J. F. Weston (1988), *Financial Theory and Corporate Policy.* 3rd edition, Reading, MA: Addison-Wesley. First edition: 1979.

Cowles, A. (1939), *Common Stock Indexes,* Principia Press, Bloomington, Indiana.

Damodaran, A. (2006), *Damodaran on Valuation*, 2nd edition. New York: John Wiley and Sons. 1st edition: (1994).

Damodaran, A. (2001a), *The Dark Side of Valuation*, New York: Prentice-Hall.

Damodaran, A. (2001b), *Corporate Finance: Theory and Practice*. 2nd edition, New York: John Wiley and Sons.

Damodaran, A. (2001c), *Corporate Finance: Theory and Practice*. 2nd *international* edition, New York: John Wiley and Sons.

Damodaran, A. (2002), *Investment Valuation*, New York: John Wiley and Sons.

Dimson, E. and P. Marsh (2001), "U.K. Financial Market Returns, 1955-2000", *Journal of Business* 74, n.1, pp. 1-31.

Dimson, E., P. Marsh and M. Staunton (2000), The Millennium Book: A Century of Returns. ABN AMRO/London Business School.

Dimson, E., P. Marsh and M. Staunton (2002), *Triumph of the Optimists: 101 Years of Global Investment Returns*. New Jersey: Princeton University Press.

Dimson, E., P. Marsh and M. Staunton (2003), "Global Evidence on the Equity Risk Premium", *Journal of Applied Corporate Finance*, 15:4, pp. 27-38.

Dimson, E., P. Marsh and M. Staunton (2004), "Irrational Optimism", *Financial Analysts Journal*, Vol. 60, No. 1, pp. 15–25.

Dimson, E., P. Marsh and M. Staunton (2006a), *Global Investment Returns Yearbook 2006*. ABN AMRO/London Business School.

Dimson, E., P. Marsh and M. Staunton (2006b), *DMS Global Returns data module*. Chicago, IL: Ibbotson Associates.

Dimson, E., P. Marsh and M. Staunton (2006c), "The Worldwide Equity Premium: A Smaller Puzzle", SSRN Working Paper No. 891620.

Donaldson, G., M. J. Kamstra, and L. A. Kramer (2006), "Estimating the Ex Ante Equity Premium", Rotman School of Management Working Paper. Available at SSRN: http://ssrn.com/abstract=945192

Doukas, J.A., C.F. Kim and C. Pantzalis (2006), "Divergence of Opinion and Equity Returns", *Journal of Financial and Quantitative Analysis* 41 (3), pp. 573-606.

Easton, P., G. Taylor, P. Shroff and T. Sougiannis (2002), "Using Forecasts of Earnings to Simultaneously Estimate Growth and the Rate of Return on Equity Investment", *Journal of Accounting Research* 40(3), pp. 657-676.

Epstein, L. G. and S.E. Zin (1991), "The Independence Axiom and Asset Returns", NBER Working Paper No. T0109.

Fama, E.F. (1981), "Stock Returns, Real Activity, Inflation and Money", *American Economic Review*, 4, pp. 545 - 65.

Fama, E.F. and K. R. French (1988), "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics* 22, pp. 3-27.

Fama, E. F. and K. R. French (1989), "Business Conditions and Expected Returns on Stocks and Bonds", *Journal of Financial Economics* 25, pp. 23-49.

Fama, E.F. and K.R. French (1992), "The Cross-Section of Expected Stock Returns", *Journal of Finance* 47, pp. 427-466.

Fama, E.F. and K.R. French (2002), "The Equity Risk Premium", *Journal of Finance* 57 no. 2, pp. 637-659.

Fama, E.F. and G. W. Schwert (1977), "Asset Returns and Inflation", *Journal of Financial Economics* 5, pp. 115-146

Faugere, C. and J. Van Erlach (2006), "The Equity Premium: Consistent with GDP Growth and Portfolio Insurance", *Financial Review,* Vol. 41, No. 4, November.

Fernandez, P. (2002), *Valuation Methods and Shareholder Value Creation*. Academic Press, San Diego, CA.

Fernandez, P. (2004), *Valoración de Empresas*. 3rd edition. Ediciones Gestion 2000. Spain. 2nd edition: 2001.

Fernandez, P. (2005), "Equivalence of Ten Different Methods for Valuing Companies by Cash Flow Discounting", *International Journal of Finance Education,* Vol. 1 (1), pp. 141-168

Fernandez, P. (2006), "The Equity Premium in Finance and Valuation Textbooks", IESE Working Paper n. 657, 15 pages.

Ferson, W. E., and C. R. Harvey (1992), "Seasonality and Consumption–Based Asset Pricing", *Journal of Finance* 47, pp. 511-552.

Fisher, L. and J. Lorie (1964), "Rates of Return on Investments in Common Stock", *Journal of Business*, vol. 37, pp. 1-21.

Glassman, J. K. and K. A. Hassett (2000), *Dow 36.000: The new strategy for profiting from the coming rise in the stock market*, Three Rivers.

Goedhart, M., T. Koller and D. Wessels (2002), "The Real Cost of Equity", McKinsey & Company, N.5, Autumn, pp. 11-15.

Goetzmann, W.N. and R.G. Ibbotson (2006), "History and the Equity Risk Premium", in R. Mehra (Ed.), *Handbook of Investments: Equity Risk Premium*. Amsterdam: Elsevier.

Goetzmann, W.N., R.G. Ibbotson and L. Peng (2001), "A New Historical Database for the NYSE 1815 to 1925: Performance and Predictability", *Journal of Financial Markets*, Vol. 4, No. 1, pp. 1–32.

Gomes, F. and A. Michaelides (2005), "Asset Pricing with Limited Risk Sharing and Heterogeneous Agents", manuscript, London School of Economics.

Gordon, M. (1962), *The Investment, Financing and Valuation of the Corporation,* Homewood, IL. Irwin.

Goyal, A. and I. Welch (2003), "Predicting the Equity Premium with Dividend Ratios", *Management Science* 49(5), 639–654.

Goyal, A. and I. Welch (2007), "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction", *Review of Financial Studies*, forthcoming.

Graham, J.R. and C.R. Harvey (2005), "The Equity Risk Premium in September 2005: Evidence from the Global CFO Outlook Survey", Working Paper, Duke University, September.

Graham, J.R. and C.R. Harvey (2007), "The Equity Risk Premium in January 2007: Evidence from the Global CFO Outlook Survey", Working Paper, Duke University, Available at SSRN: http://ssrn.com/abstract=959703

Grabowski, R.J. (2006), "Equity Risk Premium: 2006 Update", *Business Valuation Review*, Summer.

Harris, R.S. and F.C. Marston (2001), "The Market Risk Premium: Expectational Estimates Using Analysts' Forecasts", *Journal of Applied Finance*, Vol. 11.

Harris, R.S., F.C. Marston, D.R. Mishra and T.J. O'Brien (2003), "Ex Ante Cost of Equity Estimates of S&P 500 Firms: The Choice Between Global and Domestic CAPM", *Financial Management*, Vol. 32, No. 3, Autumn.

Hawawini, G. and C. Viallet (2002), *Finance for Executives*, 2nd edition, South-Western, Thompson Learning.

Heaton, J. and D.J. Lucas (1996), "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing", *Journal of Political Economy*, vol. 104, no. 3 (June):443–487.

Heaton, J. and D.J. Lucas (1997), "Market Frictions, Savings Behavior and Portfolio Choice", *Journal of Macroeconomic Dynamics,* vol. 1, no. 1:76–101.

Heaton, J. and D.J. Lucas (2000), "Stock Prices and Fundamentals", *NBER Macroeconomics Annual*.

Hodrick, R. J. (1992), "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement", *Review of Financial Studies* 5, pp. 357—386.

Ibbotson Associates (2006), *Stocks, Bonds, Bills, and Inflation, Valuation Edition, 2006 Yearbook.*

Ibbotson, R. (2002), *TIAA-CREF Investment Forum,* June.

Ibbotson, R. and P. Chen (2003), "Long-Run Stock Returns: Participating in the Real Economy", *Financial Analysts Journal*, Vol. 59, No. 1, pp. 88-98.

Ilmanen, A. (2003), "Expected returns on stocks and bonds", *Journal of Portfolio Management* 29, pp. 7-27.

Jacquier, E., A. Kane, and A.J. Marcus (2003), "Geometric or Arithmetic Mean: A Reconsideration", *Financial Analysts Journal*, Vol. 59, No. 6, pp. 46-53.

Jagannathan, R., E.R. McGrattan, and A.D. Shcherbina (2000), "The Declining U.S. Equity Premium", *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 24, pp. 3–19.

Jorion, P. and W. N. Goetzmann (1999), "Global Stock Markets in the Twentieth Century", *Journal of Finance* 54 (June), pp. 953-80.

Kadlec, C. and R. Acampora (1999), *Dow 100,000: Fact or Fiction?*, New Jersey: Prentice Hall.

Keim, D. and R. Stambaugh (1986), "Predicting Returns in the Stock and Bond Markets", *Journal of Financial Economics* 17, pp. 357—390.

Kocherlakota, N.R. (1996), "The Equity Premium: It's Still a Puzzle", *Journal of Economic Literature*, Vol. 34, pp. 42-71.

Koller, T., M. Goedhart and D. Wessels (2005), *Valuation: Measuring and Managing the Value of Companies*, 4th Edition, McKinsey & Company, Inc. Wiley.

Kothari, S. and J. Shanken (1997), "Book-to-Market, Dividend Yield, and Expected Market Returns: a Time-Series Analysis", *Journal of Financial Economics* 44, pp.169-203.

Lamont, Owen (1998), "Earnings and Expected Returns", *Journal of Finance* 53 (5), pp. 1563-1587.

Lettau, M., and S. C. Ludvigson (2001): "Consumption, Aggregate Wealth and Expected Stock Returns", *Journal of Finance* 56(3), pp. 815-849.

Levy, M. and H. Levy (1996), "The Danger of Assuming Homogeneous Expectations", *Financial Analysts Journal*, May/June, pp. 65-70.

Lewellen, J. (2004), "Predicting Returns with Financial Ratios", *Journal of Financial Economics* 74, pp. 209–235.

Li, H. and Y. Xu (2002), "Survival Bias and the Equity Premium Puzzle", *Journal of Finance* 57, pp. 1981-1993.

Maheu, J. M. and T.H. McCurdy (2006), "How Useful are Historical Data for Forecasting the Long-run Equity Premium?", University of Toronto Working Paper.

Mankiw, N.G. (1986), "The Equity Premium and the Concentration of Aggregate Shocks", *Journal of Financial Economics*, vol. 17, no. 1 (September):211–219.

Mankiw, N. G. and S. P. Zeldes (1991), "The Consumption of Stockholders and Nonstockholders", *Journal of Financial Economics* 29, pp. 97-112.

Mas-Colell, A., M. D. Whinston and J. R. Green (1995), *Microeconomic Theory*, Oxford University Press.

Mayfield, E. S. (2004), "Estimating the Market Risk Premium", *Journal of Financial Economics*.

McGrattan, R. E. and E.C. Prescott (2001), "Is the Stock Market Overvalued?" NBER, Working Paper No. 8077.

McGrattan, R. E. and E.C. Prescott (2005), "Taxes, Regulations, and the Value of U.S. and U.K. Corporations", *Review of Economic Studies* 72, pp. 767–96, July.

Mehra, R. (2003), "The Equity Premium: Why is it a Puzzle?", *Financial Analysts Journal*, Vol. 59, pp. 54–69.

Mehra, R. and E. Prescott (1985), "The Equity Premium: A Puzzle", Journal of Monetary Economics, Vol. 15, pp. 145–161.

Mehra, R. and E. Prescott (1988), "The Equity Risk Premium: A Solution?", Journal of Monetary Economics, Vol. 22, pp. 133–136.

Mehra, R. and E. Prescott (2003), "The Equity Premium in Retrospect", in G.M. Constantinides, M. Harris, and R.M. Stulz (Eds.), Handbook of the Economics of Finance: Volume 1B Financial Markets and Asset Pricing. Amsterdam: Elsevier.

Mehra, R. and E. Prescott (2006), "The Equity Premium: What have we learned in 20 years?", in R. Mehra (Ed.), Handbook of Investments: Equity Risk Premium in the Handbook of the Economics of Finance series. Amsterdam: Elsevier.

Menzly, L., J. Santos and P. Veronesi (2002), "The Time Series of the Cross Section of Asset Prices", CRSP Working Paper No. 541.

Miller, M.H. (2000), "The History of Finance: An Eyewitness Account", *Journal of Applied Corporate Finance*, Vol. 13 No. 2, pp. 8–14.

Mishel, L., J. Bernstein, and S. Allegretto (2006), *The State of Working America*, Cornell University Press.

Modigliani, F. and R. A. Cohn (1979), "Inflation, Rational Valuation and the Market", *Financial Analysts Journal*, Vol. 35, No. 2, pp, 24-44

O'Hanlon, J. and A. Steele (2000), "Estimating the Equity Risk Premium Using Accounting Fundamentals", *Journal of Business Finance & Accounting* 27 (9&10), pp. 1051-1083.

O'Neill, J., D. Wilson and R. Masih (2002), "The Equity Risk Premium from an Economics Perspective", Goldman Sachs, Global Economics Paper No. 84.

Palepu, K.G., P.M. Healy and V.L. Bernard (2004), *Business Analysis and Valuation: Using Financial Statements, Text and Cases*, 3rd ed., South-Western College Pub.

Pastor, L. and R. F. Stambaugh (2001), "The Equity Premium and Structural Breaks", *Journal of Finance*, 56, No. 4, pp. 1207-1239.

Penman, S.H. (2003), *Financial Statement Analysis and Security Valuation*. 2nd edition, McGraw-Hill. 1st ed. in 2001.

Polk, C. K., S. B. Thompson and T. Vuolteenaho (2006) 'Cross-sectional forecasts of the equity premium', *Journal of Financial Economics* 81, 101—141

Pratt, S.P. (2002), *Cost of Capital: Estimation and Applications*, 2nd edition, Wiley.

Rietz, T.A. (1988) "The Equity Risk Premium: A Solution", *Journal of Monetary Economics*, vol. 22, no. 1 (July), pp. 117–131.

Ritter, J.R. (2002), "The Biggest Mistakes We Teach", *Journal of Financial Research,* Vol. 25, pp. 159-168.

Ritter, J.R. (2005), “Economic Growth and Equity Returns”, *Pacific-Basin Finance Journal* 13, pp. 489–503
Ritter, J.R. and R. Warr (2002), "The Decline of Inflation and the Bull Market of 1982 to 1999", *Journal of Financial and Quantitative Analysis*, Vol. 37, No. 1, pp. 29-61.
Ross, S. A., R. W. Westerfield and J. F. Jaffe (2005), *Corporate Finance*, 7th edition, Homewood, IL: McGraw-Hill/Irwin. Previous editions: 1998, 1993, 1996, 1999 and 2002.
Rozeff, M.S. (1984), “Dividend Yields Are Equity Risk Premiums”, *Journal of Portfolio Management*, pp. 68-75,
Saito, M. (1995), “Limited Market Participation and Asset Pricing”, manuscript, University of British Columbia.
Schwert, G.W. (1990), “Indexes of U.S. Stock Prices from 1802 to 1987”, *Journal of Business*, Vol. 63 (1990), pp. 399–426.
Shalit, H. and S. Yitzhaki (2006), "Capital Market Equilibrium with Heterogeneous Investors", SSRN 899276 .
Shiller, R. J. (2000), *Irrational Exuberance*, Princeton University Press, Princeton New Jersey, 2000.
Siegel, J. J. (1999), “The Shrinking Equity Premium”, *Journal of Portfolio Management*, Fall, pp. 10-17.
Siegel, J. J. (2002), *Stocks for the Long Run*, 3rd edition, New York: Irwin. 2nd ed. in 1998, 1st ed. in 1994.
Siegel, J. J. (2005a), “Perspectives on the Equity Risk Premium”, *Financial Analysts Journal*, Vol. 61, No. 6, pp. 61-71.
Siegel, J.J. (2005b), The Future for Investors: Why the Tried and the True Triumph Over the Bold and the New. NY: Crown Business.
Siegel, J.J. and J.D. Schwartz (2006), “The Long-Term Returns on the Original S&P 500 Firms”, *Financial Analysts Journal*, Vol. 62, No. 1, pp. 18-31.
Siegel, J.J. and R. Thaler (1997), “The Equity Premium Puzzle”, *Journal of Economic Perspectives* 11(1), pp. 191-200.
Smith, E. L. (1926), *Common Stocks as Long Term Investments*,
Storesletten, K., C.I. Telmer and A. Yaron (1999), "Asset Pricing with Idiosyncratic Risk and Overlapping Generations", Universitat Pompeu Fabra Working Paper No. 405.
Stowe, J.D., T.R. Robinson, J.E. Pinto and D.W. McLeavey (2002), *Analysis of Equity investments: Valuation.* AIMR (Association for Investment Management and Reasearch).
Van Horne, J. C. (1986), *Fundamentals of Financial Management and Policy*, 6th edition, Englewood Cliffs, NJ: Prentice-Hall. Previous editions: 1971, 1974, 1977, 1980 and 1983.
Van Horne, J. C. (1992), *Financial Management and Policy*, 9th edition, Englewood Cliffs, NJ: Prentice-Hall. Previous editions: 1968, 1971, 1974, 1977, 1980, 1983.
Vissing-Jorgensen, A. (2002), “Limited Asset Market Participation and the Elasticity of Intertemporal Substitution, *Journal of Political Economy* 110, pp. 835-53.
Vivian, A. (2005), “The Equity Premium: 101 Years of Empirical Evidence from the UK”, University of Durham. Available at SSRN number 766184.
Welch, I. (2000), “Views of Financial Economists on the Equity Premium and on Professional Controversies”, *Journal of Business*, Vol. 73, No. 4, pp. 501-537.
Welch, I. (2001), “The Equity Premium Consensus Forecast Revisited”, Cowles Foundation Discussion Paper No. 1325. SSRN n. 285169.
Weston, J. F. and E. F. Brigham (1968), *Essentials of Managerial Finance,* Holt, Rinehart and Winston.
Weston, J. F. and T. E. Copeland (1992), *Managerial Finance,* 9th edition, The Dryden Press.
Weston, J. F., S. Chung and J. A. Siu (1997), *Takeovers, Restructuring and Corporate Governance,* 2nd edition, New Jersey: Prentice-Hall.
Weston, J. F., M.L. Mitchel and J.H. Mulherin (2004), *Takeovers, Restructuring, and Corporate Governance,* 4th edition, Pearson Education, Prentice Hall.
Williams, J.B. (1938), *The Theory of Investment Value*, Harvard University Press, Cambridge, MA.
Wilson, J. and C. Jones (2002), “An Analysis of the S&P 500 Index and Cowles’s Extensions: Price Indexes and Stock Returns, 1870–1999”, *Journal of Business*, Vol. 75, pp. 505–533.

Analyst Long-term Growth Forecasts, Accounting Fundamentals and Stock Returns

Working Paper Draft Date: 4/4/2016

Abstract:

We decompose consensus analyst long-term growth forecasts into a hard growth component that captures accounting information (asset and sales growth, profitability and equity dilution) and an orthogonal soft growth component. The soft component does not forecast future returns, and the hard component does forecast future returns, but in a perverse way. Specifically, stocks with accounting information indicating favorable long-term growth forecasts tend to realize negative future excess returns. This and other evidence we present is consistent with biased long-term growth forecasts generating stock mispricing.

Contact Info:

Gregg Fisher

Gerstein Fisher
Chief Investment Officer
565 Fifth Avenue, 27th Floor
New York, NY 10017
United States
212-968-0707 ext. 111 (Phone)

Ronnie Shah[1]

Gerstein Fisher
Senior Director of Research
565 Fifth Avenue, 27th Floor
New York, NY 10017
United States
212-968-0707 ext. 138 (Phone)

Sheridan Titman

University of Texas at Austin - Department of Finance
McCombs School of Business
Walter W. McAllister Centennial Chair in Financial Services
Austin, TX 78712
United States
512-232-2787 (Phone)

---

[1]Contact author. The views expressed here are those of the authors and not necessarily those of any affiliated institution. We thank members of Gerstein Fisher Investment Strategy Group for their research assistance. This research has benefited from discussion with Michael Clement, Yong Yu and Sunny Yang.

I. Introduction

The Gordon growth model expresses a stock's price as a function of its current dividends, a discount rate, and long-term growth expectations. Of the three relevant components of price, determining long-term growth expectations requires the most judgement and is the most likely to be subject to systematic mistakes. This paper analyzes potential errors in long-term growth expectations by examining the long-term consensus (mean) forecasts of earnings reported by sell-side analysts.[2] Consistent with earlier work, we find evidence of systematic errors in the forecasts, as well as evidence that these errors are reflected in stock prices in ways that are consistent with various return anomalies discussed in the academic finance literature.

To better understand the biases in long-term growth forecasts we decompose the forecasts into what we call a hard component, which can be explained by accounting and choice variables, and a soft component, which is the residual. Elements of the hard component include accounting ratios that capture profitability and changes in sales, as well as choices that influence asset growth and equity dilution. As we show, both components of long-term growth are related to current stock prices, suggesting that either the forecasts or the rationale used by the forecasters influence stock prices.[3] However, our evidence indicates that the forecasts of sell-side analysts are systematically biased, and that these biases may have influenced stock prices in ways that make their returns predictable.

---

[2]Analysts periodically provide forecasts of the current, one- and two-year forward EPS and a longer-term growth rate (LTG) that reflects expected annual percentage changes in EPS after the two-year EPS forecast. The exact forecast period for LTG is subjective and can vary by analyst. Da and Warachka (2011) explain that LTG reflects an analyst's perception of EPS growth over the three-year period starting two years from now.

[3]There is a large literature that links analyst long-term growth forecasts to stock prices. Easton, Taylor, Shroff and Sougiannis (2001), Bradshaw (2004), Claus and Thomas (2001), Gebhardt, Lee and Swaminathan (1998) and Nekrasov and Ogneva (2011) use analyst long-term growth as an input for a residual income valuation model to estimate the cost of capital. Bandyopadhyay, Brown and Richardson (1995) examine 128 Canadian firms and find that 60% of the variation in analyst stock price recommendations can be explained by long-term earnings growth forecasts.

The observed biases are linked to the hard component of the growth forecasts. In particular, the forecasts suggest that analysts believe profits are mean reverting, but profitability actually tends to be fairly persistent. The forecasts also indicate that analysts believe that high past sales growth is a good predictor of future earnings growth. However, we find that high sales growth is actually weakly negatively associated with future earnings growth. Endogenous firm decisions, such as the rate of asset growth, and the use of external financing, are associated with higher growth forecasts, but the relationship between these choices and actual earnings growth is actually negative. The soft component of the growth forecasts does in fact correctly predict actual growth, although in some tests the relationship is relatively weak.

The above evidence is consistent with the idea that the logic of mapping hard information to expected future growth rates may be leading investors astray. If this is the case, investors may be able to profit with trading strategies that buy stocks when the hard component of growth is unfavorable and sell when the hard component is favorable. Our evidence, which is consistent with other papers in the investment anomalies literature, indicates that this is indeed the case.

Our paper is not the first to describe biases in analyst long-term growth forecasts and relate these biases to abnormal stock returns.[4] Previous research by Dechow and Sloan (1997), Chan, Karceski and Lakonishok (2003), La Porta (1996) and Sloan and Skinner (2002) find evidence that overly optimistic equity analyst forecasts contribute to the value premium and that growth stocks underperform when high expectations are not met. Copeland, Dolgoff, and Moel (2004) show that innovations in analyst long-term growth estimates are positively correlated with contemporaneous stock returns. A more recent paper by Da and Warachka (2011) conjectures that short-term earnings forecasts are much more accurate than the long-term forecasts and shows that a strategy that exploits differences between these forecasts generates excess returns.

We contribute to this literature in a number of ways. In particular, we are the first to consider how the various types of hard information, such as endogenous choices like asset growth and equity issues may influence long-term growth forecasts. Second, we are the first to seriously consider the challenges associated with estimating realized long-term earnings growth in a sample with considerable survivorship bias – close to 1/3 of our sample has missing realized five-year earnings growth as reported by I/B/E/S. Some of the missing firms were acquired and some went bankrupt, so our sample of survivors is clearly biased. As we will describe in detail later, to address this problem, we use the market-adjusted returns measured until the firm is no longer in the database to create a proxy for EPS growth rate.

Our paper is also related to the literature that examines the relation between information disclosed in firms' financial statements and future stock returns. For example, Novy-Marx (2013) finds that highly profitable firms outperform low profit firms. Lakonishok, Shleifer and Vishny (1994) report a negative relation between sales growth and future returns. There is also a larger literature that explores whether various measures of asset growth and equity dilution explain stock returns.[5] This literature suggests two potential explanations for why analysts provide favorable long-term growth forecasts for firms growing assets and raising external equity. The first explanation, discussed in Daniel and Titman (2006), is that executives tend to raise capital when soft information about growth prospects is most favorable. If analysts tend to overreact to this soft information, then we will see a relation between favorable analyst forecasts, increases in external financing, and negative future returns. A second, somewhat more cynical explanation is that analysts issue optimistic growth forecasts for firms that are likely to be raising capital externally. The idea here is that analysts that make optimistic

[5]Pontiff and Woodgate (2008), Daniel and Titman (2006) and Bradshaw, Richardson and Sloan (2006) find that firms that repurchase shares outperform those that issue additional shares. Cooper, Gulen and Schill (2008) and Titman and Wei (2004) find evidence that asset and capital investment growth, respectively, are negatively related to future returns.

long-term growth forecast make it easier for their investment bankers to generate underwriting business.[6]

One can potentially distinguish between these explanations by examining our evidence on data both before and after the enactment of the global research analyst settlement in September 2002 (See Kadan, Madureira and Wang (2009), Clarke, Kohrana, Patel and Rau (2011) and Loh and Stulz (2011) for more information on the global research analyst settlement), which curtailed the ability of investment bankers to influence sell-side recommendations. Consistent with the idea that the settlement changed analyst behavior, we find that the relation between hard information and future returns are weaker in the post-settlement period. This evidence, however, should be interpreted with caution given the short post-global settlement sample period and confounding events such as the inclusion of certain accounting ratios in quantitative investment models (McLean and Pontiff (2014) and Chordia, Subrahmanyam and Tong (2014)) and the effect of regulation-FD (Agrawal, Chadha and Chen (2006) and Mohanram and Sunder (2006)).

The rest of this paper is organized as follows. The first section describes the data used in our analysis and the characteristics of high and low forecasted growth firms. The second section presents the decomposition of analyst long-term growth forecasts and examines the persistence of long-term growth forecasts and different accounting and valuation ratios. The third section presents the main analysis, exploring how various measures of expected growth are related to valuation ratios and realized earnings growth. The fourth section analyzes how different components of long-term growth forecasts predict future stock returns. The fifth section discusses pre- and post-Global Settlement evidence and evaluates various explanations for our results. The final section concludes.

II. Data

[6]For a discussion of this more cynical view see Cragg and Malkiel (2009), Dechow, Hutton and Sloan (2000), Lin and McNichols (1998), Teoh and Wong (2002).

Our main variable of interest, consensus analyst long-term growth (LTG), is taken from I/B/E/S and reflects the mean analyst estimate of annualized earnings growth.[7] There are a few challenges associated with using this measure as an estimate of projected growth. First, each individual analyst long-term growth estimate is updated periodically at the discretion of the analyst, which creates the possibility of stale data. However, as we show, consensus analyst growth forecasts are very persistent through time, suggesting that the individual analyst forecasts change very slowly. Second, analysts do not always produce a long-term growth estimate to go alongside their shorter-term forecasts.

The starting sample for this study includes all NYSE, AMEX and NASDAQ stocks listed on both the Center for Research in Security Prices (CRSP) return files and the Compustat annual industrial files from 1982 through 2014. Information on stock returns, market capitalizations and prices are from the CRSP database. Balance and income sheet information, shares outstanding and GICS industry codes are from the COMPUSTAT database. Analyst long-term consensus growth forecasts (LTG), current stock prices, next year's consensus EPS and actual five-year annual EPS growth rates are from Institutional Brokers Estimate System (I/B/E/S) Summary file. I/B/E/S compiles these forecasts on the third Thursday of every month.

We exclude stocks that have negative or missing book equity, missing industry codes, LTG estimates, or missing accounting data required to construct the different variables used in this study. Two of our measures require non-zero information on sales and assets in year *t-2*, which mitigates backfilling biases. While we include financial stocks, excluding those securities has very little impact on the results reported in the paper. Our final sample has an average of 2,213 firms in each year.

[7]Our empirical results are economically similar using the median consensus forecast instead of the mean.

Following Fama and French (1992), we form all of our variables at the end of June in year *t*, using fiscal year *t-1* accounting information and analyst estimates from June of year *t*. For valuation ratios such as Price/Book, we use market equity from December of year *t-1*. For EPS valuation ratios based on analyst estimates and measures of company size, we use market equity from June of year *t* to measure the information in the numerator and the denominator at the same point in time. Stock returns are adjusted for stock delisting to avoid survivorship bias, following Shumway (1997). Portfolios used in various asset pricing tests are formed once a year on the last day in June, allowing for a minimum of a six-month lag between the end of the financial reporting period and portfolio formation.

Variable definitions are as follows. Realized EPS growth (REAL EPS) is from I/B/E/S and reflects the annualized growth rate in EPS over the past five years. Equity dilution (EQDIL) is measured as the percentage growth in split-adjusted shares outstanding. Sales growth (△SALES) is constructed as the year-over-year percentage growth in revenues divided by split-adjusted shares outstanding. Asset growth (△ASSETS) is equal to the year-over-year percentage growth in assets divided by split-adjusted shares outstanding. Profitability (ROA) is defined as operating income before depreciation scaled by assets. SIZE is the logarithm of company market capitalization measured at the end of June.[8] P/B is the logarithm of the market equity to book equity. P/E_{t+1} is the logarithm of the forward price to earnings calculated as the analyst consensus EPS for the next year divided by the price per share. Change in analyst long-term earnings forecasts (△LTG) is the year-over-year change in analyst consensus long-term earnings forecasts. Each year, variables are cross-sectionally winsorized to reduce the effect of outliers by setting values greater than the 99th percentile and less than the 1st percentile to the 99th and 1st

---

[8]To calculate book equity, we use the following logic which is largely consistent with the tiered definitions used by Fama and French (1992). Book equity is equal to shareholders' equity plus deferred taxes less preferred stock. If shareholders' equity is missing, we substitute common equity. If common equity and shareholders' equity are both missing, the difference between assets and liabilities less minority interest is selected. Deferred taxes are deferred taxes and/or investment tax credit. Preferred stock is redemption value if available; otherwise, carry value of preferred stock is used. We set to zero the following balance sheet items, if missing: preferred stock, minority interest, and deferred taxes.

percentile breakpoint values, respectively. All variables are updated annually at the end of June of each year. Our variable definitions are largely consistent with previous studies.

[Insert Figure 1 Here]

Figure 1 reports the average and median annual consensus analyst long-term growth forecast (LTG) from 1982 to 2014 and five-year realized EPS annualized growth rate from 1982 to 2009. The mean estimated growth rate over this period is remarkably stable, increasing from 15.4% in 1982 to 19.7% in 2001 and then decreasing to 14.0% in 2014. The actual five-year growth rate (1982 reflects the five-year growth rate between years 1982 and 1987) fluctuates from slightly higher than 0% to 17.8%. The median cross-sectional forecast and realized earnings growth rates show a similar pattern. Realized growth tends to be high following recessions (1991, 2003, and 2008) and much lower in periods that include recessions in the five-year window.

At the end of June of each year *t* stocks are allocated into quintiles based on LTG. Table 1 reports formation period (using accounting information from year *t-1*) value-weighted summary statistics for various accounting ratios, price-ratio variables and market capitalizations for each of the five quintile portfolios. The first quintile portfolio contains the firms with the lowest expected growth; the fifth quintile portfolio contains the firms with the highest expected growth. Over our sample period, analysts expect the lowest growth firms to average 7% annualized growth in earnings per share, while the top group has average projected EPS growth rates that are four times as large. The distribution of LTG is right-skewed: the middle group (3rd quintile) has close to a 14$^{\%}$ lower growth rate than the highest growth group, but only a 7% higher growth rate than the lowest growth group.

[Insert Table 1 Here]

Although the following comparison is plagued with clear survival bias, it is useful to compare the long-term growth forecasts with realized EPS growth. Realized EPS growth does line up with projected growth – increasing monotonically from a low of 3.0% for the quintile portfolio with the lowest LTG to a high of 13.6% for the highest LTG. The average forecast error, defined as the difference between the forecast and the actual growth, also increases monotonically moving from left to right, rising from 3.9% for the lowest LTG growth to 14.4% for the highest LTG group. Even the lowest expected growth firms based on LTG miss their long-term earnings projections, although the misses are relatively small. In contrast, the highest expected growth firms have average realized growth that is more than 50% less than their ex-ante forecast.

The second section of Table 1 Panel B shows that many of the accounting variables used in our study have a meaningful relation with long-term growth forecasts. High expected growth firms tend to have greater equity dilution (EQDIL) and higher past sales (ΔSALES) and asset growth (ΔASSETS). We also observe the same asymmetry associated with expected growth rates – the highest growth group has equity dilution ratios, sales and asset growth rates that are twice as large as the 4th quintile, while the difference between the 3rd and 4th quintile is not as large. Our last non-price variable, profitability (ROA), does not appear to be related to consensus long-term analyst growth.

The third section of Table 1 Panel B examines how price-related variables are related to growth expectations. The results show that low growth rate firms are not the largest firms in our sample, with a time-series average of yearly cross-sectional mean capitalization (SIZE) of 30.9 BN, but are larger than the highest growth rate firms, which have capitalizations of 19.8 BN. High growth firms also tend to have much higher valuation ratios (P/B, P/E_{t+1}) – the highest growth group has a market capitalization that is on average 39x next-period expected earnings, while the lowest growth group has a market capitalization that is only 14x next-period expected earnings. This is consistent with the idea that greater growth opportunities are reflected in higher valuation ratios.

III. Decomposing Growth Expectations

Table 2 presents regressions that document the relation between the hard information variables and long-term growth forecasts. The first four rows of Table 2 display univariate panel regressions of LTG on different firm characteristics using annual data from 1982 to 2014. Errors are clustered by firm and year. Long-term growth is measured as of June of year *t*, while the independent variables use accounting information from fiscal year *t-1*. Similar to Table 1, equity dilution (EQDIL), sales growth (ΔSALES) and asset growth (ΔASSETS) are all positively related to LTG. The fourth variable, profitability (ROA), is negatively related to long-term growth, but is not reliably different from zero (T-stat=1.65). Past sales growth has the highest explanatory power, explaining 10% of the variation in long-term growth.

[Insert Table 2 Here]

Rows 5 through 8 report our estimates of multivariate cross-sectional regressions of LTG on the four non-price accounting variables. The regressions are run both with and without fixed effects that capture variation in long-term growth forecasts by industry and year. In most regressions, the coefficients of both the accounting variables and the industry and firm fixed effects are statistically significant, indicating that we can explain analyst long-term growth forecasts with hard information.

The positive coefficients on sales growth indicate an expectation that the past sales growth will persist into the future, which should in turn lead to future EPS growth. Higher asset growth, or growth of certain quantities on the balance sheet, such as property, plant and equipment, can indicate the firm is making presumably positive NPV investments that will generate future earnings. Equity issuances can also indicate the presence of growth opportunities due to a need for additional capital, while share repurchases may indicate the lack of growth opportunities. The negative coefficient on profitability signifies expected mean reversion, as those low profit firms are expected to have the highest growth in EPS when compared to high profit firms.

The panel regressions reported in Table 2 implicitly assumes that the multivariate relation between the hard information variables and analyst long-term consensus growth forecasts are constant over time. Figure 2 displays the time-series Fama-MacBeth coefficients of contemporaneous accounting variables from a regression explaining analyst long-term growth forecasts. As the figure shows, most relationships are stable over time and all of the equity dilution, sales and asset growth coefficients are positive. The profitability coefficient varies the most, reaching a minimum in the late 90s, during which many technology firms had poor profits but high future expected growth. There does not appear to be a large difference in the coefficient estimate before and after the global settlement (August 2002).

[Insert Figure 2 Here]

In the tests that follow, we decompose analyst long-term growth forecasts into two parts. The first component, which we call *Hard Growth*, is the fitted values from the regression reported in the last row of Table 2 and reported in Equation 1.

$$\text{Hard Growth} = 0.04 + 0.08\ \text{EQDIL} + 0.05\ \Delta\text{SALES} + 0.04\ \Delta\text{ASSETS} - 0.12\ \text{ROA} \tag{1}$$

The second component, denoted *Soft Growth*, is the difference between LTG and Hard Growth. Soft Growth reflects analyst private views or information content in LTG that is unexplained by observable accounting variables.

For our measure of Hard Growth, we use the coefficients of the independent variables from the equation reported above, but we do not include the coefficients on industry or time dummies to avoid any forward-looking bias. This assumption is not material – when we use only same period information to form hard and soft growth measures, the results presented in later sections are not materially different.

To better understand how growth expectations are incorporated into market prices, Table 3 estimates the relation between the components of long-term growth and two valuation ratios. Panel A reports results for log price-to-book (P/B) and Panel B reports results for log of forward earnings-to-price (P/E_{t+1}). The first four rows of each panel examine the relation between the valuation ratios and the four accounting ratios. For the P/B ratio, each of the four accounting variables is significantly positively related, with R^2 ranging from 0.11 to 0.29. Given P/B ratio reflects the market's expectations of growth opportunities: the coefficients on the positive indicators of growth (EQDIL, ΔASSETS, ΔSALES) have the correct sign, while the coefficient on the negative indicator of growth, ROA, has the incorrect sign, although it has the lowest t-statistics of the four variables. For the P/E_{t+1} ratio displayed in Panel B, the three variables that indicate growth all have the predicted positive sign, although sales growth is not statistically significant. ROA has a negative sign and is statistically significant after controlling for industry variation.

[Insert Table 3 here]

The last four rows of each panel in Table 3 use Hard Growth (the fitted values from the last regression reported in Table 2) and Soft Growth (the difference between LTG and Hard Growth or the residual of the same regression) as independent variables. For both valuation ratios, we find that Soft Growth has a positive and highly significant relation with value. Hard Growth is also positive and significant in most regressions, but the relationships are not as strong. Indeed, all of the regressions are consistent with both the hard and soft information in the analyst forecasts being incorporated into market prices.

IV. Do Growth Estimates Predict Future Earnings Growth?

We next examine whether the soft and hard components of forecasted earnings growth actually predict realized earnings growth (REAL EPS). I/B/E/S and Dechow and Sloan (1997) estimate realized earnings growth over the past five years using an AR(1) regression of log (EPS) using six annual observations between years *t* and *t+5*, where year *t* is the reference year that LTG is measured. Hence, one can estimate the extent to which long term growth forecasts and the various components of expect growth predict actual growth.

Unfortunately, sample selection bias creates a major problem for this analysis. Estimating realized earnings growth requires future realizations of non-negative EPS values, and a number of firms in the sample experience negative earnings and a number of other firms drop out of our sample. Specifically, in our sample from 1982 to 2009, we have five-year earnings growth rates for only two-thirds of the original sample (41,957 out of 63,842 firm-years). For those stocks with five-year earnings growth data (REAL EPS), 97.4% have a full 60 months of stock returns, and the average compound return is 14.4% per year for this sample. In comparison, only 22.5% of stocks with missing REAL EPS data have 60 months of stock returns – those firms with 60 months of data, but missing REAL EPS data, have stock returns that averaged only 5.37% per year.

Clearly, the firms with missing data performed worse than those that stayed in our data base. However, firms leave the sample for a variety of reasons, such as mergers, as well as bankruptcy and negative future earnings. Hence, in addition to losing firms that do very poorly, we lose some because the firms did very well – as a result, the bias should affect both low and high expected growth firms. Indeed, we find that 42% of the high expected growth firms (top quintile based on LTG each year) and 27% of low expected growth firms (lowest quintile) have missing five-year earnings growth information.

Heckman's (1979) two-stage selection model provides a potential solution for this sample selection problem. However, this approach requires an instrument that is correlated with whether or not REAL EPS is missing but which is uncorrelated with actual EPS growth. Unfortunately, we have not been able to come up with a good instrument. What we do instead is come up with proxies for the missing data. Specifically, we calculate the five-year market-adjusted return $R_{i,MAR(t,t+5)}$ as the difference between the compound annual five-year stock return $R_{i(t,t+5)}$ measured from July of year t to June of year $t+5$ less the compound annual market return $R_{Mkt(t,t+5)}$ measured over the same period.[9]

$$R_{i,MAR(t,t+5)} = R_{i(t,t+5)} - R_{Mkt(t,t+5)} \qquad [2]$$

Figure 3 reports value-weighted, market-adjusted returns $R_{MAR(t,t+5)}$ for decile portfolios formed by ranking stocks on I/B/E/S five-year realized EPS growth rate (REAL EPS). We include all stocks that have non-missing EPS data. Moving from left-to-right, the average five-year market-adjusted return rises from -19.0% to 8.6%. The monotonic relation between the EPS growth and stock returns is consistent with Ball and Brown (1968), Ball, Kothari and Watts (1993), Daniel and Titman (2006) and suggests that return information is a good proxy for EPS growth.

[Insert Figure 3 Here]

The approach we take fills in missing earnings data, which reflect close to 1/3 of our sample, with estimates based on observed stock returns. Specifically, our matching process involves calculating the percentile rank of $R_{MAR(t,t+5)}$ for a given year using all firms (including those with missing REAL EPS), defined as the percent of firms with a lower $R_{MAR(t,t+5)}$, and takes values between 0 and 100. We then do the same exercise calculating the percentile rank of REAL EPS using the sample of non-missing firms from Figure 3.

[9] When a firm has less than 60 months of data, we use the available return data to estimate compound annual market-adjusted returns.

For each missing REAL EPS observation, we then assign the average five-year EPS growth rate estimated in the same year for the REAL EPS percentile rank that corresponds to the same percentile rank of $R_{MAR(t,t+5)}$. Our procedure matches a distressed firm with poor stock returns and missing EPS growth rate, potentially due to negative earnings or a bankruptcy a low EPS growth rate. Similarly, the procedure matches a firm that has high stock returns and a missing five-year EPS growth rate, possibly due to a corporate action such as a merger, with a high EPS growth rate.

Figure 4 displays a histogram of $R_{MAR(t,t+5)}$ for those firms with missing REAL EPS data. This figure provides a sense of the distribution of market-adjusted stock returns for the sample with missing data and whether firms are matched to low or high realized EPS growth rates. The matched firms often have very low or very high market-adjusted returns – 22% of the missing sample in which $R_{MAR(t,t+5)}$ was in the bottom decile of future average returns, while 19% were in the top decile. In contrast, only 11% of the missing sample had future five-year returns that were either in the fifth or sixth deciles.

We examine why firms have missing REAL EPS. For those firms in the highest decile of market-adjusted returns, 93% were delisted because of a merger or acquisition. Among those in the lowest decile of market-adjusted returns, almost all of those firms were either delisted over the next five years because of bankruptcy or had negative earnings over the five-year period.

[Insert Figure 4 Here]

Table 4 reports results for a panel regression of 5-year realized EPS growth (REAL EPS) on our measures of hard and soft information. When REAL EPS is missing, we assign a future EPS growth rate as described above. Errors are clustered by industry and firm, which help to correct for the overlapping nature of estimating realized EPS growth over five years. The first two rows display results without inclusion of LTG; the third and fourth rows include LTG. In our fourth specification reported on the fourth row, we find equity dilution (T-stat=7.41), sales growth (T-stat=2.67) and asset growth (T-stat=2.16) are all significantly negatively related to actual growth, despite being positively related to forecasted growth. Profitability is also reliably positively related to actual growth (T-stat=5.02), even though profitability loads negatively on forecasted growth. We also find a negative relation between LN (P/B) ratio (T-stat=3.11) and realized growth, suggesting that growth stocks have lower earnings growth when compared to value stocks. After including industry and year dummies, the coefficient on analyst long-term growth (T-stat=1.00) is no longer significant, indicating that analyst long-term estimates are relatively poor predictors of actual earnings growth after controlling for hard information, and industry and year fixed effects.

[Insert Table 4 Here]

The last two rows of Table 4 report regression results of hard and soft growth on realized five-year earnings growth. In our first specification in row 5, we find a negative and significant relation between hard growth (T-stat=4.39), and realized earnings growth. We also find a significant positive relation between soft growth (T-stat=2.58) and realized earnings growth. After including industry and year dummies reported in the last row of Table 4, the coefficient on soft growth declines from 0.11 to 0.02 and is no longer significantly different from zero (T-stat=0.63). A straightforward extension of our analysis (which, for the sake of brevity, we do not report) is that hard accounting information also explains analyst forecast errors; i.e. the difference between the realized 5-year earnings growth and the analyst long-term consensus growth forecast.

To understand the importance of these results, recall that Table 2 shows that sales and asset growth and equity dilution variables are positively related to analyst long-term growth expectations, while profitability is negatively related. Table 4 illustrates the opposite: profitability is positively related to actual earnings growth, but sales and asset growth and equity dilution is negatively related. These results are consistent with a bias in how analysts and markets perceive hard information when making earnings growth forecasts and setting prices.

Analysts, and by extension financial markets, may make mistakes due to the way they interpret the persistence of certain accounting variables. Increasing sales and high profitability is generally associated with greater earnings growth. Similarly, endogenous variables such as asset growth and equity dilution may indicate future investment or the presence of growth opportunities. In Figure 3, we report Spearman rank correlations for each variable and their future values to examine the persistence of different variables that are related to growth expectations. The x-axis reflects the number of years between the current and future variable values. Correlations for each measure decline as more time elapses.

[Insert Figure 3 Here]

Our results suggest that analysts make mistakes when interpreting the persistence of accounting information while setting growth expectations. The "level" variables based on ratios of balance sheet information or market prices (ROA, P/B, P/E_{t+1}) tend to have high persistence, initially ranging from 0.70 to 0.84 for a one-year lag (t+1) and falling to 0.43 to 0.62 for a five-year lag (t+5). Value companies tend to stay value companies, and profitable firms tend to stay profitable. In contrast, the "change" variables, or those variables based on differences in balance sheet quantities (EQDIL, ΔASSETS, ΔSALES), exhibit far less persistence: one-year lag correlations are between 0.41 to 0.27 and decline to 0.20 to 0.11 for a five-year lag. Analyst long-term growth (LTG) is also very persistent, with serial correlations that decline from 0.84 (one-year) to 0.61 (five-year).

The correlations reported in Table 2 and Equation 1 show how analysts expect certain accounting quantities will affect future earnings growth. For example, profitability has a negative loading on LTG, indicating that analysts believe that low profit firms today will have higher earnings growth and hence high future profits. In reality, profitability is fairly persistent and low profit firms do not have higher earnings growth when compared to high profit firms. Sales growth also has a positive correlation with analyst long-term earnings growth forecasts indicating that analysts expect sales growth will persist in the future, even though it is actually not very persistent and a negative (weak) indicator of actual earnings growth. Similarly, endogenous variables such as asset growth and equity dilution which should reflect growth opportunities load positively on LTG. However, these indicators of growth are also not very persistent and are actually negatively related to actual earnings growth.

As we show, there is a tendency for these mistakes to at least partially correct over the following year. Table 5 reports regressions of year-over-year changes in analyst consensus long-term growth (LTG) on accounting and manager choice variables. The first four rows show that change variables (equity dilution, asset and sales growth) are associated with strong negative revisions in LTG. The coefficient on the fourth variable, ROA, does not predict innovations in LTG. Our composite variable, Hard Growth, also predicts when LTG forecasts will be revised downwards.

[Insert Table 5 Here]

If LTG forecasts do in fact reflect market beliefs, and if their revisions can be predicted with the Hard Growth component, then one might conjecture that the Hard Growth component also predicts returns. As we show in the next section, this is indeed the case.

V. Do Errors in Growth Forecasts Lead to Return Predictability?

Our final analysis, reported in Table 6, examines how the different components of long-term growth forecasts explain differences in average stock returns. Panel A of the Table reports average value-weighted returns for portfolios formed on LTG, Hard Growth and Soft Growth for those firms with available LTG and accounting data. Consistent with

Jung, Shane and Yang (2012), we find that analysts' consensus long-term growth expectations are unrelated to future stock returns. Our measure of Hard Growth, however, is strongly negatively related to average returns. Average returns for value-weighted portfolios formed on Hard Growth reported in the 2nd row of Table 6 Panel A decline from 1.19 for decile 1 (lowest growth) to 1.04 for decile 9. The last decile, which includes the firms with the highest Hard Growth indicators (low profitability, high external financing, high asset and sales growth), has monthly returns that are 55 basis points lower than the previous decile; the difference between the top and bottom decile is -0.60% per month (T-stat=2.66). In contrast, the last row of Table 6 Panel A shows that Soft Growth, which reflects analysts' views that is unrelated to accounting information, is unrelated to stock returns.

[Insert Table 6 Here]

Panels B and C of the table report these same portfolio returns for smaller firms and for a larger sample that also includes firms that do not have LTG data. Panel B, which reports returns on the smallest half of the firms (based on market capitalization), shows stronger results – the average return of the top decile is 0.86% less per month (T-stat=3.88) when compared to the average return of the bottom decile. Panel C examines a larger data on firms with data available to measure Hard Growth, but including firms that may not have LTG forecasts. Not requiring LTG estimates doubles the sample size to an average of 4,045 firms per month. As we show, with this larger sample that more closely reflects the samples used in earlier studies of these return anomalies, we find a very strong relation between our estimate of hard growth and stock returns – the average return of a portfolio that is long the highest decile of hard growth firms and short the lowest decile of hard growth firms is -0.79% (T-stat = 3.38).

[Insert Table 7 Here]

Table 7 reports results from Fama-MacBeth regressions of monthly returns on our hard and soft growth measures, with controls for firm size and book-to-market. There is evidence of a weak size (insignificant in all regressions) and stronger value effect (significant in every regression except one) in our sample. In the first regression on the left of the table, LTG is not related to average returns. The second regression includes variables that capture accounting information and manager decisions. We find a significant and positive relation between equity dilution (T-stat=5.25) and asset growth (T-stat=4.39) and average returns. The coefficient of sales growth (T-stat=1.86) is positive and the coefficient of profitability (T-stat=1.66) is negative, the significance of each is marginal. Including LTG in the third regression causes the significance of all the variables to increase – with sales growth (T-stat=2.12) and profitability (T-stat=2.16) now significantly different from zero at the 5% level. The t-statistics and coefficients on the hard information variables reported in the 4th regression are even stronger after including fixed effects that capture differences in industry returns each month.

The final two regressions examine how hard and soft growth relate to average returns. The results largely mirror those reported in Table 6, with LTG and soft growth not related to average returns while hard growth is strongly negatively related to average returns. The Fama-MacBeth approach equal-weights stock returns in each cross-section, compared to the value-weighted portfolio returns reported in the previous table. Our results suggest that hard growth generates a larger difference in returns among smaller stocks when compared to larger stocks, which is consistent with the results presented in Table 6 Panels B and C.

VI. The Effect of the Global Analyst Research Settlement on Long-term Growth Forecasts

The results presented in the previous sections suggest the market misinterprets hard information that signals high growth leading to underperformance, particularly for firms with the most extreme growth forecasts. One possibility explored in Dechow, Hutton and Sloan (2000) is that analysts hype those firms to gain more investment banking business and make it easier for firms to issue equity or debt. An alternative explanation is that managers tend to invest when intangible information is positive and that investors tend to over-react to intangible information (Daniel and Titman (2006)). Manager choice variables such as equity dilution and asset growth signal favorable or unfavorable intangible information, which leads to return predictability.

Rule NASD 2711 and NYSE 472, better known as the Global Analyst Research Settlement, were regulations to reduce the ability of investment banks to influence analysts' stock recommendations. The ruling required the analysts to provide disclosure of any conflict they (or their firm) may have with the recommended stock. We follow Kadan, Madureira and Wang (2009), Clarke, Khorana, Patel and Rau (2011) and Loh and Stulz (2011) by assigning the period starting with September 2002 as the post-global settlement. Analyzing our tests pre- and post-global settlement allows us to better understand how analysts change how (i) analysts form their forecasts, (ii) forecasts are incorporated into market prices, (iii) actual earnings growth is related to hard and soft information, and (iv) whether hard and soft information still has the ability to predict future stock returns.

Our decomposition is important, as we are able to explain how analysts, markets and actual earnings growth differentially react to information on long-term growth forecasts. The competing explanations provided by Dechow, Hutton and Sloan (2000) and Daniel and Titman (2006) are more relevant for managerial decisions related to capital issuance and retirements, or the level of capital expenditures and are less relevant for firm characteristics that are largely out of the control of the manager, such as sales growth or profitability.

Returning to Figure 2, we do not find meaningful differences in the way analysts form their long-term growth expectations: changes in sales and asset growth and equity dilution is positively related to LTG, while ROA is negatively related to LTG. Our results suggest that Global Settlement did not change how analysts process hard information.

[Insert Table 8 Here]

Table 8 replicates the main analyses in our paper for the pre-Global Settlement period from July 1982 to August 2002 and the post-Global Settlement period from September 2002 to December 2014. In our analysis presented in Table 8, we do not include ΔSALES and ROA as independent variables and instead focus on the manager choice variables that related to the competing explanations for our results: EQDIL and ΔASSETS. Table 8 Panel A reports our split-sample results for the panel regressions from Tables 3 and 4. In the early period, we find a very strong correlation between asset growth and the natural log of the price-to-book ratio (T-stat=12.79), consistent with Fama and French (2015), who find a high correlation between HML (low price-to-book less high price-to-book factor) and CMA (low asset growth less high asset growth), and a weaker but still statistically positive relation between log price-to-book and equity dilution (T-stat=2.75). In the later period, we find the coefficient on equity dilution becomes negative (T-stat=6.54), and there is still a positive relation with asset growth (T-stat=7.67). The weaker results in the post-global settlement period for manager choice variables help explain why Hard Growth (T-stat=0.49) is insignificantly positively related to price-to-book ratio.

For the natural log of forward earnings-to-price ratios reported in rows 5 through 8 of Table 8 Panel A, we find a positive correlation between both manager choice variables and price-to-book ratio in the pre-GS period, but the asset growth's coefficient sign flips in the post-GS period. Despite the negative relation between △ASSETS and LN (P/B), the coefficient on Hard Growth (T-stat=2.21) in the later period is still significantly different from zero.

The next four rows display regression results for the pre- and post-GS periods for regressions predicting five-year realized earnings growth. Before global settlement, price-to-book ratio is significantly negative related to actual EPS growth (T-stat=2.66, 3.70), while after global settlement price-to-book is unrelated to actual EPS growth (T-stat=0.64, 0.70). The coefficient on asset growth is significantly negative in the early period (T-stat=2.14), but becomes insignificant in the later period (T-stat=0.50). Equity dilution is a little stronger in the later period, when compared to the earlier period. We find a slightly higher Hard Growth coefficient estimate in the post-global settlement period (0.64) when compared to the pre-global settlement period (0.70).

The last four rows reports split-sample regression results predicting year-over-year changes in LTG. In both sub-periods, we find that equity dilution and asset growth predict negative innovations in LTG, but the coefficient on equity dilution in the post-GS period while significantly different from zero is roughly half of what it was in the pre-GS period. We also find that hard growth is associated with negative future changes in LTG in both sub-periods.

Table 8 Panels B and C report pre- and post-GS period average returns for value-weighted portfolios formed on various growth measures. The return earned by going long firms in the highest decile of equity dilution and going short the lowest decile of equity dilution declines from -0.90% (T-stat=4.47) in the earlier period to -0.43% (T-stat=1.81) in the later period. The long/short return for asset growth is negative and marginally significant in the early period (-0.53), but is positively and insignificant in the later period (0.24%). These results help explain why the difference between the highest decile portfolio and lowest decile portfolio of Hard Growth in the early period is -0.74% (T-stat=2.25) in the early period, but shrinks to -0.36% (T-stat=1.49) in the later period.

As we show, soft growth which reflects analysts' private views are positively related to valuations (P/B, P/E_{t+1}), is (weakly) positively related to actual growth, and does not explain stock returns. Our findings suggest that this component of analyst long-term growth is accurately incorporated into market prices, and that when those growth expectations are met there is no material return predictability. There is also very little change in how soft growth is related to valuations and actual earnings growth pre- and post-global settlement.

In contrast, analysts in the post-global settlement period still assign higher growth expectations to firms with low profitability, high past sales and asset growth and high external financing *despite the regulation*'s *potential influence* on the bias of these estimates. Firms with these characteristics also experience negative revisions in long-term growth forecasts in the post-GS period. Our evidence suggests regulation did not materially change how analysts interpret hard information when making long-term growth forecasts – thus, either the analysts are still trying to gain investment banking business by issuing overly optimistic growth forecasts, or are making genuine mistakes when setting long-term earnings growth expectations. However, it is hard to draw conclusions due to the small sample size of the post-GS period.

Our findings suggest the market, however, isn't fooled by this analyst behavior after August 2002 and potentially learned from the mistakes made when setting prices during the dot-com period between 1998 and 2002 as the relation between hard growth and the log of the price-to-book ratio is weaker. Hard information is a negative predictor of realized earnings growth in both sample periods. In the post-GS period, we find weaker evidence that hard information predicts future returns, which suggests our results are driven by former hypothesis related to analysts hyping stock prices to win investment banking business. However, there is an alternative explanation related to certain market participants exploiting profitability, asset growth or external financing factors to correct and profit from investor mistakes related to mispricing associated with long-term growth forecasts. Of course, we cannot rule out that the weaker results in the latter period are a result of a small sample size instead of a shift in investor behavior or other informed traders exploiting this mispricing.

VII. Conclusion

There is now substantial evidence linking various income statement and balance sheet items to future excess stock returns. While it is possible that these excess returns are associated with systematic sources of risk that investors wish to avoid, the magnitudes of the observed abnormal returns and the Sharpe ratios that can be obtained by exploiting the strategies are simply too large to be consistent with equilibrium risk premia. In other words, during our sample period, the evidence suggests that the consensus views of investors were incorrect along some meaningful dimensions.

To explore this hypothesis, we use the consensus analyst long-term earnings growth forecast as a proxy for growth expectations and examine how these expectations are influenced by various accounting variables. Our focus is on two variables that are under the direct control of a firm's management – the extent to which the firm issued or repurchased its shares and the extent to which it grew is assets and two variables that management can only indirectly control – the sales growth and profitability of the firm. As we show, these variables explain the consensus long-term growth forecasts of analysts, and as such, they also influence stock prices. However, the sign of the correlation between these variables and realized earnings growth is inconsistent with the correlation between these variables and both analyst long-term earnings growth forecasts and firm valuations. Thus, high market prices reflect faulty growth expectations and sorting stocks on these accounting variables produces meaningful differences in average returns.

It would be nice to have better intuition about why the analysts and investors made these mistakes. One possibility, explored in a number of papers, is that analysts bias their earnings forecasts to cater to firms that are likely to need future investment banking services. Another possibility is that market prices influence management choices. If the market and the analyst community view the firm favorably, the firm is more likely to raise capital, grow its assets, and may feel less compelled to increase sales and profitability. In other words, the favorable view of the market may in some cases sow its own seeds of destruction. Finally, it's possible that the analysts simply made mistakes in our sample period.

While we have made a preliminary exploration of these issues by looking at how long-term earnings growth forecasts have changed over time, our results are not conclusive. Hopefully, future research can help better understand the cause of these earnings forecast errors.

References

Agrawal, Anup, Sahiba Chadha, and Mark A. Chen. "Who Is Afraid of Reg FD? The Behavior and Performance of Sell-Side Analysts Following the SEC's Fair Disclosure Rules*." The Journal of Business 79.6 (2006): 2811-2834.

Bandyopadhyay, Sati P., Lawrence D. Brown, and Gordon D. Richardson. "Analysts' use of earnings forecasts in predicting stock returns: Forecast horizon effects." International Journal of Forecasting 11.3 (1995): 429-445.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny. "A model of investor sentiment." Journal of financial economics 49.3 (1998): 307-343.

Bauman, W. Scott, and Richard Dowen. "Growth projections and common stock returns." Financial Analysts Journal 44.4 (1988): 79-80.

Bradshaw, Mark T., Scott A. Richardson, and Richard G. Sloan. "The relation between corporate financing activities, analysts' forecasts and stock returns." Journal of Accounting and Economics 42.1 (2006): 53-85.

Bradshaw, Mark T. "How do analysts use their earnings forecasts in generating stock recommendations?" The Accounting Review 79.1 (2004): 25-50.

Chan, Louis KC, Jason Karceski, and Josef Lakonishok. "The level and persistence of growth rates." The Journal of Finance 58.2 (2003): 643-684.

Chordia, Tarun, Avanidhar Subrahmanyam, and Qing Tong. "Have capital market anomalies attenuated in the recent era of high liquidity and trading activity?" Journal of Accounting and Economics 58.1 (2014): 41-58.

Clarke, Jonathan E., Ajay Khorana, Ajay Patel, and P. Raghavendra Rau. "Independents' day? Analyst behavior surrounding the Global Settlement." Annals of Finance 7, no. 4 (2011): 529-547.

Claus, James, and Jacob Thomas. "Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets." Journal of Finance (2001): 1629-1666.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill. "Asset growth and the cross-section of stock returns." The Journal of Finance 63.4 (2008): 1609-1651.

Copeland, Tom, Aaron Dolgoff, and Alberto Moel. "The role of expectations in explaining the cross-section of stock returns." Review of Accounting Studies 9.2-3 (2004): 149-188.

Cragg, John G., and Burton G. Malkiel. Expectations and the structure of share prices. University of Chicago Press, 2009.

Da, Zhi, and Mitch Warachka. "The disparity between long-term and short-term forecasted earnings growth." Journal of Financial Economics 100.2 (2011): 424-442.

Daniel, Kent, and Sheridan Titman. "Market reactions to tangible and intangible information." The Journal of Finance 61.4 (2006): 1605-1643.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. "Investor psychology and security market under-and overreactions." the Journal of Finance 53.6 (1998): 1839-1885.

Dechow, Patricia M., and Richard G. Sloan. "Returns to contrarian investment strategies: Tests of naive expectations hypotheses." Journal of Financial Economics 43.1 (1997): 3-27.

Dechow, Patricia M., Amy P. Hutton, and Richard G. Sloan. "The Relation between Analysts' Forecasts of Long-Term Earnings Growth and Stock Price Performance Following Equity Offerings." Contemporary Accounting Research17.1 (2000): 1-32.

Easton, Peter, et al. "Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment." Journal of Accounting Research 40.3 (2002): 657-676.

Fama, Eugene F., and Kenneth R. French. "The cross-section of expected stock returns." the Journal of Finance 47.2 (1992): 427-465.

Fama, Eugene F., and Kenneth R. French. "A five-factor asset pricing model." Journal of Financial Economics 116.1 (2015): 1-22.

Finger, Catherine A., and Wayne R. Landsman. "What do analysts' stock recommendations really mean?" Review of Accounting and Finance 2.2 (2003): 67-87.

Frankel, Richard, and Charles MC Lee. "Accounting valuation, market expectation, and cross-sectional stock returns." Journal of Accounting and economics 25.3 (1998): 283-319.

Gebhardt, William R., Charles MC Lee, and Bhaskaran Swaminathan. "Toward an implied cost of capital." Journal of accounting research 39.1 (2001): 135-176.

Heckman, James J. "Sample selection bias as a specification error."Econometrica: Journal of the econometric society (1979): 153-161.

Hong, Harrison, Terence Lim, and Jeremy C. Stein. "Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies." The Journal of Finance 55.1 (2000): 265-295.

Jung, Boochun, Philip B. Shane, and Yanhua Sunny Yang. "Do financial analysts' long-term growth forecasts matter? Evidence from stock recommendations and career outcomes." Journal of Accounting and Economics 53.1 (2012): 55-76.

Kadan, Ohad, Leonardo Madureira, Rong Wang, and Tzachi Zach. "Conflicts of interest and stock recommendations: The effects of the global settlement and related regulations." Review of Financial Studies 22, no. 10 (2009): 4189-4217.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. "Contrarian investment, extrapolation, and risk." The journal of finance 49.5 (1994): 1541-1578.

La Porta, Rafael. "Expectations and the cross-section of stock returns." The Journal of Finance 51.5 (1996): 1715-1742.

Lin, Hsiou-wei, and Maureen F. McNichols. "Underwriting relationships, analysts' earnings forecasts and investment recommendations." Journal of Accounting and Economics 25.1 (1998): 101-127.

Loh, Roger K., and René M. Stulz. "When are analyst recommendation changes influential?." Review of Financial Studies 24, no. 2 (2011): 593-627.

McLean, D., Pontiff, J. 2014. Does academic research destroy return predictability? Working paper.

Mohanram, Partha S., and Shyam V. Sunder. "How Has Regulation FD Affected the Operations of Financial Analysts?*." Contemporary Accounting Research 23.2 (2006): 491-525.

Novy-Marx, Robert. "The other side of value: The gross profitability premium."Journal of Financial Economics 108.1 (2013): 1-28.

Pontiff, Jeffrey, and Artemiza Woodgate. "Share Issuance and Cross-sectional Returns." The Journal of Finance 63.2 (2008): 921-945.

Ramnath, Sundaresh, Steve Rock, and Philip B. Shane. "A review of research related to financial analysts' forecasts and stock recommendations." Available at SSRN 848248 (2008).

Skinner, Douglas J., and Richard G. Sloan. "Earnings surprises, growth expectations, and stock returns or don't let an earnings torpedo sink your portfolio." Review of accounting studies 7.2-3 (2002): 289-312.

Sharpe, Steven A. "How does the market interpret analysts' long-term growth forecasts?" Journal of Accounting, Auditing & Finance 20.2 (2005): 147-166.

Stickel, Scott E. "Analyst incentives and the financial characteristics of Wall Street darlings and dogs." The journal of investing 16.3 (2007): 23-32.

Shumway, Tyler. "The delisting bias in CRSP data." Journal of Finance (1997): 327-340.

Teoh, Siew Hong, and Tak Jun Wong. "Why New Issues and High-Accrual Firms Underperform: The Role of Analysts' Credulity." Review of Financial Studies 15.3 (2002): 869-900.

Titman, Sheridan, Kuo-Chiang Wei, and Feixue Xie. "Capital investments and stock returns." Journal of financial and Quantitative Analysis 39.04 (2004): 677-700.

Figure 1. Average Consensus Analyst Long-term Growth Estimates and Realized 5-year EPS Growth Rate from 1982 to 2014. The figure plots cross-sectional mean and median estimates for LTG and REAL EPS by year. LTG is the mean estimate of all analysts' expectations of the future EPS annual growth rate measured in the 3rd week of June of year *t*. REAL EPS is the five-year average annualized realized EPS growth rate between year *t* and year *t+5*.

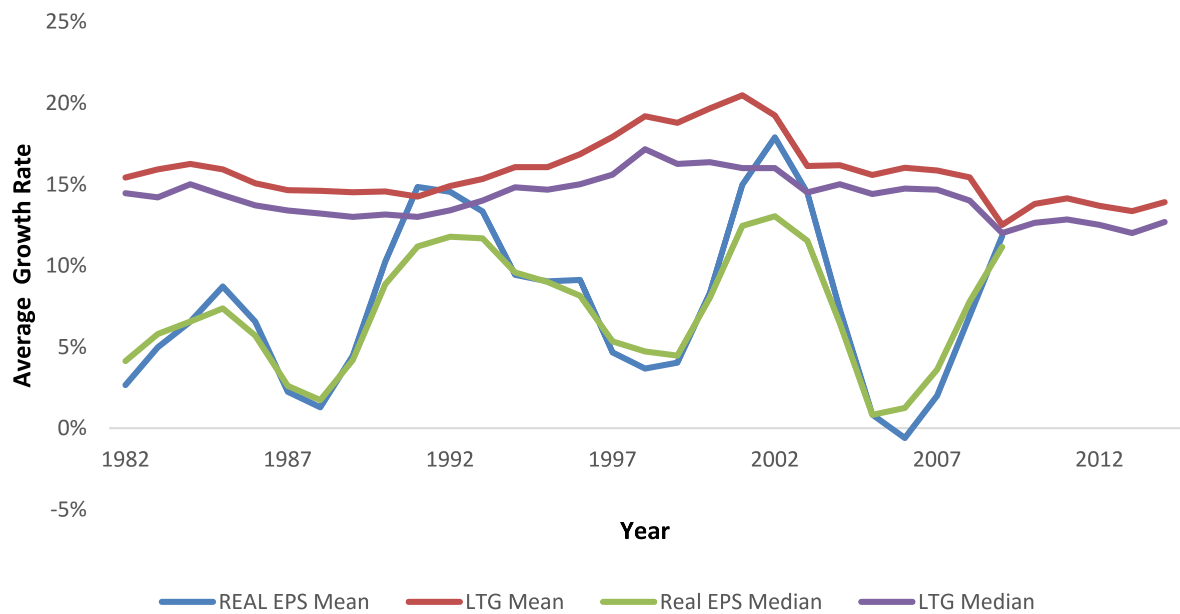


Table 1. Sample Summary Statistics from 1982 to 2014. This table presents summary statistics for firms that meet the restrictions described in the data section. The first panel describes the distribution of analyst long-term growth forecasts, LTG. At the end of June of each year *t*, stocks are ranked on LTG and then allocated to five groups, each with an equal number of stocks. The second panel reports value-weighted averages for LTG, 5-year realized earnings growth, accounting ratios, valuation ratios and market capitalization for each quintile portfolio using information available at the portfolio formation date. Variable definitions are as follows. LTG measures the mean estimate of all analysts' expectations of the future EPS annual growth rate measured in the 3rd week of June of year *t*. REAL EPS is the five-year average annualized future EPS growth rate between year *t* and year *t+5*. EqDil (equity dilution) is the percentage change in split-adjusted shares outstanding from year *t-2* to year *t-1*. ΔSales (sales growth) is the percentage change in revenues per split-adjusted share from year *t-2* to year *t-1*. ΔAssets (asset growth) is the percentage change in assets per split-adjusted share from *t-2* to *t-1*. ROA (profitability) is operating income in year *t-1* divided by assets for year *t-1*. SIZE x 10^9 is market capitalization (in millions) as of June of year *t*. P/B (price/book ratio) is market capitalization as of December of year t-1, divided by book equity in year *t-1*. P/E_{t+1} (price/forward earnings ratio) is price per share divided by fiscal year 1 analyst consensus earnings per share measured in the 3rd week of June of year *t*. The sample has an average of 2,213 firms per year.

Panel A. Average Analyst Long-Term Growth Statistics

p1	Median	Mean	p99	σ
0.010	0.142	0.158	0.484	0.084

Panel B. Average Firm Characteristics by Analyst Long-Term Growth Quintile

	1	2	3	4	5
Growth Variables					
LTG	0.070	0.111	0.141	0.181	0.280
REAL EPS	0.030	0.057	0.070	0.087	0.136
Non-Price Variables					
EQDIL	0.024	0.018	0.015	0.037	0.076
ΔSALES	0.048	0.070	0.098	0.155	0.311
ΔASSETS	0.059	0.091	0.122	0.181	0.335
ROA	0.140	0.145	0.170	0.188	0.171
Price Variables					
SIZE x 10^9	30.91	32.93	26.55	23.34	19.80
P/B	1.98	3.18	3.70	4.80	6.54
P/E_{t+1}	14.31	16.15	19.04	23.60	39.00

Table 2. Panel Regression Explaining Long-Term Growth from 1982 - 2014. This table reports results from panel regressions of analyst long-term growth (LTG) on past accounting growth measures. LTG is the mean estimate of all analysts' expectations of the EPS annual growth rate between year *t+2* to year *t+5* measured in the 3rd week of June of year *t*. EQDIL (equity dilution) is the percentage change in split-adjusted shares outstanding from fiscal year-end in *t-2* to *t-1*. ΔSALES (sales growth) is the percentage change in revenues per split-adjusted share from *t-2* to *t-1*. ΔASSETS (asset growth) is the percentage change in assets per split-adjusted share from year *t-2* to year *t-1*. ROA (profitability) is operating income in year *t-1* divided by assets in year *t-1*. N is the average number of stocks each year. Certain regressions use industry (Based on GICs 10 sector definitions) and year fixed effects. T-statistics are reported in parentheses based on robust standard errors that are clustered by firm and industry. The number of firm-year observations is 74,130.

	Intercept	EQDIL	ΔSALES	ΔASSETS	ROA	R^2	Industry Fixed Effect?	Year Fixed Effect?
Coefficient	0.16	0.12				0.04	No	No
t-stat	(11.75)	(4.02)						
Coefficient	0.15		0.08			0.10	No	No
t-stat	(11.35)		(13.56)					
Coefficient	0.15			0.08		0.07	No	No
t-stat	(10.62)			(12.68)				
Coefficient	0.17				-0.11	0.02	No	No
t-stat	(8.23)				(1.65)			
Coefficient	0.15	0.10	0.06	0.05	-0.11	0.17	No	No
t-stat	(8.23)	(9.36)	(13.99)	(8.12)	(1.87)			
Coefficient	0.07	0.09	0.05	0.04	-0.12	0.34	Yes	No
t-stat	(20.92)	(7.50)	(10.46)	(13.40)	(4.54)			
Coefficient	0.14	0.09	0.06	0.05	-0.10	0.20	No	Yes
t-stat	(10.77)	(11.18)	(15.13)	(7.68)	(1.85)			
Coefficient	0.04	0.08	0.05	0.04	-0.12	0.37	Yes	Yes
t-stat	(7.56)	(8.43)	(10.52)	(14.23)	(4.64)			

Figure 2. Coefficient Estimates from Annual Regressions Explaining Long-Term Growth from 1982 - 2014. This figure plots the time-series of coefficients from a Fama-Macbeth regression of analyst long-term growth on equity dilution, sales growth, asset growth, profitability variables and industry dummies. LTG measures the mean estimate of all analysts' expectations of the EPS annual growth rate between year *t+2* to year *t+5* measured in the 3rd week of June of year *t*. EQDIL (equity dilution) is the percentage change in split-adjusted shares outstanding from fiscal year-end in *t-2* to *t-1*. ΔSales (sales growth) is the percentage change in revenues per split-adjusted share from *t-2* to *t-1*. ΔAssets (asset growth) is the percentage change in assets per split-adjusted share from *t-2* to *t-1*. ROA (profitability) is operating income in *t-1* divided by assets in *t-1*.

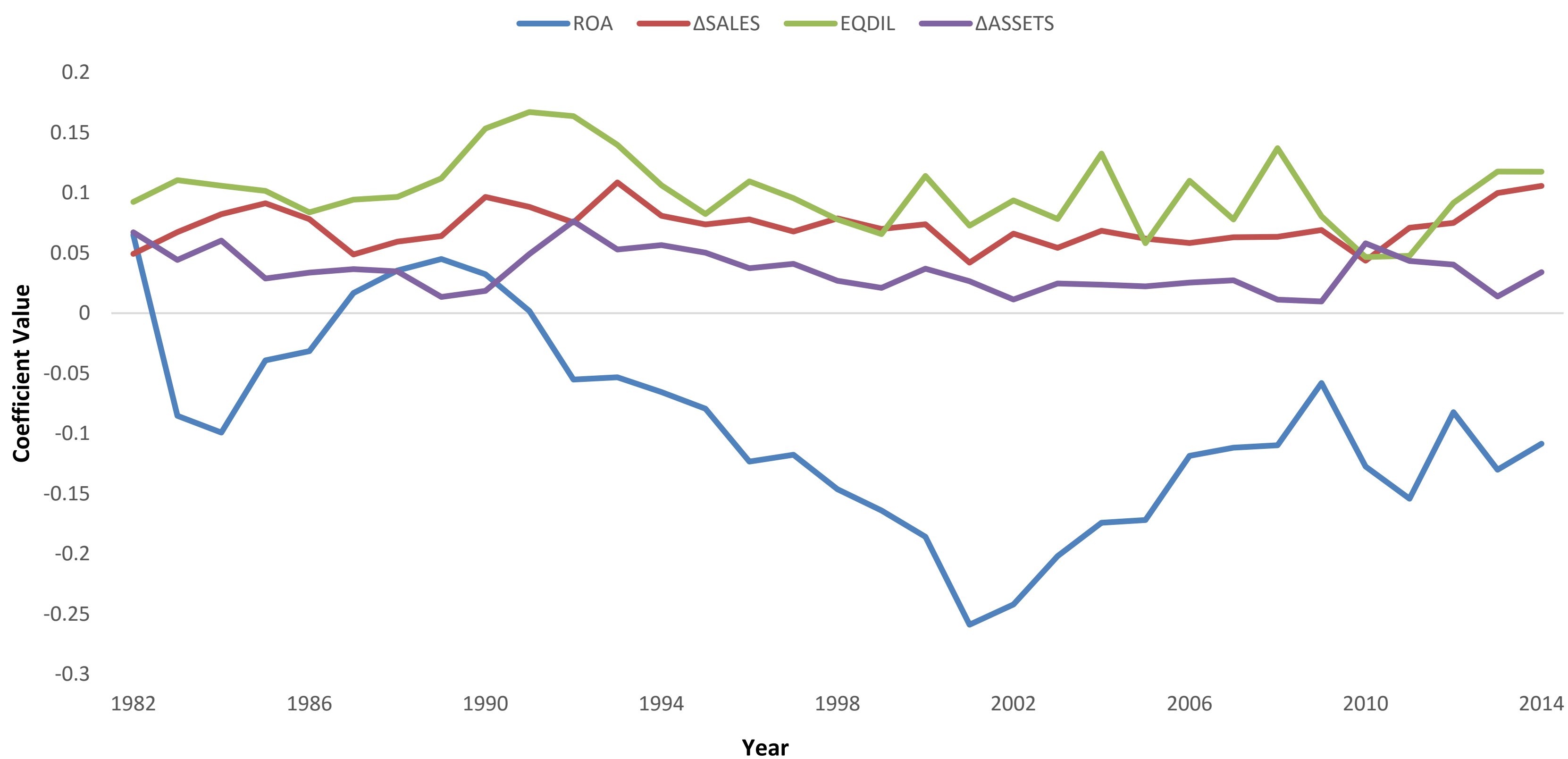


Table 3. Panel Regression Explaining Price-to-Book and Price-to-Forward Earnings Valuation Ratios from 1982 to 2014. The dependent variable for the regression is either the natural log of P/B ratio (Panel A) or the natural log of the P/E_{t+1} ratio (Panel B). P/B (price/book ratio) is market capitalization as of December of year t-1, divided by book equity in year *t-1*. P/E_{t+1} (price/forward earnings ratio) is price per share divided by fiscal year 1 analyst consensus earnings per share measured in the 3rd week of June of year *t*. EqDil (equity dilution) is the percentage change in split-adjusted shares outstanding from fiscal year-end in *t-2* to *t-1*. ΔSales (sales growth) is the percentage change in revenues per split-adjusted share from *t-2* to *t-1*. ΔAssets (asset growth) is the percentage change in assets per split-adjusted share from *t-2* to *t-1*. ROA (profitability) is operating income in *t-1* divided by assets for *t-1*, Hard Growth is the fitted value from the last regression listed in Table 2 and Soft Growth is equal to LTG minus Hard Growth. The independent variables are constructed using financial statement data from the fiscal period ending in year t-1. N is the average of firms each year. For brevity, the intercept is not reported. Robust standard errors are clustered by firm and industry.

Panel A. P/B

	EQDIL	ΔSALES	ΔASSETS	ROA	Hard Growth	Soft Growth	R^2	Industry Fixed Effect?	Year Fixed Effect?	N
Coefficient	0.38	0.40	0.26	1.60			0.11	No	No	2,213
t-stat	(5.98)	(6.18)	(7.16)	(2.59)						
Coefficient	0.33	0.40	0.26	1.81			0.20	No	Yes	2,213
t-stat	(4.43)	(6.53)	(7.46)	(3.02)						
Coefficient	0.33	0.31	0.22	1.71			0.21	Yes	No	2,213
t-stat	(5.06)	(7.75)	(9.95)	(2.82)						
Coefficient	0.28	0.31	0.22	1.85			0.29	Yes	Yes	2,213
t-stat	(3.84)	(7.92)	(9.38)	(3.11)						
Coefficient					2.02	3.74	0.16	No	No	2,213
t-stat					(3.14)	(11.74)				
Coefficient					1.38	3.01	0.27	Yes	Yes	2,213
t-stat					(2.89)	(11.73)				

Panel B. P/E$_{t+1}$

	EQDIL	ΔSALES	ΔASSETS	ROA	Hard Growth	Soft Growth	R^2	Industry Fixed Effect?	Year Fixed Effect?	N
Coefficient	0.21	0.06	0.14	-0.62			0.02	No	No	2,022
t-stat	(5.34)	(0.90)	(3.16)	(0.86)						
Coefficient	0.21	0.06	0.14	-0.43			0.13	No	Yes	2,022
t-stat	(4.94)	(0.84)	(3.87)	(0.61)						
Coefficient	0.14	0.01	0.09	-1.25			0.14	Yes	No	2,022
t-stat	(3.39)	(0.16)	(2.71)	(3.69)						
Coefficient	0.14	0.01	0.10	-1.10			0.23	Yes	Yes	2,022
t-stat	(3.05)	(0.12)	(3.41)	(3.53)						
Coefficient					2.20	2.80	0.14	No	No	2,022
t-stat					(3.44)	(7.85)				
Coefficient					2.10	2.32	0.28	Yes	Yes	2,022
t-stat					(4.24)	(8.39)				

Figure 3. Value-weighted Average Market-Adjusted Return for Portfolios Formed on Realized EPS Growth Rate from 1982 to 2009. At the end of June of year *t*, stocks are allocated to ten portfolios according to realized EPS growth rate (REAL EPS). The figure reports the average value-weighted (using market capitalization as of the end of June in year *t*), market-adjusted five-year return measured over the 60 months starting in July of year *t*. There is an average of 1,498 firms per year with non-missing five-year EPS growth rates.

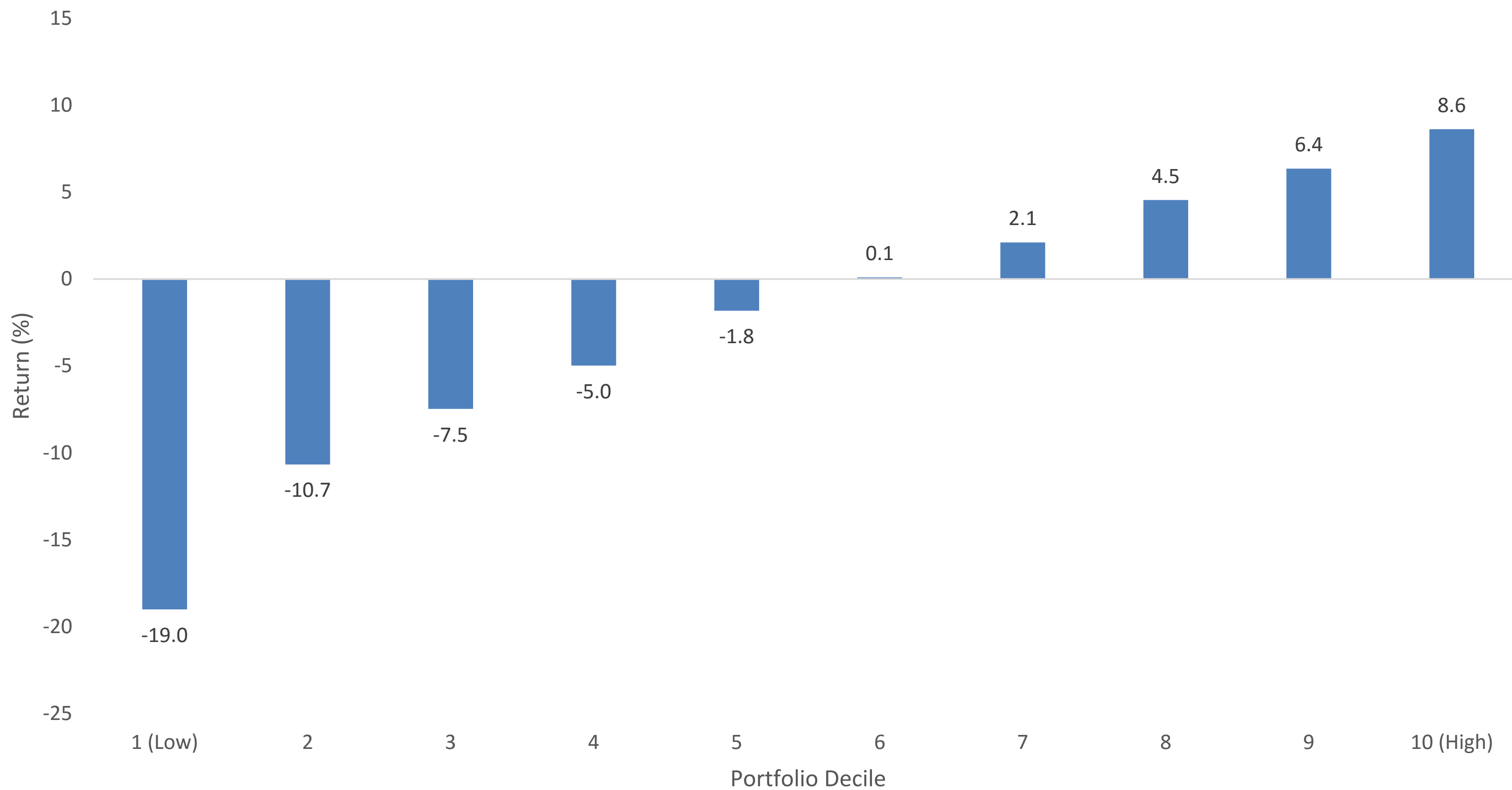


Figure 4. Histogram of Five-year Market-adjusted Returns with Missing EPS Five-year Growth Rates from 1982 to 2009. This figure reports the percentage of firm-years with missing realized earnings (REAL EPS) information, by market-adjusted return decile. There are 21,885 firm-years with future stock returns that have missing five-year EPS growth rates that were assigned EPS growth rates using our matching technique.

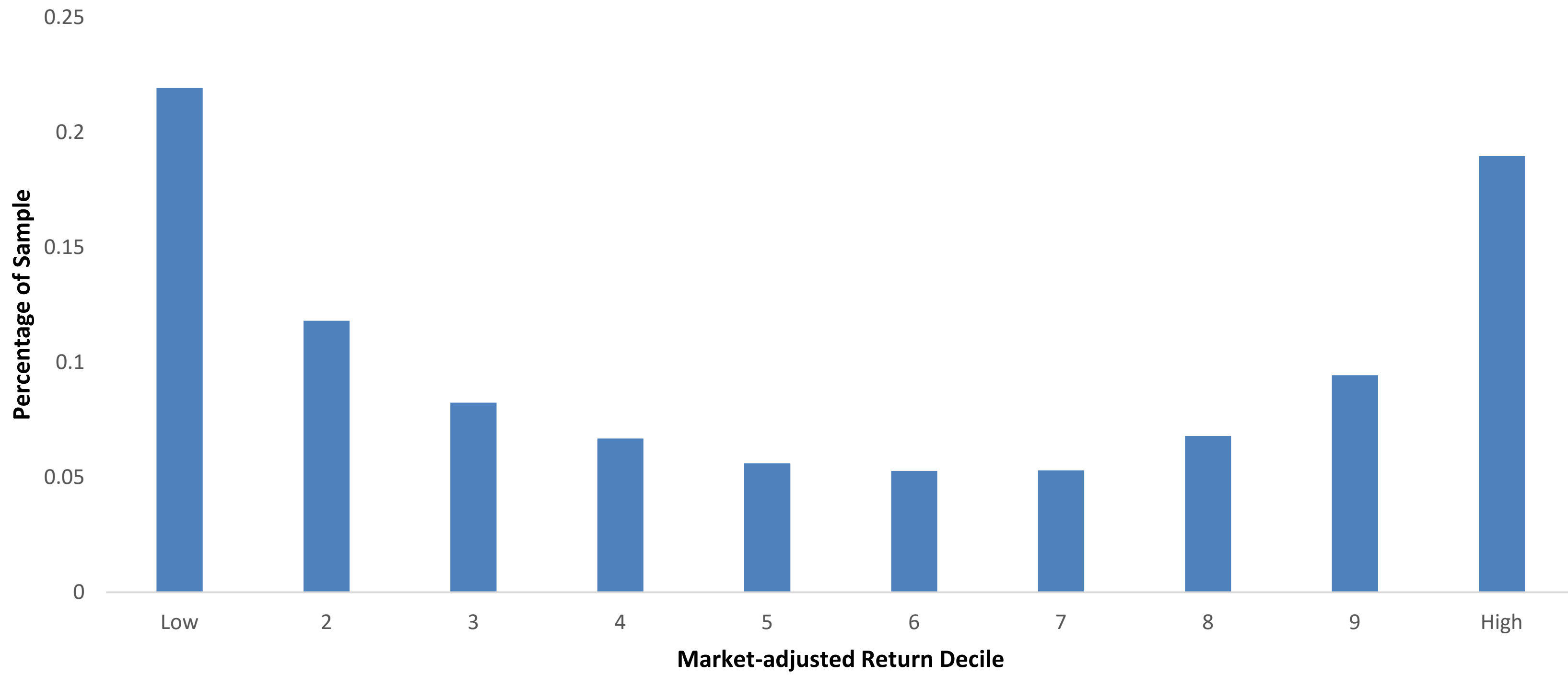


Table 4. Panel Regression Explaining Realized Earnings Growth from 1982 to 2014. The dependent variable for the regression is realized earnings growth (REAL EPS), which is the five-year annualized EPS growth rate. EQDIL is equity dilution measured as the percentage change in adjusted shares outstanding over the previous year. ΔSALES is the percentage change in split-adjusted revenues over the previous year. ΔASSETS is the percentage change in split-adjusted assets over the previous year. ROA is profitability, measured as operating income before depreciation divided by assets. LTG is measured as of the 3rd week in June of year t, while the independent variables are constructed using financial statement data from the fiscal period ending in year t-1. T-statistics, reported in parentheses, are based on robust standard errors that are clustered by firm and industry. For brevity, the intercept is not reported.

	LTG	EQDIL	ΔSALES	ΔASSETS	ROA	Hard Growth	Soft Growth	LN(P/B)	R^2	Ind & Year Fixed Effect?	N
Coefficient		-0.09	-0.02	-0.03	0.05			0.00	<.01	No	2,280
t-stat		(6.33)	(1.67)	(2.44)	(1.83)			(0.79)			
Coefficient		-0.10	-0.02	-0.03	0.12			-0.01	0.05	Yes	2,280
t-stat		(7.37)	(2.19)	(2.12)	(5.13)			(2.79)			
Coefficient	0.11	-0.10	-0.02	-0.03	0.07			-0.01	0.02	No	2,280
t-stat	(2.60)	(6.78)	(2.61)	(2.62)	(3.21)			(1.91)			
Coefficient	0.03	-0.10	-0.02	-0.03	0.13			-0.02	0.05	Yes	2,280
t-stat	(0.99)	(7.40)	(2.66)	(2.12)	(5.06)			(3.11)			
Coefficient						-0.52	0.11	-0.01	<.01	No	2,280
t-stat						(4.39)	(2.58)	(1.84)			
Coefficient						-0.61	0.02	-0.01	0.05	Yes	2,280
t-stat						(6.09)	(0.63)	(2.24)			

Figure 5. Persistence of Variables that Explain Growth from 1982 to 2009. This figure plots the average time-series Spearman correlation for different variables and their 1-, 2-, 3-, 4- and 5-year lag values using annual data. LTG measures the mean estimate of all analysts' expectations of the EPS annual growth rate between year *t+2* to year *t+5* measured in the 3rd week of June of year *t*. EQDIL (equity dilution) is the percentage change in split-adjusted shares outstanding from fiscal year-end in *t-2* to *t-1*. ΔSALES (sales growth) is the percentage change in revenues per split-adjusted share from *t-2* to *t-1*. ΔASSETS (asset growth) is the percentage change in assets per split-adjusted share from *t-2* to *t-1*. ROA (profitability) is operating income in *t-1* divided by assets for *t-1*. B/M (book/market ratio) is book equity in year *t-1* divided by market equity in December of *t-1*. P/B is market capitalization in December *t-1* divided by book equity in year *t-1*. P/E_{t+1} is the price per share in June *t*, divided by analyst EPS estimate for the next year *t+1*.

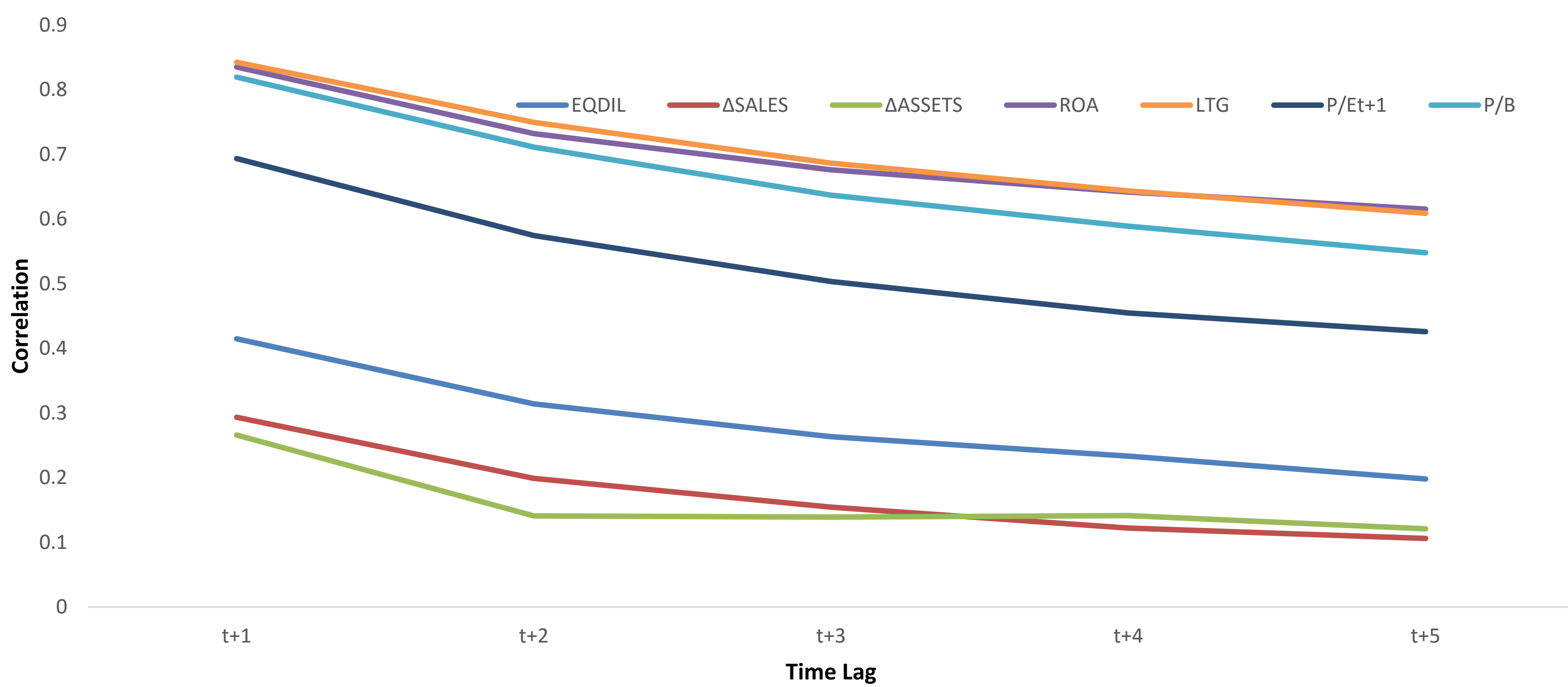


Table 5. Panel Regression Explaining Changes in Long-term Growth Estimates from 1982 to 2013. The dependent variable for the regression is the year-over-year change in analyst long-term growth forecasts ($LTG_{t+1} – LTG_t$) measured in the 3rd week of June of year *t*. EqDil (equity dilution) is the percentage change in split-adjusted shares outstanding from fiscal year-end in *t-2* to *t-1*. ΔSales (sales growth) is the percentage change in revenues per split-adjusted share from *t-2* to *t-1*. ΔAssets (asset growth) is the percentage change in assets per split-adjusted share from *t-2* to *t-1*. ROA (profitability) is operating income in *t-1* divided by assets for *t-1*, Hard Growth is the fitted value from the last regression listed in Table 2. The independent variables are constructed using financial statement data from the fiscal period ending in year t-1. N is the average of firms each year. For brevity, the intercept is not reported. Robust standard errors are clustered by firm and industry.

	EQDIL	**ΔSALES**	**ΔASSETS**	**ROA**	**Hard Growth**	**R^2**	**Industry Fixed Effect?**	**Year Fixed Effect?**	**N**
Coefficient	-0.02	-0.02	-0.01	0.00		0.03	No	No	1,929
t-stat	(7.81)	(5.91)	(8.21)	(0.31)					
Coefficient	-0.02	-0.02	-0.01	0.00		0.05	No	Yes	1,929
t-stat	(8.44)	(6.13)	(7.85)	(0.11)					
Coefficient	-0.02	-0.02	-0.01	0.00		0.03	Yes	No	1,929
t-stat	(7.62)	(5.74)	(7.82)	(0.41)					
Coefficient	-0.02	-0.02	-0.01	0.00		0.05	Yes	Yes	1,929
t-stat	(8.31)	(5.91)	(7.32)	(0.25)					
Coefficient					-0.24	0.02	No	No	1,929
t-stat					(5.40)				
Coefficient					-0.23	0.05	Yes	Yes	1,929
t-stat					(6.30)				

Table 6. Value-weighted Monthly Returns for Portfolios Formed on Long Term Growth Measures from July 1982 to December 2014. At the end of June of year *t*, stocks are allocated to ten portfolios based on the decile breakpoints for LTG (analyst long-term growth estimate), Hard Growth (fitted values from the last regression in Table 2) and Soft Growth (LTG minus Explained Growth). Panel A presents results for the original sample of firms with non-missing LTG. Panel B presents results for the bottom half of firms in the original sample based on market capitalization at the end of June of each year. Panel C reports results for all firms listed in CRSP/Compustat (including those with missing LTG data) that have valid data to construct EQDIL, ΔSALES, ΔASSETS, ROA and positive book equity. T-statistics are reported in parentheses to the right of each estimate. Monthly returns are reported in percentages.

Panel A. Original Sample

	1	2	3	4	5	6	7	8	9	10	10-1	*t-stat*	n
LTG	1.14%	1.10%	1.15%	1.12%	1.03%	1.08%	1.13%	1.25%	0.89%	1.15%	0.01%	(0.02)	2,153
Hard Growth	1.19%	1.18%	1.07%	1.22%	1.08%	1.23%	0.95%	1.05%	1.04%	0.59%	-0.60%	(2.66)	2,153
Soft Growth	0.98%	1.06%	1.15%	1.06%	1.22%	0.96%	1.06%	1.21%	1.02%	1.31%	0.33%	(0.96)	2,153

Panel B. Small Firms Only

	1	2	3	4	5	6	7	8	9	10	10-1	*t-stat*	n
LTG	1.24%	1.29%	1.23%	1.30%	1.29%	1.39%	1.28%	1.10%	1.17%	1.06%	-0.18%	(0.54)	1,077
Hard Growth	1.41%	1.44%	1.49%	1.27%	1.28%	1.37%	1.13%	1.36%	1.12%	0.55%	-0.86%	(3.88)	1,077
Soft Growth	1.18%	1.18%	1.14%	1.24%	1.25%	1.28%	1.32%	1.32%	1.23%	1.22%	0.05%	(0.15)	1,077

Panel C. All Firms (Includes Missing LTG Data Firms)

	1	2	3	4	5	6	7	8	9	10	10-1	*t-stat*	n
Hard Growth	1.16%	1.18%	1.11%	1.12%	1.11%	1.20%	1.02%	0.99%	0.98%	0.37%	-0.79%	(3.38)	4,045

Table 7. Fama-MacBeth Regressions of Monthly Returns on Growth, Size and Book/Market Measures from July 1982 to December 2014. This table reports the results of a set of Fama-MacBeth regressions of monthly returns on lagged growth measures, equity dilution, sales and asset growth, profitability, size and the book-to-market ratio. N is the average number of firms in the sample each year. LTG is the mean estimate of all analysts' expectations of the EPS annual growth rate between year *t+2* to year *t+5* measured in the 3rd week of June of year *t*. EQDIL (equity dilution) is the percentage change in split-adjusted shares outstanding from fiscal year-end in *t-2* to *t-1*. ΔSALES (sales growth) is the percentage change in revenues per split-adjusted share from *t-2* to *t-1*. ΔASSETS (asset growth) is the percentage change in assets per split-adjusted share from year *t-2* to year *t-1*. ROA (profitability) is operating income in year *t-1* divided by assets in year *t-1*. LN (Size) is the natural log of the market capitalization. LN (P/B) is the natural log of the price-to-book ratio. Hard Growth is the fitted value from the last regression listed in Table 2 and Soft Growth is equal to LTG minus Hard Growth. N is the average number of stocks each year. Certain regressions use industry dummies (based on GIC's 10 sector definitions). T-statistics are reported in parentheses to the right of each estimate and are based on Newey West corrected standard errors with a lag of 12 months. Monthly returns are reported in percentages.

	1		**2**		**3**		**4**		**5**		**6**	
Intercept	0.016	(2.18)	0.019	(2.52)	0.015	(2.16)	0.013	(2.33)	0.019	(2.72)	0.017	(3.19)
LTG	0.002	(0.17)			0.012	(1.25)	0.007	(1.11)				
EQDIL			-0.014	(5.25)	-0.015	(5.58)	-0.013	(5.62)				
ΔSALES			-0.002	(1.86)	-0.003	(2.12)	-0.003	(3.13)				
ΔASSETS			-0.005	(4.39)	-0.005	(4.51)	-0.005	(4.55)				
ROA			0.009	(1.66)	0.010	(2.18)	0.015	(2.96)				
Hard Growth									-0.072	(4.65)	-0.079	(5.54)
Soft Growth									0.012	(1.20)	0.007	(0.97)
Ln(SIZE)	0.000	(0.43)	0.000	(0.95)	0.000	(0.59)	0.000	(0.43)	0.000	(0.60)	0.000	(0.49)
Ln(P/B)	-0.001	(1.98)	-0.001	(1.01)	-0.002	(2.39)	-0.001	(1.98)	-0.002	(2.26)	-0.002	(2.81)
Ind Fixed Effect?	No		No		No		Yes		No		Yes	
R^2	0.03		0.03		0.04		0.08		0.03		0.08	
N	2,154		2,154		2,154		2,154		2,154		2,154	

Table 8. Pre- and Post-Global Settlement (August 2002) Split-Sample Regressions and Value-weighted Portfolio Returns from July 1982 to December 2014. This table replicates key results in earlier tables for different sample periods. Pre-GS refers to the period from July 1982 to August 2002, and post-GS refers to the period from September 2002 to December 2014. Panel A displays panel regression results similar to Tables 3 and 4; Panels B and C display average value-weighted returns for portfolios formed on various growth forecasts similar to analysis presented in Table 5. LTG is the mean estimate of all analysts' expectations of the EPS annual growth rate between year *t+2* to year *t+5* measured in the 3rd week of June of year *t*. EQDIL (equity dilution) is the percentage change in split-adjusted shares outstanding from fiscal year-end in *t-2* to *t-1*. ΔASSETS (asset growth) is the percentage change in assets per split-adjusted share from year *t-2* to year *t-1*. LN (Size) is the natural log of the market capitalization. LN (P/B) is the natural log of the price-to-book ratio. Hard Growth is the fitted value from the last regression listed in Table 2 and Soft Growth is equal to LTG minus Hard Growth. N is the average number of stocks each year. The regressions in Panel A include year and industry fixed effects (based on GIC's 10 sector definitions). T-statistics reported are double-clustered by firm and industry. Monthly returns shown in Panels B and C are reported in percentages.

Panel A. Panel Regression Split-Sample Results

	Dependent Variable	EQDIL	ΔASSETS	Hard Growth	Soft Growth	LN (P/B)	R^2	Time Period	N	Table Reference
Coefficient	LN (P/B)	0.09	0.42				0.23	Pre-GS	2,250	3A
t-stat		(2.75)	(12.79)							
Coefficient	LN (P/B)	-0.37	0.62				0.20	Post-GS	2,140	3A
t-stat		(6.54)	(7.67)							
Coefficient	LN (P/B)			1.60	3.38		0.30	Pre-GS	2,250	3A
t-stat				(5.04)	(12.08)					
Coefficient	LN (P/B)			0.66	2.27		0.21	Post-GS	2,140	3A
t-stat				(0.49)	(8.18)					
Coefficient	LN (P/E_{t+1})	0.19	0.12				0.24	Pre-GS	2,078	3B
t-stat		(3.06)	(2.94)							
Coefficient	LN (P/E_{t+1})	0.36	-0.13				0.11	Post-GS	1,923	3B
t-stat		(3.85)	(2.57)							
Coefficient	LN (P/E_{t+1})			2.09	2.37		0.32	Pre-GS	2,078	3B
t-stat				(4.66)	(12.28)					
Coefficient	LN (P/E_{t+1})			2.36	3.18		0.18	Post-GS	1,923	3B
t-stat				(2.21)	(4.85)					
Coefficient	REALEPS	-0.10	-0.03			-0.01	0.05	Pre-GS	2,255	4

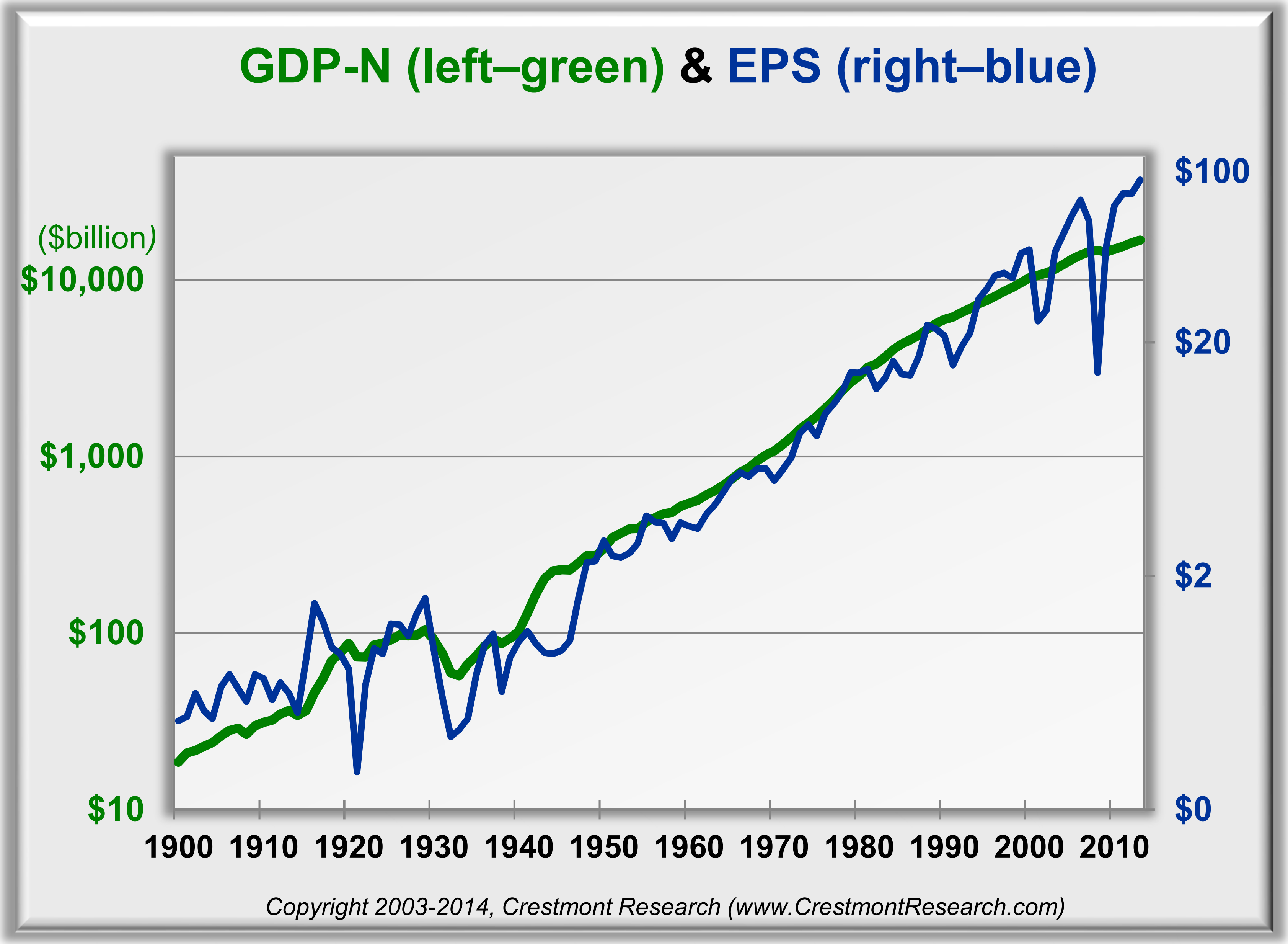
t-stat		(6.39)	(2.14)			(2.66)				
Coefficient	REALEPS	-0.13	-0.01			0.01	0.08	Post-GS	2,357	4
t-stat		(5.18)	(0.50)			(0.64)				
Coefficient	REALEPS			-0.57	0.04	-0.02	0.05	Pre-GS	2,255	4
t-stat				(4.77)	(0.90)	(3.70)				
Coefficient	REALEPS			-0.75	0.04	0.01	0.08	Post-GS	2,357	4
t-stat				(4.40)	(0.62)	(0.70)				
Coefficient	ΔLTG	-0.02	-0.02				0.04	Pre-GS	1,962	5
t-stat		(7.59)	(11.16)							
Coefficient	ΔLTG	-0.01	-0.02				0.03	Post-GS	1,842	5
t-stat		(3.54)	(11.54)							
Coefficient	ΔLTG			-0.24			0.05	Pre-GS	1,962	5
t-stat				(6.03)						
Coefficient	ΔLTG			-0.20			0.03	Post-GS	1,842	5
t-stat				(6.29)						

Panel B. Table 6 Pre-GS (July 1982 - August 2002)

	1	2	3	4	5	6	7	8	9	10	10-1	*t-stat*	N
LTG	1.30%	1.25%	1.37%	1.30%	1.23%	1.19%	1.35%	1.20%	0.84%	1.15%	-0.15%	(0.28)	2,173
Hard Growth	1.37%	1.31%	1.21%	1.37%	1.17%	1.46%	1.09%	1.22%	1.06%	0.63%	-0.74%	(2.25)	2,173
Soft Growth	1.15%	1.24%	1.36%	1.23%	1.37%	1.12%	1.12%	1.14%	1.11%	1.37%	0.23%	(0.48)	2,173
EQDIL	1.65%	1.40%	1.31%	1.21%	1.24%	1.43%	1.33%	1.05%	0.81%	0.75%	-0.90%	(4.47)	2,173
ΔASSETS	1.33%	1.21%	1.10%	1.48%	1.23%	1.22%	1.44%	1.29%	1.08%	0.81%	-0.53%	(1.78)	2,173

Panel C. Table 6 Post-GS (September 2002 – December 2014)

	1	2	3	4	5	6	7	8	9	10	10-1	*t-stat*	N
LTG	0.88%	0.85%	0.78%	0.83%	0.72%	0.91%	0.78%	1.35%	0.98%	1.15%	0.27%	(0.70)	2,122
Hard Growth	0.89%	0.98%	0.85%	0.98%	0.95%	0.87%	0.72%	0.76%	0.99%	0.53%	-0.36%	(1.49)	2,122
Soft Growth	0.72%	0.72%	0.80%	0.73%	0.92%	0.79%	1.02%	1.19%	0.98%	1.20%	0.48%	(1.20)	2,122
EQDIL	0.94%	0.68%	0.86%	0.92%	0.85%	1.10%	1.17%	0.80%	0.95%	0.51%	-0.43%	(1.81)	2,122
ΔASSETS	1.13%	1.26%	1.36%	1.26%	1.40%	1.07%	1.09%	1.22%	1.05%	1.38%	0.24%	(0.48)	2,122


GDP-N (left–green) & EPS (right–blue)
($billion)
$10,000
$1,000
$100
$10
$100
$20
$2
$0
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
2010
Copyright 2003-2014, Crestmont Research (www.CrestmontResearch.com)

	SFF		CBO		
	2015-2025	2015	0.042		
Real GDP	2.51	2016	0.046		Real GDP
CPI Inlfatio	2.14	2017	0.045		CPI Inlfation
	4.65	2018	0.042		
		2019	0.042		
		2020	0.042		
		2021	0.042		
		2022	0.042		
		2023	0.042		
		2024	0.042		
		2025	0.042		
			0.042636		

EIA		SSA	
2012-2040		2014-2090	
2.40%	2014		17557
2.10%	2090		499900
4.50%	76		
			4.5%

The real cost of equity

The inflation-adjusted cost of equity has been remarkably stable for 40 years, implying a current equity risk premium of 3.5 to 4 percent

Marc H. Goedhart, Timothy M. Koller, and Zane D. Williams

As central as it is to every decision at the heart of corporate finance, there has never been a consensus on how to estimate the cost of equity and the equity risk premium.[1]

Conflicting approaches to calculating risk have led to varying estimates of the equity risk premium from 0 percent to 8 percent—although most practitioners use a narrower range of 3.5 percent to 6 percent. With expected returns from long-term government bonds currently about 5 percent in the US and UK capital markets, the narrower range implies a cost of equity for the typical company of between 8.5 and 11.0 percent. This can change the estimated value of a company by more than 40 percent and have profound implications for financial decision making.

Discussions about the cost of equity are often intertwined with debates about where the stock market is heading and whether it is over- or undervalued. For example, the run-up in stock prices in the late 1990s prompted two contradictory points of view. On the one hand, as prices soared ever higher, some investors expected a new era of higher equity returns driven by increased future productivity and economic growth. On the other hand, some analysts and academics suggested that the rising stock prices meant that the risk premium was declining. Pushed to the extreme, a few analysts even argued that the premium would fall to zero, that the Dow Jones industrial average would reach 36,000 and that stocks would earn the same returns as government bonds. While these views were at the extreme end of the spectrum, it is still easy to get seduced by complex logic and data.

We examined many published analyses and developed a relatively simple methodology that is both stable over time and overcomes the shortcomings of other models. We estimate that the real, inflation-adjusted cost of equity has been remarkably stable at about 7 percent in the US and 6 percent in the UK since the 1960s. Given current, real long-term bond yields of 3 percent in the US and 2.5 percent in the UK, the implied equity risk premium is around 3.5 percent to 4 percent for both markets.

The debate

There are two broad approaches to estimating the cost of equity and market risk premium. The first is historical, based on what equity investors have earned in the past. The second is forward-looking, based on projections implied by current stock prices relative to earnings, cash flows, and expected future growth.

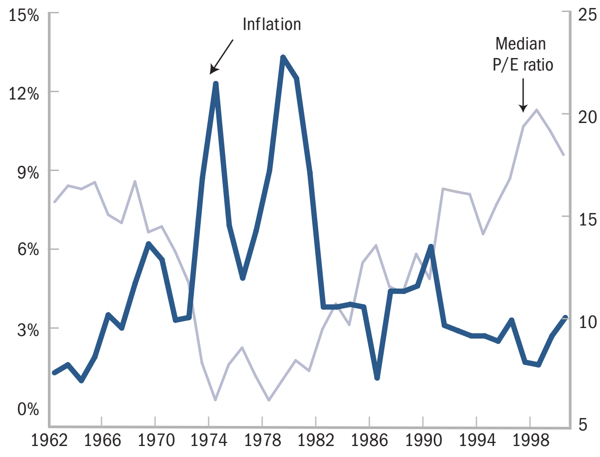
The latter is conceptually preferable. After all, the cost of equity should reflect the return expected (required) by investors. But forward-looking estimates are fraught with problems, the most intractable of which is the difficulty of estimating future dividends or earnings growth. Some theorists have attempted to meet that challenge by surveying equity analysts, but since we know that analyst projections almost always overstate the long-term growth of earnings or dividends,[2] analyst objectivity is hardly beyond question. Others have built elaborate models of forward-looking returns, but such models are typically so complex that it is hard to draw conclusions or generate anything but highly unstable results. Depending on the modeling assumptions, recently published research suggests market risk premiums between 0 and 4 percent.[3]

Unfortunately, the historical approach is just as tricky because of the subjectivity of its assumptions. For example, over what time period should returns be measured—the previous 5, 10, 20, or 80 years or more? Should average returns be reported as arithmetic or geometric means? How frequently should average returns be sampled? Depending on the answers, the market risk premium based on historical returns can be estimated to be as high as 8 percent.[4] It is clear that both historical and forward-looking approaches, as practiced, have been inconclusive.

Overcoming the typical failings of economic models

In modeling the behavior of the stock market over the last 40 years,[5] we observed that many real economic variables were surprisingly stable over time (including long-term growth in corporate profits and returns on capital) and that much of the variability in stock prices related to interest rates and inflation (Exhibit 1). Building on these findings, we

Exhibit 1. US median P/E vs. inflation



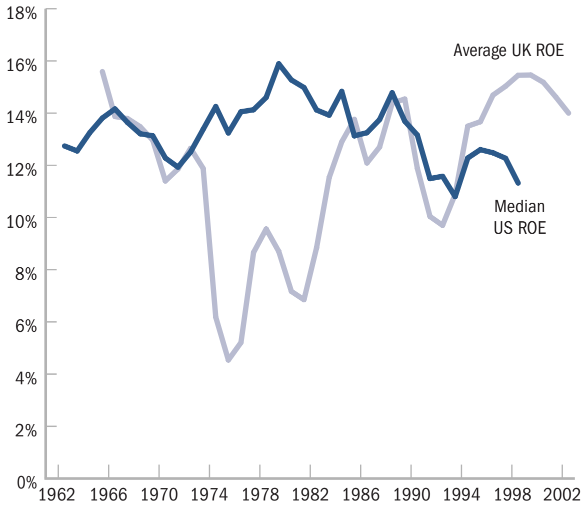
Source: McKinsey analysis

developed a simple, objective, forward-looking model that, when applied retrospectively to the cost of equity over the past 40 years, yielded surprisingly stable estimates.

Forward-looking models typically link current stock prices to expected cash flows by discounting the cash flows at the cost of equity. The implied cost of equity thus becomes a function of known current share values and estimated future cash flows (see sidebar, "Estimating the cost of equity"). Using this standard model as the starting point, we then added three unique characteristics that we believe overcome the shortcomings of many other approaches:

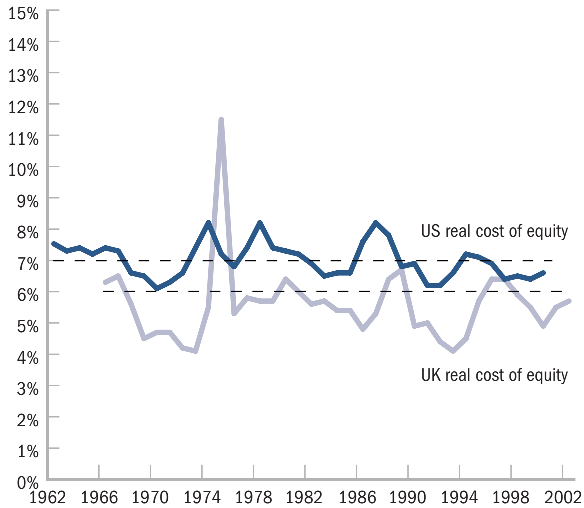
1. *Median stock price valuation.* For the US, we used the value of the median company in the S&P 500 measured by P/E ratio as an estimate of the market's overall valuation at any point in time. Most researchers have used the S&P 500 itself, but we argue that the S&P 500 is a value-weighted index that has been distorted at times by a few highly valued companies, and therefore does not properly

Exhibit 2. Return on book equity (ROE)



Source: McKinsey analysis

Exhibit 3. Annual estimates of the real cost of equity



Source: McKinsey analysis

reflect the market value of typical companies in the US economy. During the 1990s, the median and aggregate P/E levels diverged sharply. Indeed by the end of 1999, nearly 70 percent of the companies in the S&P 500 had P/E ratios below that of the index as a whole. By using the median P/E ratio, we believe we generate estimates that are more representative for the economy as a whole. Since UK indices have not been similarly distorted, our estimates for the UK market are based instead on aggregate UK market P/E levels.

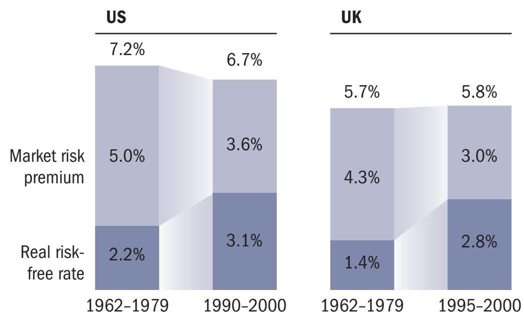
2. *Dividendable cash flows*. Most models use the current level of dividends as a starting point for projecting cash flows to equity. However, many corporations have moved from paying cash dividends to buying back shares and finding other ways to return cash to shareholders, so estimates based on ordinary dividends will miss a substantial portion of what is paid out. We avoid this by discounting not the dividends paid but the cash flows available to shareholders after new investments have been funded. These are what we term "dividendable" cash flows to investors that might be paid out through share repurchases as ordinary dividends, or temporarily held as cash at the corporate level.

We estimate dividendable cash flows by subtracting the investment required to sustain the long-term growth rate from current year profits. This investment can be shown to equal the projected long-term profit growth (See sidebar, "Estimating the cost of equity") divided by the expected return on book equity. To estimate the return on equity (ROE), we were able to take advantage of the fact that US and UK companies have had fairly stable returns over time. As Exhibit 2 shows, the ROE for both US and UK companies has been consistently about 13 percent per year,[6] the only significant exception being found in UK returns of the late 1970s.

3. *Real earnings growth based on long-term trends*. The expected growth rate in cash flow

The stability of the implied inflation-adjusted cost of equity is striking. Despite a handful of recessions and financial crises over the past 40 years . . . equity investors have continued to demand about the same cost of equity in inflation-adjusted terms.

Exhibit 4. Decomposition of the inflation-adjusted cost of equity



Source: McKinsey analysis

and earnings was estimated as the sum of long-term real GDP growth plus expected inflation. Corporate profits have remained a relatively consistent 5.5 percent of US GDP over the past 50 years. Thus, GDP growth rates are a good proxy for long-term corporate profit growth. Real GDP growth has averaged about 3.5 percent per year over the last 80 years for the US and about 2.5 percent over the past 35 years for the UK. Using GDP growth as a proxy for expected earnings growth allows us to avoid using analysts' expected growth rates.

We estimated the expected inflation rate in each year as the average inflation rate experienced over the previous five years.[7] The nominal growth rates used in the model for each year were the real GDP growth combined with the contemporary level of expected inflation for that year.

Results

We used the above model to estimate the inflation-adjusted cost of equity implied by stock market valuations each year from 1963 to 2001 in the US and from 1965 to 2001 for the UK (Exhibit 3). In the US, it consistently remains between 6 and 8 percent with an average of 7 percent. For the UK market, the inflation-adjusted cost of equity has been, with two exceptions, between 4 percent and 7 percent and on average 6 percent.

The stability of the implied inflation-adjusted cost of equity is striking. Despite a handful of recessions and financial crises over the past 40 years including most recently the dot.com bubble, equity investors have continued to demand about the same cost of equity in inflation-adjusted terms. Of course, there are deviations from the long-term averages but they aren't very large and they don't last very long. We interpret this to mean that stock markets ultimately understand that despite ups and downs in the broad economy, corporate earnings and economic growth eventually revert to their long-term trend.

We also dissected the inflation-adjusted cost of equity over time into two components: the inflation-adjusted return on government bonds and the market risk premium. As Exhibit 4 demonstrates, from 1962 to 1979 the expected

Estimating the cost of equity

To estimate the cost of equity, we began with a standard perpetuity model:

$$P_t = \frac{CF_{t+1}}{k_e - g} \tag{1}$$

where P_t is the price of a share at time t, CF_{t+1} is the expected cash flow per share at time $t + 1$, k_e is the cost of equity, and g is the expected growth rate of the cash flows. The cash flows, in turn, can be expressed as earnings, E, multiplied by the payout ratio:

$$CF = E(\textit{payout ratio})$$

Since the payout ratio is the share of earnings left after reinvestment, replacing the payout ratio with the reinvestment rate gives:

$$CF = E(1 - \textit{reinvestment rate})$$

The reinvestment rate, in turn, can be expressed as the ratio of the growth rate, g, to the expected return on equity:

$$\textit{reinvestment rate} = \frac{g}{ROE}$$

And thus the cash flows can be expressed as:

$$CF = E\left(1 - \frac{g}{ROE}\right) \tag{2}$$

We then combined formulas (1) and (2) to get the following:

$$\frac{P_t}{E_{t+1}} = \frac{1 - \frac{g}{ROE}}{k_e - g} \Rightarrow k_e = \frac{E_{t+1}}{P_t}\left(1 - \frac{g}{ROE}\right) + g \tag{3}$$

If the inflation embedded in k_e and g is the same, we can then express equation 3 as:

$$k_{er} = \frac{E_{t+1}}{P_t}\left(1 - \frac{g}{ROE}\right) + g_r \tag{4}$$

Where k_{er} and g_r are the inflation-adjusted cost of equity and real growth rate, respectively. We then solved for k_{er} for each year from 1963 through 2001, using the assumptions described in the text of the article.

inflation-adjusted return on government bonds appears to have fluctuated around 2 percent in the US and around 1.5 percent in the UK. The implied equity risk premium was about 5 percent in both markets.[8] But in the 1990s, it appears that the inflation-adjusted return on both US and UK government bonds may have risen to 3 percent, with the implied equity risk premium falling to 3 percent and 3.6 percent in the UK and US respectively.

We attribute this decline not to equities becoming less risky (the inflation-adjusted cost of equity has not changed) but to investors demanding higher returns in real terms on government bonds after the inflation shocks of the late 1970s and early 1980s. We believe that using an equity risk premium of 3.5 to 4 percent in the current environment better reflects the true long-term opportunity cost for equity capital and hence will yield more accurate valuations for companies. **MoF**

Marc H. Goedhart (Marc_Goedhart@McKinsey.com) *is associate principal in McKinsey's Amsterdam office,* ***Timothy M. Koller*** (Tim_Koller@McKinsey.com) *is a principal in McKinsey's New York office, and* ***Zane D. Williams*** (Zane_Williams @McKinsey.com) *is a consultant in McKinsey's Washington, D.C., office,*

[1] Defined as the difference between the cost of equity and the returns investors can expect from supposedly risk-free government bonds.

[2] *See* Marc H. Goedhart, Brendan Russel, and Zane D. Williams, "Prophets and profits?" *McKinsey on Finance*, Number 2, Autumn 2001.

[3] *See,* for example, Eugene Fama and Kenneth French, "The Equity Premium," *Journal of Finance*, Volume LVII, Number 2, 2002; and Robert Arnott and Peter Bernstein, "What Risk Premium is 'Normal'," *Financial Analysts Journal*, March/ April, 2002; James Claus and Jacob Thomas, "Equity premia as low as three percent?" *Journal of Finance*, Volume LVI, Number 5, 2001.

[4] *See,* for example, *Ibbotson and Associates,* Stock, Bonds, Bills and Inflation: 1997 Yearbook.

[5] *See* Timothy Koller and Zane Williams, "What happened to the bull market?" *McKinsey on Finance*, Number 1, Summer 2001.

[6] One consequence of combining a volatile nominal growth rate (due to changing inflationary expectations) with a stable ROE is that the estimated reinvestment rate varies tremendously over time. In the late 1970s, in fact, our estimates are near 100 percent. This is unlikely to be a true representation of actual investor expectations at the time. Instead, we believe it likely that investors viewed the high inflation of those years as temporary. As a result, in all of our estimates, we capped the reinvestment rate at 70 percent.

[7] This assumption is the one that we are least comfortable with, but our analysis seems to suggest that markets build in an expectation that inflation from the recent past will continue (witness the high long-term government bond yields of the late 1970s).

[8] There is some evidence that the market risk premium is higher in periods of high inflation and high interest rates, as was experienced in the late 1970s and early 1980s.

McKinsey on Finance is a quarterly publication written by experts and practitioners in McKinsey & Company's Corporate Finance & Strategy Practice. It offers readers insights into value-creating strategies and the translation of those strategies into stock market performance. This and archive issues of *McKinsey on Finance* are available on line at http://www.corporatefinance.mckinsey.com

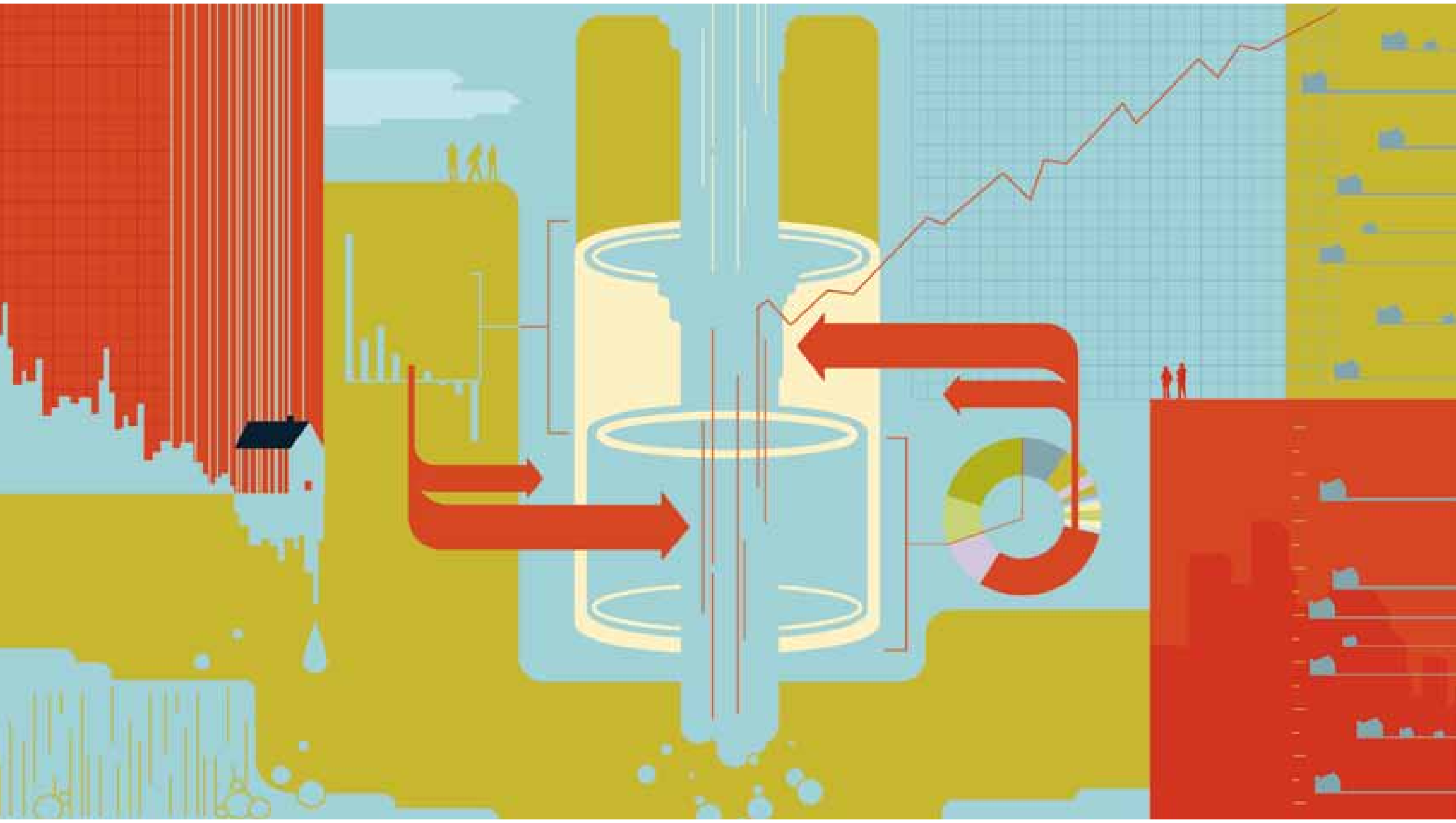
McKinsey & Company is an international management consulting firm serving corporate and government institutions from 85 offices in 44 countries.

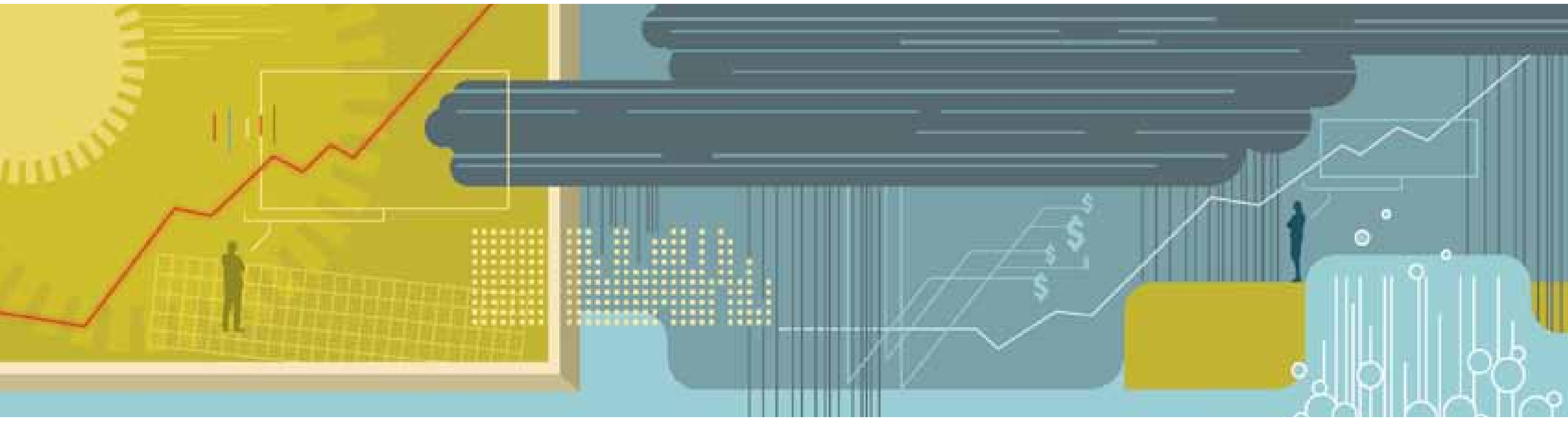
McKinsey&Company

McKinsey on Finance

Number 35,
Spring 2010

Perspectives on Corporate Finance and Strategy





Equity analysts: Still too bullish

After almost a decade of stricter regulation, analysts' earnings forecasts continue to be excessively optimistic.

Marc H. Goedhart, Rishi Raj, and Abhishek Saxena

No executive would dispute that analysts' forecasts serve as an important benchmark of the current and future health of companies. To better understand their accuracy, we undertook research nearly a decade ago that produced sobering results. Analysts, we found, were typically overoptimistic, slow to revise their forecasts to reflect new economic conditions, and prone to making increasingly inaccurate forecasts when economic growth declined.[1]

Alas, a recently completed update of our work only reinforces this view—despite a series of rules and regulations, dating to the last decade, that were intended to improve the quality of the analysts' long-term earnings forecasts, restore investor confidence in them, and prevent conflicts of interest.[2] For executives, many of whom go to great lengths to satisfy Wall Street's expectations in their financial reporting and long-term strategic moves, this is a cautionary tale worth remembering.

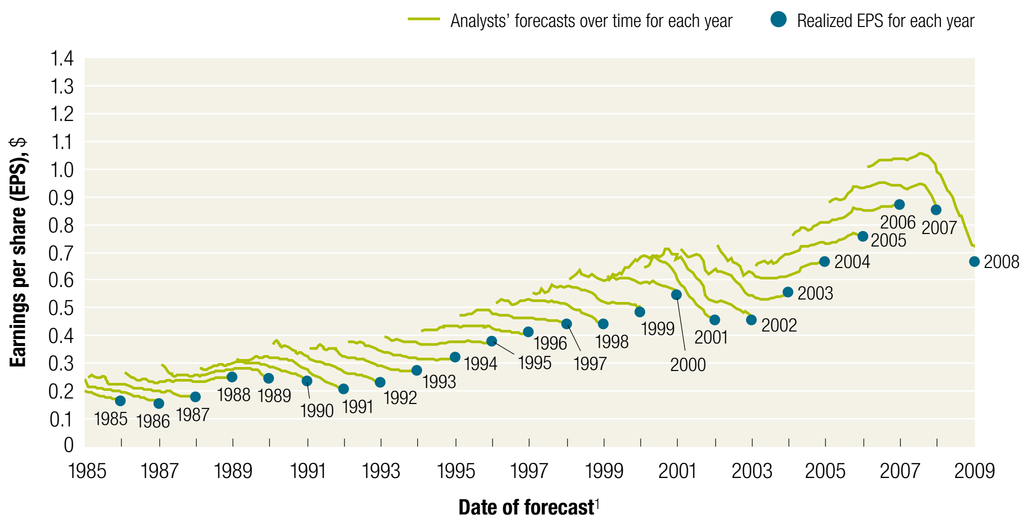
Exceptions to the long pattern of excessively optimistic forecasts are rare, as a progression of consensus earnings estimates for the S&P 500 shows (Exhibit 1). Only in years such as 2003 to 2006, when strong economic growth generated actual earnings that caught up with earlier predictions, do forecasts actually hit the mark.

Exhibit 1

Off the mark

With few exceptions, aggregate earnings forecasts exceed realized earnings per share.

S&P 500 companies



[1] Monthly forecasts.

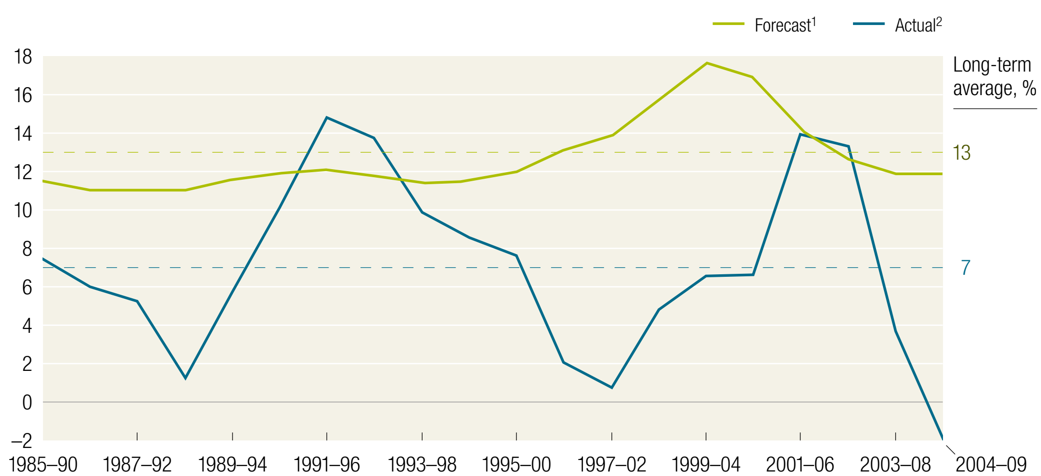
Source: Thomson Reuters I/B/E/S Global Aggregates; McKinsey analysis

Exhibit 2

Overoptimistic

Actual growth surpassed forecasts only twice in 25 years—both times during the recovery following a recession.

Earnings growth for S&P 500 companies, 5-year rolling average, %



[1] Analysts' 5-year forecasts for long-term consensus earnings-per-share (EPS) growth rate. Our conclusions are same for growth based on year-over-year earnings estimates for 3 years.

[2] Actual compound annual growth rate (CAGR) of EPS; 2009 data are not yet available, figures represent consensus estimate as of Nov 2009.

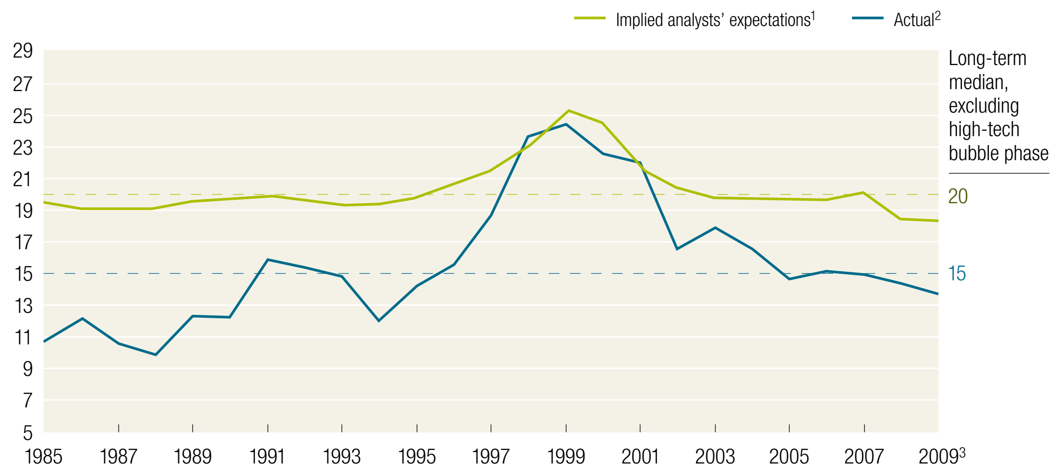
Source: Thomson Reuters I/B/E/S Global Aggregates; McKinsey analysis

Exhibit 3

Less giddy

Capital market expectations are more reasonable.

Actual P/E ratio vs P/E ratio implied by analysts' forecasts, S&P 500 composite index



[1] P/E ratio based on 1-year-forward earnings-per-share (EPS) estimate and estimated value of S&P 500. Estimated value assumes: for first 5 years, EPS growth rate matches analysts' estimates then drops smoothly over next 10 years to long-term continuing-value growth rate; continuing value based on growth rate of 6%; return on equity is 13.5% (long-term historical median for S&P 500), and cost of equity is 9.5% in all periods.
[2] Observed P/E ratio based on S&P 500 value and 1-year-forward EPS estimate.
[3] Based on data as of Nov 2009.

Source: Thomson Reuters I/B/E/S Global Aggregates; McKinsey analysis

This pattern confirms our earlier findings that analysts typically lag behind events in revising their forecasts to reflect new economic conditions. When economic growth accelerates, the size of the forecast error declines; when economic growth slows, it increases.[3] So as economic growth cycles up and down, the actual earnings S&P 500 companies report occasionally coincide with the analysts' forecasts, as they did, for example, in 1988, from 1994 to 1997, and from 2003 to 2006.

Moreover, analysts have been persistently overoptimistic for the past 25 years, with estimates ranging from 10 to 12 percent a year,[4] compared with actual earnings growth of 6 percent.[5] Over this time frame, actual earnings growth surpassed forecasts in only two instances, both during the earnings recovery following a recession (Exhibit 2). On average, analysts' forecasts have been almost 100 percent too high.[6]

Capital markets, on the other hand, are notably less giddy in their predictions. Except during the market bubble of 1999–2001, actual price-to-earnings ratios have been 25 percent lower than implied P/E ratios based on analyst forecasts (Exhibit 3). What's more, an actual forward P/E ratio[7] of the S&P 500 as of November 11, 2009—14—is consistent with long-term earnings growth of 5 percent.[8] This assessment is more reasonable, considering that long-term earnings growth for the market as a whole is unlikely to differ significantly from growth in GDP,[9] as prior McKinsey research has shown.[10] Executives, as the evidence indicates, ought to base their strategic decisions on what they see happening in their industries rather than respond to the pressures of forecasts, since even the market doesn't expect them to do so.

1 Marc H. Goedhart, Brendan Russell, and Zane D. Williams, "Prophets and profits," mckinseyquarterly.com, October 2001.

2 US Securities and Exchange Commission (SEC) Regulation Fair Disclosure (FD), passed in 2000, prohibits the selective disclosure of material information to some people but not others. The Sarbanes–Oxley Act of 2002 includes provisions specifically intended to help restore investor confidence in the reporting of securities' analysts, including a code of conduct for them and a requirement to disclose knowable conflicts of interest. The Global Settlement of 2003 between regulators and ten of the largest US investment firms aimed to prevent conflicts of interest between their analyst and investment businesses.

3 The correlation between the absolute size of the error in forecast earnings growth (S&P 500) and GDP growth is –0.55.

4 Our analysis of the distribution of five-year earnings growth (as of March 2005) suggests that analysts forecast growth of more than 10 percent for 70 percent of S&P 500 companies.

5 Except 1998–2001, when the growth outlook became excessively optimistic.

6 We also analyzed trends for three-year earnings-growth estimates based on year-on-year earnings estimates provided by the analysts, where the sample size of analysts' coverage is bigger. Our conclusions on the trend and the gap vis-à-vis actual earnings growth does not change.

7 Market-weighted and forward-looking earnings-per-share (EPS) estimate for 2010.

8 Assuming a return on equity (ROE) of 13.5 percent (the long-term historical average) and a cost of equity of 9.5 percent—the long-term real cost of equity (7 percent) and inflation (2.5 percent).

9 Real GDP has averaged 3 to 4 percent over past seven or eight decades, which would indeed be consistent with nominal growth of 5 to 7 percent given current inflation of 2 to 3 percent.

10 Timothy Koller and Zane D. Williams, "What happened to the bull market?" mckinseyquarterly.com, November 2001.

Marc Goedhart (Marc_Goedhart@McKinsey.com) is a consultant in McKinsey's Amsterdam office; **Rishi Raj** (Rishi_Raj@McKinsey.com) and **Abhishek Saxena** (Abhishek_Saxena@McKinsey.com) are consultants in the Delhi office.

Big Issues for Investors to Think About

Jim O'Neill
Chairman,
Goldman Sachs
Asset Management
jim.oneill@gs.com
+44(20)7774-2699

Anna Stupnytska
anna.stupnytska@gs.com
+44(20)7774-5061

James Wrisdale
james.wrisdale@gs.com
+44(20)7774-5282

Introduction

Judging from the questions we are getting, many investors are already thinking about 2013 as we move into the fourth quarter. As such, in our remaining editions of *Monthly Insights* for this year, we will update our 2013 GDP growth forecasts, introduce our first look at 2014 and provide a summary of our new *Strategy Series* on Global Fixed Income Benchmarking to be published later in the year. Ahead of these plans and against the background of 2013, we provide some supplementary thoughts to our September *Monthly Insights*, in which we assessed our views about the main perceived and widely discussed 'risks' in the global markets and how investors should be thinking about them in terms of their approach to asset allocation.

In the context of all the perceived uncertainties about the current and future investment environment, we take a closer look at the Equity Risk Premium (ERP) as a tool for determining asset allocation views. We show that, despite the rally in equity markets through the summer, the ERP still seems to be rather high. While there are certainly a number of caveats to using this metric, including the issues related to the concept of the risk-free rate, we argue that for medium- to long-term investors, it remains a good time to be favouring equities over bonds in asset allocation decisions. This view is also generally supported by one of the most conservative equity valuation metrics.

We also touch on the popular topic of regulation, especially as it relates to its uncertain impact on the likelihood of future economic recovery and the attraction of equities relative to 'safe' assets. In addressing this, we highlight the key distinction between better regulation and more regulation. We also argue that concerns about the regulatory environment seem to be currently holding back the risk appetite of both corporates and financial investors, in spite of the lower rate environment driven by central banks.

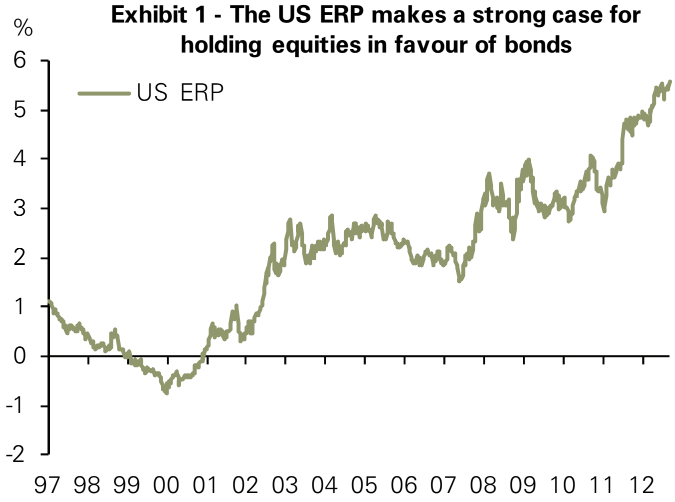
Finally, once more, we briefly review the three big macro issues that are particularly important: the US economic outlook, the Euro area crisis and China's economic adjustment from 'quantity' to 'quality' of growth. In the US, the recent positive signs in the data are somewhat offset uncertainty about the fiscal cliff. In the Euro area, while tricky issues are again back to the fore, it seems to us that both the ECB and Europe's key policymakers continue to be committed to muddling through their considerable challenges. In China, the cyclical environment remains disappointing but scope for monetary easing and upside growth surprises remains wide. Even taking into account all the risks, China's equity market seems attractively valued.

Asset Allocation in an Uncertain World

A frequently repeated assertion from client meetings is that the current investing environment is highly uncertain and, as a result, it is an especially difficult time to assess asset allocation and styles of investing. While we believe part of this concern is warranted, we often wonder whether this is more a reflection of people's state of mind rather than anything that is truly different. At the core of this comment is the following question we encourage all investors to ask: Is the environment really so different now that everything is more uncertain, or is it simply that, as a result of events since 2008, we are now all aware that so many things are uncertain? And perhaps as a related point, in hindsight we now know that things were not as certain as we originally believed over the last decade or, more specifically, for much of the time from 2001 to 2007.

From an investment perspective, we relate the uncertainties about the future and investment alternatives directly to the concept of the ERP. The ERP is generally defined as the excess equity return being offered at any moment in time beyond that of the risk-free rate. *Table 1* shows our estimate of the latest ERP globally and in some key regions of the world. A positive ERP is generally offered to compensate investors for the risk of holding equities over the risk-free rate. A rising ERP, as has been observed over the last five years globally, can generally be regarded as a symptom of investors requiring a higher rate of return to compensate for the (perceived) growing risks. A high ERP, especially relative to some past norm, suggests that investors perceive the future to be riskier. So with respect to discussions with clients, the real question is: Is the ERP high (and attractive) enough to compensate for all these risks that we talk about? We believe that there is quite a lot of evidence that the ERP has some mean reverting tendencies, and when the ERP is relatively high compared to some period in the past, then in fact it is usually a good time to invest in riskier assets, especially equities.[1] In this context, at the end of the 1990s the US ERP was rather low as *Exhibit 1* shows. This coincided with a period where many people did not seem to be as aware that there was a lot of uncertainty lying ahead. Put another way, is it more comforting for an investor when all the trustees or advisors are worried about a lot, or worried about very little?

Exhibit 1 - The US ERP makes a strong case for holding equities in favour of bonds



Source: Datastream and GSAM calculations

Today, despite the rally in equity markets through the summer, the ERP still seems to be rather high as we discussed in the September *Monthly Insights*. Our starting position is that this suggests for medium- to long-term investors, now is a good time to be favouring equities over bonds in asset allocation decisions.

Table 1 - Equity risk premia are elevated across the world

%	Real GDP Growth	Real Earnings Growth	+	Dividend Yield	=	Expected Real Return	-	Real Bond Yield	=	Implied ERP
US	2.5	2.5		2.2		4.7		-0.8		**5.5**
UK	2.3	2.3		3.5		5.8		-1.2		**7.0**
Europe ex UK	2.0	2.0		3.7		5.7		0.0		**5.6**
Japan	1.5	1.5		2.6		4.1		0.5		**3.6**
Brazil	5.0	5.0		4.2		9.2		3.4		**5.8**
China	8.0	8.0		4.1		12.1		0.5		**11.6**
India	8.0	8.0		1.5		9.5		4.1		**5.3**
Russia	5.0	5.0		4.0		9.0		1.7		**7.4**
GDP-weighted										
Advanced	2.1	2.1		2.9		5.0		-0.3		**5.3**
BRICs	7.0	7.0		3.8		10.8		1.7		**9.1**
World	3.5	3.5		3.1		6.6		0.2		**6.4**
PPP-weighted										
Advanced	2.2	2.2		2.8		5.0		-0.4		**5.4**
BRICs	7.3	7.3		3.6		10.9		1.7		**9.1**
World	4.2	4.2		3.1		7.3		0.5		**6.9**

Source: Datastream and GSAM calculations. As of 10/10/12

[1] "Finding 'Fair Value' in Global Equities: Part I", GS G*lobal Economics Paper* No. 179, February 2009 .

Supplementary Issues to the ERP Conundrum: The Risk-free Rate

Amongst many issues there are two that we think are important to consider. Firstly, many people often argue that the main reason the ERP is so high is primarily due to the fact that the so-called risk-free rate is so low. Along with that observation comes the question of whether the US (or other government) 10-year bond yield can actually be considered a genuine risk-free rate at such low levels and also whether we can have confidence in identifying a risk-free rate anymore. While aspects of these questions are connected, there are separate issues to consider.

In terms of the perceived low-level of both nominal and real 10-year bond yields in the US and globally, it is the case that they are low compared to the past. In a more forward-looking context, it may be the case that these low yields reflect some kind of pessimism about the economic potential of the future. The case of Japan gives some support for such a pessimistic conclusion, and if all of the US, Europe, Growth Markets and other parts of the world faced a future similar to the one Japan experienced from 1989 to 2012, then it would be difficult to treat such views lightly. In such a world, while the ERP is high compared to its past, it could be set to rise further as the future earnings growth from the world 'disappoints' suggesting that to invest in equities would become even riskier.

The second issue we wish to raise is very close to the ERP concept, and it concerns the notion of cyclically-adjusted Price to Earnings ratios (CAPEs) for many regional stock markets. *Table 2* shows our latest estimates. In our view, the CAPE approach is a rather conservative method of assessing the value of equity markets. As can be seen, despite the erratic rally since 2008, and including the rally of the last four months, several markets appear to be 'cheap' relative to their own history and also relative to expectations of future earnings. In particular, both continental Europe and a number of Growth Markets seem quite attractive. While, of course, there are some valid caveats to using this metric (just as with any other valuation method), we find the strong complementary signals of CAPEs and the ERP to be especially compelling. Of course, it is true that the future may be so bleak in so many different parts of the world, that future earnings will never be able to match earnings of the past. But you would have to be quite confident about this, rather than just 'worried'.

So What is the Risk-free Rate?

While it may be the case that 10-year bond yields are not a true indicator of the risk-free rate, the question is, what other financial measure can one find? It is possible, in fact perhaps likely, that the 10-year bond yield is still a good indicator for this purpose. Certainly the fact that in 2012, a number of government bond yields have dropped and some others—primarily those of peripheral Europe—have risen, suggests that markets do have the ability to differentiate and apply suitable risk premia to different governments.

Table 2 - Equity Valuations

	Latest CAPE	Forward PE	Deviation from avg.	CAPE vs FY1 PE
USA	22.9	12.8	22%	78%
Mexico	22.0	16.5	13%	33%
Indonesia	19.5	13.4	-10%	45%
Japan	18.8	11.6	-59%	62%
India	17.5	13.7	-21%	28%
Canada	16.9	13.0	-12%	30%
Australia	15.8	12.4	0%	27%
Korea	13.9	8.8	-16%	58%
China	12.6	9.0	-27%	41%
Germany	12.4	10.5	-35%	18%
UK	11.2	10.6	-18%	6%
France	10.3	10.5	-48%	-2%
Brazil	10.1	10.1	-34%	0%
Turkey	9.9	10.1	-35%	-2%
Italy	6.9	9.0	-69%	-24%
Russia	6.9	5.1	-51%	37%
Spain	6.4	10.7	-64%	-40%

Source: Datastream and GSAM calculations

Of course, some observers would respond that the genuine risk inherent in many Western bond markets (the US, UK, Japan amongst others) is only being curtailed or disguised by the fact that the central banks of these countries have become such active buyers of their own bond markets. Therefore, the argument goes that they are not appropriate measures of a risk-free rate. While one can understand why this observation is offered, it is a fact that central banks frequently influence the relative attractiveness of their own bond markets by their decisions, rather than it being just a feature of life since unconventional monetary policy or quantitative easing. By definition, when central banks adjust their interest rates as part of the 'norm', they will have an impact on their own bond markets.

In any case, in terms of selecting different markets for asset allocation purposes, if it is true that the actions of the Federal Reserve, Bank of Japan, Bank of England, European Central Bank (ECB) et al are artificially inflating bond prices, then it simply adds to our view that equities are relatively more attractive than bonds.

In this context, as and when, Western central banks decide that they can start to 'exit' from their current unconventional monetary policies and/or choose to raise short-term interest rates, it would seem quite likely that longer-term bond yields, including 10-year yields would rise. Some may argue that in terms of the currently high ERP, such an increase in bond yields would in fact reduce the ERP back to 'normal' levels without having any positive impact on equity markets, even if on a relative basis they perform better than bonds. We think that there is a danger of too much simplification in such an environment. Undoubtedly, on days when bond yields may rise significantly as part of some return to 'normality,' it would seem conceivable that equities may suffer. But presumably it would depend more specifically on why bond yields were rising and the valuation of specific equity markets at the time.

As we discussed in the September *Monthly Insights*, it is possible that if the US economy positively surprises in 2012, and survives the threat of the 'fiscal cliff' (more on this to follow), then the US ERP will return to normal, more by a rise in bond yields, than by a rally in equities. This is because US equities do not appear to be particularly cheap from a CAPE valuation perspective. But it is very questionable to apply the same thought process elsewhere. In the Euro area, for instance, one could easily imagine that a return to 'normality' in terms of German bond yields would probably be associated with a stabilisation of the Economic and Monetary Union (EMU) crisis, which would likely be rather positive for undervalued equities in the region. Similarly, Chinese equities that appear to be cheap would presumably cheer in light of clear evidence that Chinese GDP growth may achieve 7-8% in 2013 given all the current fears of a hard landing.

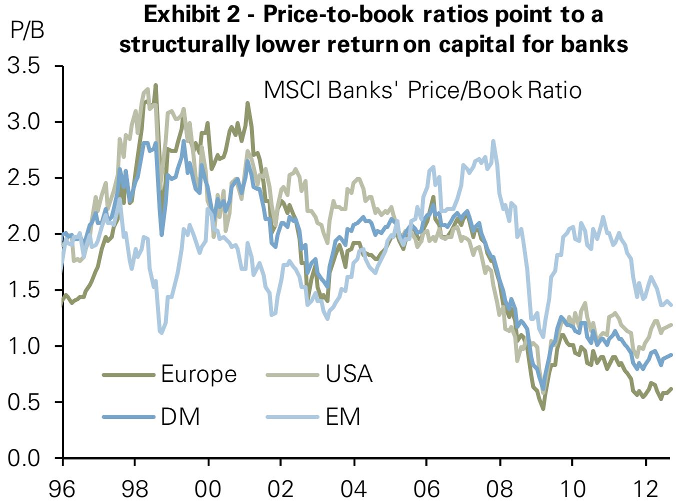
Regulatory Issues and Investing

One other issue that is frequently raised by investors is the broad topic of regulation, and whether increased regulation is both reducing the likelihood of future economic recovery and the attraction of equities relative to 'safe' assets. There are a number of aspects to the regulation debate, of which perhaps three seem especially critical. At the core of all the issues is the key distinction between better regulation and more regulation. We all would probably agree that better regulation would help sustainable growth but in the search for such, it is often tempting for policymakers to bias quantity over quality. More regulation can result in less economic growth if it stifles the risk-taking decision capability of corporate leaders and entrepreneurs.

In the post 2008 environment, the most vocal debate concerning regulatory issues obviously relates to banks, and the financial sector in general, in the US and Europe, with a myriad of implemented and planned changes. While critical aspects of various policies are quite different, the one common theme is essentially that financial intermediaries, especially banks that have a traditional role in lending money and accepting deposits, will be forced to hold more capital. As a result of this and some other changes, the return on capital for banks is widely expected to be less. Much—if not all—of finance theory would suggest that a lower return on capital is consistent with less economic growth. However, there are two pertinent arguments at this point in time that would challenge this common view. Firstly, if the regulatory changes contribute to a less volatile return on bank capital, even if it is perceived to be lower, might this not be helpful for the future of bank earnings or those that own their equity? And secondly, of course, markets have already taken onboard the anticipated path to a lower return on capital, or at least one would imagine judging from most bank equity valuation metrics as *Exhibit 2* shows. It is also probably worth adding that it is dangerous to generalise about all banks globally. The environment facing US-based banks is quite different from those in Europe (better capitalised perhaps), and the environment surrounding banks originating in Growth Markets is certainly different from that in the US and Europe.

A second regulatory concern relates to the broader behaviour of western corporate leaders in general and the current tendency of many to hold large amounts of cash, which one might imagine could be better deployed by investing, for acquisition or other uses. It often seems as though corporate leaders share the concerns of many financial investors. Frequently when quizzed, however, they seem eager to point out that their own company outlook is fine, but they worry about broad economic uncertainties, especially, in Europe an the US. Quite often, we discuss these issues at our regular internal CIO call and encourage our Fundamental Equity investment teams to ask their corporate contacts in a variety of industries: What would it take to make them want to spend and invest more? The answers invariably involve comments regarding uncertainty about the future, and as part of this, frequent reference to concerns about the regulatory environment both in terms of issues about taxation, disclosure and reporting standards. In our view, based on the frequency that we hear such comments, it indeed seems that this is one of the more legitimate concerns contributing to caution from corporate CEO's, especially in the US.

Exhibit 2 - Price-to-book ratios point to a structurally lower return on capital for banks



Source: Datastream

The third regulatory issue to consider is more specific to financial investors, and in particular, to pension funds and insurance companies post 2008. Aspects of this issue also relate to the previous two. The zest to impose the so-called mark-to-market accounting in recent years, while justifiable for the activities of financial institutions and in particular their trading books, seems much less rational when you ponder the purpose of pension funds and insurance companies, whose goals tend to include longer-term horizons and planning. Such institutions are natural holders of less liquid assets, which by definition are harder to value on any one day as there are no liquid markets for some of those assets. By introducing stricter regulatory guidelines to these investors, there appears to be some evidence that it discourages them from holding riskier assets than they otherwise might due to tougher capital requirements. In addition, as we learnt from our detailed survey of insurance companies globally, a notable minority are searching for higher-yielding investments to somehow help give them the returns necessary to satisfy their mandates in spite of the lower rates driven by central banks.

The Three Big Macro Issues

As we discussed in our September *Monthly Insights*, there are many macro issues that both warrant concern and of course get plenty of attention. Of them, we believe that there are three that are particularly important; the US economic outlook against the backdrop of the so-called fiscal cliff, the Euro area crisis and China's economic adjustment from 'quantity' to 'quality' of growth. Here is a brief summary of our latest thoughts on each.

The US Outlook

As we await the presidential election outcome in a few weeks with suddenly opinion polls looking rather tight and less certainty about the winner. What remains unclear, even if markets had confidence about the election outcome, is whether there will be a sizable and early compromise on fiscal policy in order to both reduce the risk of an undesirable significant fiscal tightening in 2013 and to have credible plans for medium- to longer-term fiscal consolidation. It is a very delicate balancing act, and it remains uncertain as to how policymakers will deal with both. We are assuming that some sort of compromise to avoid excessive 2013 tightening will be found, although it is not clear whether efforts to boost long-term fiscal credibility will be addressed.

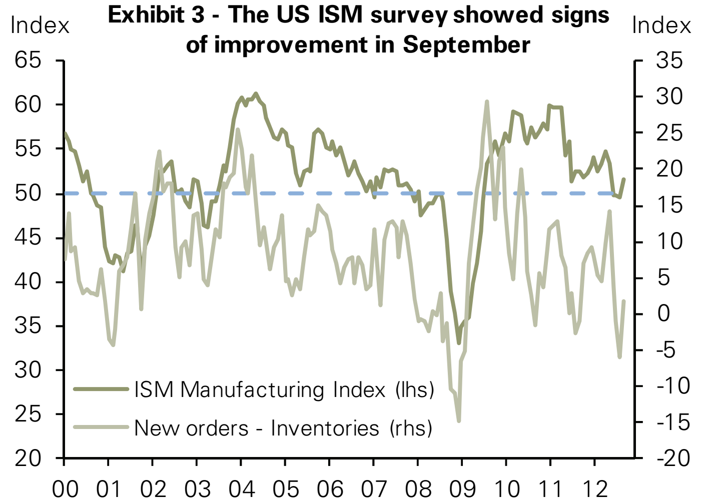
In the meantime, as always, we are watching all the incoming data releases, and since our September *Monthly Insights*, the news flow has been better. In addition to the surprising drop in US unemployment rate to 7.8%, both the manufacturing and service sectors ISM surveys positively surprised. In particular, the key new order and inventory components of the manufacturing survey showed a notable improvement as can be seen in *Exhibit 3*. While there are downside risks to the US outlook, there is also upside potential.

Euro Area Issues

Since the calm of August and early September, tricky issues are back to the fore in the Euro area. In particular, what to do about Greece and Spain in terms of their financing is at the top of the agenda, and the early signals about major agreed steps towards a banking union across the Euro area appear to be somewhat questionable. Each of those three topics, as well as other less predictable ones, could flare up further.

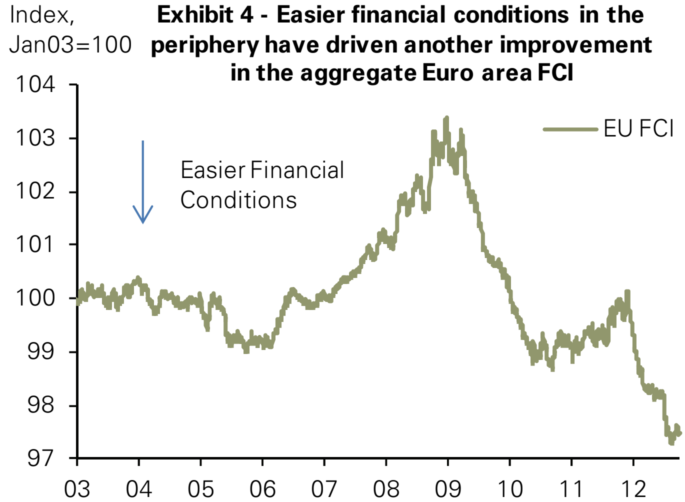
However, we believe that there are two key forces that investors need to remember when thinking about these (and other) tough challenges in the Euro area. Firstly, in announcing the Outright Monetary Transactions (OMT), the ECB made it quite clear that they intend to both offset any tendency of financial markets to impair the monetary mechanism of the ECB, and to reduce any implied Euro 'break-up' risk. Both of these concerns suggest that the ECB will fight more aggressively to avoid any fresh unwarranted tightening of the Euro area financial conditions. As can be seen in *Exhibit 4*, Euro area financial

Exhibit 3 - The US ISM survey showed signs of improvement in September



Source: Haver Analytics

Exhibit 4 - Easier financial conditions in the periphery have driven another improvement in the aggregate Euro area FCI



Source: GS Global ECS Research

conditions have eased considerably since the end of June. In this regard, it is interesting to see that the September Purchasing Managers' Index (PMI) surveys showed some signs of stabilisation of the Euro area recession, although given the scale of previous declines, one needs to be careful about giving too much attention to just one month. The detailed geographical breakdown did show some improvements in the so-called peripheral economies, notably Italy.

The second important issue is that the German Chancellor, Angela Merkel, despite considerable domestic opposition, has backed the ECB. We continue to interpret this decision as a 'more Europe' kind of judgement, which suggests that at a minimum, Europe's key policymakers will continue to try and muddle through their considerable challenges.

China

At the time of writing, we await the release of China's Q3 GDP estimate as well as all the September monthly economic data. Based on our favoured GS proprietary economic indicators, the GS China Activity Index (GSCA) and the China Financial Conditions Index (FCI), it seems as though the near-term economic data will surprise on the downside, as *Exhibit 5* illustrates. Our forecasts for Chinese GDP growth for both 2012 and 2013 are below the consensus.

Against this near-term headwind, there are three important forces that point in the opposite direction. Firstly, our long-term optimism towards China includes an assumption that China will 'only' grow by 7.1% this decade, and we have been assuming this for a few years. Once market participants get away from the expectation of 10%-type real GDP assumptions, this will make it easier for China to positively surprise. Secondly, in the context of expectations, financial conditions and leading indicators of cyclical growth, China's low inflation rate makes it quite likely that policymakers will be able to stimulate growth if necessary, consistent with their assumption of 7% GDP growth to deliver their 12th five-year plan. Thirdly, as shown earlier in the discussion of CAPE, China's equity market seems attractively valued, even taking into account all the risks.

Exhibit 5 - An easing in China financial conditions would provide support to economic activity

Source: GS Global ECS Research

Investing. Seeking Return and Minimising Risk

As we discussed in our September *Monthly Insights*, some investors have become so concerned with future economic and financial risks that they frequently think about so-called 'tail risks' and the notion of disaster hedging. We prefer to think about risk and return in the context of our ERP metric and more well-established valuation concept such as CAPE (although we would add that at current implied and actual volatilities, option-based strategies to protect downside risks are quite cheap).

In addition, against the background of a world economy increasingly being driven by economic activity in Growth Markets, our recommended approach to both global equity and fixed income investing is to seek equity and fixed income benchmarks with higher Growth Market allocations than in established market-cap based benchmarks. These alternative approaches offer more exposure to stronger growth and key fundamentals.

Appendix

GDP Growth Forecasts: GSAM vs Consensus

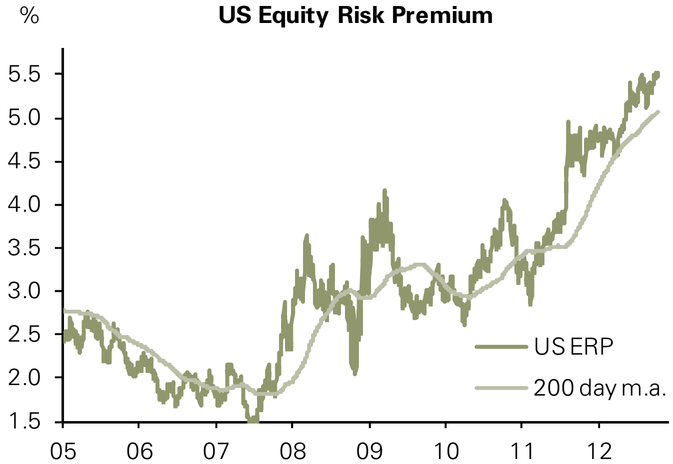
	Trend GSAM	2011	2012		2013	
			GSAM	Consensus*	GSAM	Consensus*
US	**2.5**	1.7	1.8	2.2	2.5	2.1
UK	**2.3**	0.7	-0.2	-0.3	1.8	1.3
Canada	**2.5**	2.5	2.2	2.0	2.5	2.0
Euroland	**2.0**	1.5	-0.4	-0.5	1.0	0.1
Japan	**1.5**	-0.7	1.9	2.4	1.3	1.3
Brazil	**5.0**	2.7	2.0	1.6	5.0	4.0
China	**8.0**	9.2	7.4	7.7	8.0	8.1
India	**8.0**	6.9	6.2	5.9	8.0	6.9
Russia	**5.0**	4.2	4.0	3.8	5.0	3.7
Mexico	**3.0**	3.9	3.6	3.9	3.8	3.5
Korea	**4.8**	3.6	2.5	2.6	4.8	3.5
Indonesia	**5.8**	6.5	6.0	6.1	6.1	6.1
Turkey	**5.0**	8.5	3.0	2.9	5.0	4.3
Advanced	**2.2**	1.3	1.0	1.2	1.8	1.3
BRICs	**7.3**	7.4	6.2	6.2	7.4	6.9
Growth Markets	**6.8**	7.0	5.7	5.7	6.9	6.4
World	**4.2**	3.8	3.0	3.2	4.1	3.6

*As of Sep 2012. Source: GSAM and Consensus Economics

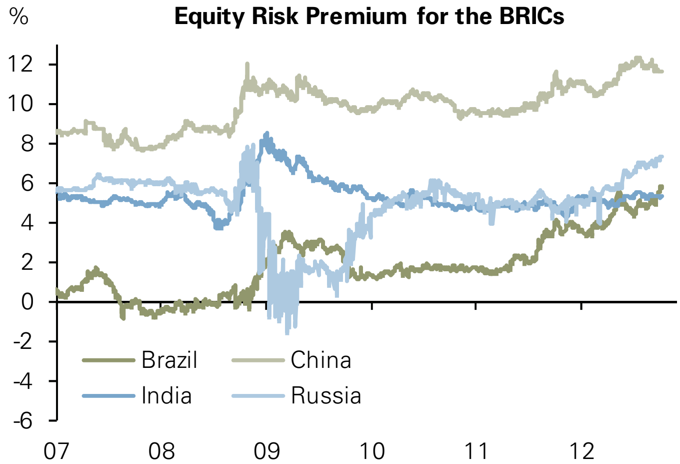
Global Equity Risk Premium*

	Real GDP Growth Trend	Real Earnings Growth	+	Dividend Yield	=	Expected Real Return	-	Real Bond Yield	=	Implied ERP	Expected Inflation	Expected Nominal Return
US	2.5	2.5		2.2		4.7		-0.8		**5.5**	2.0	6.7
UK	2.3	2.3		3.5		5.8		-1.2		**7.0**	2.0	7.8
Europe ex UK	2.0	2.0		3.7		5.7		0.0		**5.6**	2.0	7.7
Japan	1.5	1.5		2.6		4.1		0.5		**3.6**	1.0	5.1
Brazil	5.0	5.0		4.2		9.2		3.4		**5.8**	4.5	13.7
China	8.0	8.0		4.1		12.1		0.5		**11.6**	3.0	15.1
India	8.0	8.0		1.5		9.5		4.1		**5.3**	4.0	13.5
Russia	5.0	5.0		4.0		9.0		1.7		**7.4**	6.0	15.0
GDP-weighted												
Advanced	2.1	2.1		2.9		5.0		-0.3		**5.3**	1.8	6.8
BRICs	7.0	7.0		3.8		10.8		1.7		**9.1**	3.8	14.6
World	3.5	3.5		3.1		6.6		0.2		**6.4**	2.4	9.0
PPP-weighted												
Advanced	2.2	2.2		2.8		5.0		-0.4		**5.4**	1.9	6.9
BRICs	7.3	7.3		3.6		10.9		1.7		**9.1**	3.7	14.6
World	4.2	4.2		3.1		7.3		0.5		**6.9**	2.6	9.9

*As of 09 Oct 2012. Source: GSAM calculations



Source: GSAM calculations



Source: GSAM calculations

Disclaimers

For professional investors only—not for distribution to the general public.

This material does not constitute research, investment advice or trade recommendations. These views may not represent the views of GSAM's portfolio management teams, the Global Investment Research, and/or any other departments/divisions of Goldman Sachs and its affiliates which may differ.

This material has been prepared by GSAM and is not a product of Global Investment Research. Investors are urged to consult with their financial advisors before buying or selling any securities. Views and opinions are current as of the date of this presentation and may be subject to change. GSAM has no obligation to provide any updates or changes.

Prospective investors should inform themselves as to any applicable legal requirements and taxation and exchange control regulations in the countries of their citizenship, residence or domicile which might be relevant.

Past performance is not indicative of future results, which may vary. No part of this material may, without GSAM's prior written consent, be (i) copied, photocopied or duplicated in any form, by any means, or (ii) distributed to any person that is not an employee, officer, director, or authorised agent of the recipient.

There may be conflicts of interest relating to GSAM and its service providers, including Goldman Sachs and its affiliates, who are engaged in businesses and have interests other than that of managing, distributing and otherwise providing services to GSAM. These activities and interests include potential multiple advisory, transactional and financial and other interests in securities and instruments that may be purchased or sold by GSAM, or in other investment vehicles that may purchase or sell such securities and instruments. These are considerations of which investors should be aware. Additional information relating to these conflicts is set forth in GSAM's Conflicts of Interest Policy.

Any references to indices, benchmarks or other measures of relative market performance over a specified period of time are provided for your information only. Although certain information has been obtained from sources believed to be reliable, we do not guarantee its accuracy, completeness or fairness. We have relied upon and assumed without independent verification, the accuracy and completeness of all information available from public sources.

Economic and market forecasts presented herein reflect our judgment as of the date of this material and are subject to change without notice. These forecasts do not take into account the specific investment objectives, restrictions, tax and financial situation or other needs of any specific client. Actual data will vary and may not be reflected here. These forecasts are subject to high levels of uncertainty that may affect actual performance. Accordingly, these forecasts should be viewed as merely representative of a broad range of possible outcomes. These forecasts are estimated, based on assumptions, and are subject to significant revision and may change materially as economic and market conditions change. Goldman Sachs has no obligation to provide updates or changes to these forecasts. Case studies and examples are for illustrative purposes only and are not actual results. If any assumptions used do not prove to be true, results may vary substantially.

The concept of the ERP stems from a risk-return tradeoff, in which a higher rate of return is required to entice investors to take on riskier investments. The risk-free rate in the market is often quoted as the rate on longer-term government bonds, which are considered risk free because of the generally accepted notion that there is a low chance that the government will default on its loans. On the other hand, an investment in equities is far less guaranteed, as companies may suffer downturns or go out of business. The size of the premium will vary as the risk in a particular stock, or in the stock market as a whole, changes. High-risk investments are typically compensated with a higher premium.

Please note that neither Goldman Sachs Asset Management International nor any other entities involved in the Goldman Sachs Asset Management (GSAM) business maintain any licenses, authorisations or registrations in Asia (other than Japan), except that it conducts businesses (subject to applicable local regulations) in and from the following jurisdictions: Hong Kong, Singapore, Malaysia, Korea, and India. This material has been issued or approved by Goldman Sachs Canada, in connection with its distribution in Canada; in the United States by Goldman, Sachs & Co. This material has been issued for use in or from Hong Kong by Goldman Sachs (Asia) L.L.C, in or from Singapore by Goldman Sachs (Singapore) Pte. (Company Number: 198602165W), and in or from Korea by Goldman Sachs Asset Management Korea Co. Ltd. This material has been issued or approved in Japan for the use of professional investors defined in Article 2 paragraph (31) of the Financial Instruments and Exchange Law by Goldman Sachs Asset Management Co., Ltd. This material has been approved in the United Kingdom solely for the purposes of Section 21 of the Financial Services and Markets Act 2000 by Goldman Sachs Asset Management International, which is authorised and regulated by the Financial Services Authority (FSA). In Germany and Austria this document is presented to you by Goldman Sachs AG, regulated by the Bundesanstalt für Finanzdienstleistungsaufsicht (BaFin). In France this document is presented to you by Goldman Sachs Paris Inc. et Cie, regulated by the Autorité de Marchés Financiers (AMF). In Switzerland this document is presented to you by Goldman Sachs Bank AG, regulated by the Eidgenössische Finanzmarktaufsicht (FINMA). In Ireland this document is presented to you by Goldman Sachs Bank (Europe) plc, regulated by the Irish Financial Services Regulatory Authority (IFSRA). In Italy this document is presented to you by Goldman Sachs International, Italian Branch, regulated by the Commissione Nazionale per le Società e la Borsa (CONSOB). In Spain this document is presented to you by Goldman Sachs International, Spanish Branch, with the address at Calle Maria de Molina, 6, planta 5, regulated and registered at Comision Nacional Del Mercado de Valores (CNMV) with number 28.

Compliance code: 83309.STR.MED.OTU

BUSINESS
INSIDER

http://www.businessinsider.com/jim-oneill-global-macro-outlook-2013-4#current-erp-equity-risk-premium-levels-continue-to-indicate-that-equity-markets-are-still-quite-attractive-in-many-parts-of-the-world-22

JIM O'NEILL'S FINAL PRESENTATION: Goldman Guru Presents His View Of The World From Now Through 2020

Share

22/23

'Current ERP (equity risk premium) levels continue to indicate that equity markets are still quite attractive in many parts of the world.'

Goldman Sachs Asset Management

Current Estimates for the Equity Risk Premium*

	Real GDP Growth	Real Earnings Growth	+	Dividend Yield	=	Expected Real Return	-	Real Bond Yield	=	Implied ERP	Expected Inflation	Expected Nominal Return
US	2.5	2.5		2.1		4.6		-0.7		5.3	2.0	6.6
UK	2.3	2.3		3.2		5.5		-1.9		7.4	2.0	7.5
Europe ex UK	2.0	2.0		3.4		5.4		-0.1		5.5	2.0	7.4
Japan	1.5	1.5		1.7		3.2		-0.5		3.6	1.0	4.2
Brazil	5.0	5.0		3.7		8.7		3.9		4.8	4.5	13.2
China	7.5	7.5		4.1		11.6		0.4		11.2	3.0	14.6
India	7.5	7.5		1.6		9.1		3.9		5.2	4.0	13.1
Russia	5.0	5.0		4.8		9.8		0.8		9.0	6.0	15.8
GDP-weighted												
Advanced	2.2	2.2		2.5		4.7		-0.5		5.2	1.8	6.5
BRICs	6.7	6.7		3.8		10.5		1.5		9.0	3.8	14.3
World	3.5	3.5		2.9		6.4		0.0		6.3	2.4	8.8
PPP-weighted												
Advanced	2.2	2.2		2.5		4.7		-0.6		5.3	1.9	6.6
BRICs	6.9	6.9		3.6		10.5		1.6		9.0	3.7	14.3
World	4.1	4.1		3.0		7.0		0.3		6.7	2.6	9.6

* As of 11 April 2013

Source: GSAM Calculations

BEFORE THE FEDERAL COMMUNICATIONS COMMISSION

IN THE MATTER OF	)
	)
AMERICAN TELEPHONE AND TELEGRAPH COMPANY	) CC DOCKET NO. 79-63
	)
PETITION FOR MODIFICATION OF	)
PRESCRIBED RATE OF RETURN	)

PREPARED DIRECT TESTIMONY

DR. MYRON J. GORDON

AND

DR. LAWRENCE I. GOULD

APRIL, 1980

III. COST OF EQUITY CAPITAL

It is widely accepted that a public utility should earn a return on capital that allows it to raise the capital necessary to meet the demand for its services without an adverse effect on current shareholder stock. Such a rate of return is called the utility's cost of capital. A return in excess of that rate burdens the consumer with prices which are excessive and causes an unjustified transfer of income from the consuming public to the shareholders of the utility. It also encourages the utility to increase costs and prices further by overinvesting in plant facilities. On the other hand, a return on capital below the required return may discourage the utility from raising sufficient capital to meet demands for service, causing consumers to suffer an impairment in the quantity and quality of service. Therefore, if the return allowed by the Commission is either too high or too low, the result is less than satisfactory to the consumer. The testimony which follows is offered with a view to estimating as closely as possible the actual required return on capital (also called the cost of capital) and, with some care, to avoiding any bias in either direction.

In measuring the cost of capital from each source, the cost of debt and the cost of preferred capital pose few problems. It is clear that the utility must pay the embedded interest on its outstanding debt and the prescribed dividend on the preferred stock. Both of these measurements involve perfectly straightforward calculations. Somewhat more controversial is the problem of determining the cost of common equity capital.

A. General Principles

A utility's cost of common equity capital is the return or yield that investors on average require on its common stock as implied in the price that they are willing to pay to hold the stock. This implied yield is the cost of common equity capital, because the existing shareholders neither gain nor lose as a consequence of additional investment and financing, regardless of the method of financing, as long as the return the company earns on its common equity is equal to the return investors require on the stock. By contrast, when the allowed return on common is above the return investors require, each dollar of additional financing raises the value of the existing shares. Conversely, when the utility's operating income less interest on debt, income taxes, and preferred dividends does not leave a return on common equity equal to the return investors require on the stock, we not only have a depressed stock price because of the low return, but, in addition, each dollar of additional investment and financing further depresses the price.

The theoretical basis for the conclusion just stated has been fully developed,[1] but a simple analogy goes a long way in demonstrating the point. Ignoring operating costs, a bank that borrows at 8% and lends at 10% adds 2% of the amount borrowed and loaned to the earnings of the bank's shareholders. The more the bank borrows and lends with this 2% spread, the more it increases future earnings on and the current value of its common stock. The return that investors require on a utility's common stock is, in one form or another, what must be paid for additional equity funds, and if the company earns more on the money than it must pay to get the funds, the excess adds to the earnings on and value of the existing shares. Conversely, if the company earns a lower rate of return than it pays on additional funds, the difference comes out of the pockets of the existing shareholders.

While the management of a utility may not be able to prevent a regulatory agency from allowing it a rate of return on capital below its costs of capital, it will, quite understandably, be reluctant to compound the mis-

---

[1] For an extensive discussion, see M.J. Gordon, The Cost of Capital to a Public Utility, Michigan State University, East Lansing, Michigan, 1974.

fortunes of its shareholders by further depressing the stock price through undertaking further investment in the face of an inadequate return on capital. A difference between the return on capital and its cost is fully reflected in the return on common equity, since the bondholders and preferred shareholders are assured of receiving their prescribed returns on capital regardless of the allowed rate on total capital. However, the long-run dependence of the value of a public utility's stock on the service provided to its customers could make it advisable for the company to undertake essential capital expenditures in the face of a small and hopefully temporary unfavorable difference between the allowed rate of return and the cost of capital.

Management's own commitment to continued growth or its reluctance to face the problems of a sharp curtailment in growth may persuade it to continue a high rate of investment in the face of an unsatisfactory rate of return. However, this amounts to an appropriation of shareholder wealth in pursuit of managerial objectives, and sooner or later the shareholders may turn to a new management that is more solicitous of stockholder welfare.

B. <u>Measurement of DCF Cost of Equity Capital</u>

The principles used to measure the cost of common equity are the same as those used in measuring the yield which investors require on debt or the yield required on outstanding preferred stock. However, in the case of debt and preferred stock, the payments to investors are relatively certain and, thus, amenable to objective calculation. However, the future dividend payments on a share of stock are uncertain, and determination of the expected yield required by investors requires the use of a more complex, yet still relatively simple and very reliable, method for dealing with the problem at hand.

This method is called the DCF (Discounted Cash Flow) Method for computing the cost of equity capital.[1] It represents the valuation of a share of stock by the expression:

$$P_0 = \frac{D_1}{(1+k)} + \frac{D_2}{(1+k)^2} + \ldots + \frac{D_t}{(1+k)^t} + \ldots + \frac{D_\infty}{(1+k)^\infty} \; . \qquad (1)$$

In this expression:

P_0 = the current price per share;

D_t = the expected value of the dividend the share will pay at the end of period t; and

k = the yield or return investors require on the share.

---

[1] This method was developed by Myron J. Gordon in an article in <u>Management Science</u> in 1956 and was first introduced in testimony in the American Telephone and Telegraph Co. Case, F.C.C. Docket 16258, 1966.

If the future dividends are expected to grow at the rate of g each period, Equation (1) reduces to:

$$P_0 = \frac{D_1}{k-g} \tag{2}$$

Solving Equation (2) for k results in an expression for the yield that investors require:

$$k = \frac{D_1}{P_0} + g. \tag{3}$$

In other words, to measure the expected return that investors require we may take the sums of the dividend yield and the expected rate of growth in the dividend.

An alternate approach to Equation (1) for the price of a share is:

$$P_0 = \frac{D_1 + P_1}{1 + k} . \tag{4}$$

Here, we take as the future payments the next period's dividend and the end-of-period price. However, $P_1 = P_0(1+g)$, and this substitution plus a little algebra results in Equation (2). Hence, the two approaches to share valuation result in the same measurement equation for share yield.

In order to use Equation (3), we need to measure both the dividend yield and the expected rate of growth in the dividend.

1. Measurement of Dividend Yield

The term for dividend yield in the Eq. (3) expression for a share's yield is the forecast dividend for the coming period, D_1, divided by the current price, P_0. The value assigned to P_0 should be the price of the share at the time the share yield is being estimated. The rationale for using the current price is that at each point in time it reflects all the information available to a company's investors regarding future dividends. Hence, the yield investors require on any date is the discount rate that equates on that date the current price and the expected stream of future dividends. To use an average of share prices over some prior time period for P_0 would result in a value for k without meaning, that is, it would not provide the average value for k over the prior time period. Furthermore, to obtain an average value for k over some prior time period, one must average the values of share yield -- not of share price.

D_1 is the forecast dividend for the coming year if dividends are paid annually. Common practice, however, is to pay dividends quarterly, in which case D_t in Eq. (1), the fundamental expression for share price, is a quarterly dividend. The value of k that satisfies Eq. (1) is the quarterly yield on the share, and the g in Eqs. (2) and (3) is the quarterly rate at which the dividend is expected to grow.

Because it is customary and convenient to think in terms of annual and not quarterly figures for rate of return and growth statistics, annualized figures will be used here. Annualized figures are simply four times quarterly figures. That is, if the current price of a share is P_0=\$50,000, and if its forecast dividend for the coming quarter is D_1=\$1.25, the quarterly dividend yield is \$1.25/\$50.00=2.5%, and the annualized dividend yield is 10%.

We all know from bank advertisements that when interest is compounded more frequently than once a year, two annual interest rates may be computed. To illustrate, an interest rate of 15% per year with the interest compounded quarterly means that a dollar left on deposit for a year will have 3.75% added to the balance at the end of each quarter, and the balance in the account at the end of the year will be \$1.1587. In other words, a 15% interest rate compounded quarterly will earn interest equal to 15.87% of the balance at the start of the year.

What does this imply for arriving at a rate of return equal to the cost of equity capital? If the quarterly yield at which a public utility share sells is 3.75%, should the utility be allowed to earn for the year a rate of return on common equity of 15% or something more? The answer is: (1) more than 15%, if the rate of return the company earns is calculated on the basis of the common equity at the start of the year; and (2) only 15%, if the rate of return on common equity is calculated by averaging its values at the start and at the end of the year. This statement is proved in Schedule 27. The latter method represents common practice and the practice followed here. Hence, in arriving at the cost of equity capital, the correct figure for the dividend-yield term in Eq. (3) is the annualized value of the forecast dividend for the coming quarter divided by the current price.

2. Measurement of Expected Growth

A difficult problem is the determination of the long-run dividend growth expectations of investors. In other words, what is the expected rate of growth in future dividends per share, g, in which investors on average believe?

To solve the problem, it is essential to understand the determinants of long-run expected dividend growth. If a company is expected to earn a rate of return of r on its common equity, and if it retains the fraction b of its earnings, then each year its earnings per share can be expected to increase by the fraction br of its earnings per share in the previous year. Thus, br is an excellent measure of the expected rate of growth in future earnings per share. If the company is expected to have a stable retention ratio and, therefore, a stable dividend payout ratio, it follows that br is also an excellent measure of the expected rate of growth in future dividends per share. That is:

$$g = br. \tag{5}$$

This relationship is illustrated in Schedule 18. There the hypothetical initial common equity or book value per share = \$10.00, r = .10 and b = .4. The first period earnings are expected to be \$1.00 per share and the expected dividend is \$.60. The retained earnings raise the book value of equity to \$10.40 at the start of the second year, and r times that is \$1.04, which is equal to the earnings per share the second year. The dividend in the second year is expected to be \$.624, and so on through time. The earnings, dividends, and stock price are expected to grow at the rate br = (.4) (.10) = .04 in every future year.

If investors require an 8% return on the stock, the initial price is:

$$P_0 = \frac{D_1}{k-g} = \frac{\$.60}{.08-.04} = \$15.00. \tag{6}$$

Similarly, the expected share price after one year is:

$$P_1 = \frac{D_2}{k-g} = \frac{\$.624}{.08-.04} = \$15.60 \tag{7}$$

The price in subsequent periods rises by 4% as long as the yield investors require on the share remains equal to 8%.

In fact, a company's return and retention rates do not remain constant over time. However, if investors expect that a company will on average earn a return of r and retain the fraction b of its earnings, they will expect the dividends, earnings, and price to grow at a rate br due to retention of earnings.

Stock financing will be a further cause of expected growth if the company is expected to issue new shares and if the stock's market price is greater than book value. Conversely, when a company is expected to engage in stock financing through the sale of stock at share prices below book value, ignoring the stock financing results in an overestimate of growth and share yield. If the company is expected to engage in little or no stock financing, or if stock financing is expected to occur only when the market value is close to book value, the expected rate of growth in the earnings, dividends, and price per share is g = br. As will be shown later, we may ignore stock financing and only consider growth due to retention of earnings.

If two conditions are satisfied, the best estimate of g is obtained either from the company's current values of b and r or from weighted averages of their recent values. These two conditions are: stock financing may be ignored for either of the reasons stated above, and there is no information other than the past values of b and r which can be used to forecast their future values.

The sharp rise in energy prices and other costs over the past decade have had a disruptive influence on the electric utility industry, and they have created situations in which there are obvious reasons why past values of b and r should not be projected into the future. In two recent cases, the DCF formula was adapted to deal with the peculiar circumstances of each case.[1] Similarly, as will be shown below, the recent dramatic change in anticipated inflation provides information which should be used to modify the past values of b and r in order to obtain a more accurate forecast of expected growth.

3. <u>Alternative Measures of Expected Growth</u>

It might be thought that past rates of growth in

---

1 Testimony of Myron J. Gordon, Boston Edison Company Case No. DPU 19300, Commonwealth of Massachusetts, Department of Public Utilities, 1977; and Testimony of Myron J. Gordon, Public Service Company of New Mexico Case No. 1419, New Mexico Public Service Commission, 1979.

either earnings, dividends, or price could be used as estimates of g, the forecast rate of future growth in dividends. However, these past rates of growth are most unreliable due to extraneous influences on them, such as changes in the rate of return on common equity, changes in the retention rate, or changes in the yield required by investors in the case of price changes. The potential error in using past growth in <u>earnings</u> to estimate g is illustrated in Schedule 19, where the hypothetical company's return on common equity is 10% in the first three periods and 15% in the last three periods. With a retention rate of 40% and a return rate of 15% the growth rate is 6% in the last three years. This is a reasonable estimate of the expected future growth rate as of the end of the 6th year. However, with the 56% growth rate due to the rise in the return rate in the fourth year, a simple average of the five annual past growth rates in earnings is in excess of 15%. Clearly, this type of estimate of future growth rates cannot be used with any reliability at all, especially now when public utilities have received frequent upward adjustments in their allowed rates of return over the past five years. To do so would be to expect the company's rate of return on common equity to increase by 50% about every five years. This would be a ridiculous forecast, which the use of b and r would make readily apparent.

It can also be demonstrated that a change in the dividend payout rate makes the past rate of growth in dividends an incorrect basis for predicting g. Assume that a company has been earning a rate of return on its common stock of r = .10, that it has been retaining the fraction b = .60 of its earnings, and that, as a consequence, its dividend has been growing at the rate br = (.60)(.10) = .06. If the company were to raise the fraction of earnings it pays in dividends so that b falls to .25, the rate of growth in the dividend would then fall to br = (.25) (.10) = .025. However, over the period that spans the rise in the dividend payout rate, the dividend would have grown at an even higher rate than the prior 6%. It would only be correct to project the past rate of growth in the dividend into the future on the highly implausible assumption that the company is expected periodically to raise its payout rate. Therefore, unless there is convincing evidence to the contrary, <u>current expectations of b and r provide the best basis for forecasting future growth.</u>

C. Cost of Equity Capital for AT&T

Under the method we have advocated for estimating future growth, the DCF formula for a company's cost of equity capital is:

$$k = \frac{D_1}{P_0} + br. \qquad (8)$$

To arrive at a company's current value of k, the current value of each of the quantities on the right-hand side of Equation (8) must be determined. This is done below for AT&T. As we will see, obtaining estimates of these values is extremely difficult in the turbulence of today's capital markets.

1. Dividend Yield

We argued above that the projected dividend yield is appropriate for setting the allowed rate of return on equity. The current quarterly dividend payable on April 1, 1980, is \$1.25. The Value Line forecast for dividends over the next 12 months has been reduced from \$5.20 in June, 1979, to a current forecast of \$5.00.[1] Value Line reduced its forecast dividend even though it was aware of AT&T's stated intent to maintain shareholders real dividend income against inflation.[2] For the last few years AT&T has followed a policy of raising its dividend in the first quarter. With the recent declaration of the dividend to be paid on April 1, 1980 maintained at \$1.25,

---

[1] Value Line, March 15, 1980.

[2] Value Line, February 1, 1980.

the Value Line estimate appears reasonable, and we will use a dividend of $5.00, equal to the annualized value of the current quarterly dividend of $1.25.[1]

We have also argued that we should use the share price on the date for which the estimate was made. Since this testimony was finalized on March 29, 1980, we will use the company's closing price on the previous day, that is, P_o = $48.50, which results in a dividend yield of $5.00/ $48.50 = 10.31%.

Ordinarily, for periods of up to a few months, the price of a public utility share only fluctuates in a narrow range, and the choice among the prevailing prices is usually of no particular significance. However, the impact of inflation during the second half of 1979 and the actions and statements of the Federal Reserve Board and other government officials (beginning in October and culminating in President Carter's recent anti-inflation program) have had a striking impact on the capital markets. Short-term interest rates have risen sharply, and the yields and prices on long-term securities have fluctuated dramatically. In particular, as can be seen in Schedule 20, AT&T's stock fell from $57.83 on June 30, 1979, to $55 on September 30, 1979. Since then it has decreased steadily to a low of $45 on March 7, 1980, before rising to the current price of $48.50 on March 28, 1980. During the same period its dividend

[1] Projection of a higher dividend in the current economic environment would require a downward revision in the growth rate forecasts below.

yield rose steadily from 8.99% on June 30, 1979, to the current projected yield of 10.31%. This was due mainly to the effects of its dropping share price, but also to the reduction in its projected dividend from $5.20 to $5.00.

Through their impact on the dividend yield, the date and the share price used to arrive at AT&T's cost of equity capital have a material impact on the value obtained for k. In other words, in a period over which interest rates fluctuate widely, share prices and the cost of equity capital also fluctuate widely. At the time this testimony was prepared, the reaction to President Carter's anti-inflation program was unknown. Although our estimated dividend yield of 10.31% represents our best estimate at this time, the unfolding reaction to the President's program may cause AT&T's dividend yield to vary considerably over the next few months.

2. <u>Growth Rate - Past Financial Data</u>

In order to arrive at AT&T's growth rate, we require the retention rate, b, and the rate of return on common equity, r, that investors may reasonably expect.

As a first step, let us estimate b and r using only historical data. Schedule 21 shows the underlying data for the years 1975 to 1979 that is needed to calculate b and r.

For the rate of return on common equity that investors expect, we first not that a simple average of the five values of r_t (row 5) from 1975 to 1979 is 11.81%. However, inspection of the annual values reveal that although r was abnormally depressed in 1975, its values for the next three years exhibited a definite upward trend, and then only declined slightly in 1979. Investors now might well believe that the material rise in the cost of capital between 1975 and 1979 justifies the rates of return the company realized in the more recent years, in which case they would rely primarily on the 1978 and 1979 figures in forecasting the company's future rate of return. A simple average of these figures is 13.05% and it seems reasonable that investors might conclude that 13% represents the best estimate of the long-term return AT&T is expected to earn on common equity.

For the retention rate that investors expect, we first note that a simple average of the five values of b_t (row 9) from 1975 to 1979 is 37.23%. However, this average is affected by the low retention rate in 1975, and in recent years, 1977-1979, the retention rate has averaged 38.93%. It seems reasonable that on the basis of this data, investors might use these recent years, and arrive at 39% as the best estimate of AT&T's retention ratio.

Combining the above values (obtained by using historical values in Equation (8) for P_0, D_1, b, and r) provides an estimate of AT&T's cost of equity capital as of March 28, 1980, of:

$$k = \frac{D_1}{P_o} + br$$

$$= \frac{\$\,5.00}{\$48.50} + (.39)(.13)$$

$$= .1031 + .0507 = 15.38\% .$$

However, before accepting this result it may be instructive to pose the following question: What would have been the estimate for k as of June 30, 1979?

3. Growth Rate - Recent Developments

On June 30, 1979, Value Line estimated that AT&T's 1979 earnings would be \$8.00 per share. The actual value of earnings per share for 1979 was \$8.04. Since we would have been reluctant to estimate k at that time without 1979 data, we would have relied on the Value Line forecast to complete the 1979 annual data, a procedure we have used in the past. Since the Value Line estimates were extremely close to the actual 1979 results, using these estimates and the historical data would have produced the same estimates of b and r obtained previously. It is obvious that if the data and analysis do not change materially, we would obtain the same measurement of the growth rate at any point between June 30, 1979, and March 28, 1980.

The estimates which would have been obtained on two previous dates are provided below:

Date	D_1/P_0 +	br =	k
June 30, 1979	8.99%	5.07%	14.06%
November 19, 1979	9.39%	5.07%	14.46%

An estimate is provided for November 19, 1979, for comparative purposes, since an estimate of k was obtained for Rochester Telephone Co. on that date of 14.85%.[1] The difference in k between Rochester Telephone and AT&T may be attributed to AT&T's slightly lower business risk due to its greater diversification.

The problem can now be easily seen. The estimate of 15.38% obtained for AT&T is correct only if we assume that the large increase in the expected rate of inflation (which raised the dividend yield on AT&T from 8.99% on June 30, 1979, to 10.31% on March 28, 1980) had no effect on the anticipated growth in the dividend.

It is extremely unlikely that investors believe that to be true. The rise in the expected rate of inflation has not only increased interest rates, but also the expected rate at which AT&T's other costs of production, such as materials and labor, will grow. A continued expectation that the company will earn a return on common of 13% and retain 39% of earnings would require the belief that the rate of growth in its revenues will rise to match

---

[1] Myron J. Gordon, Direct Testimony, Before the State of New York Public Utility Commission, In the Matter of Rochester Telephone Co., November 20, 1979.

the rise in the rate of growth of its costs. However, if investors fear that the regulatory process will not be fully responsive to the increase in the rate at which the company's costs are rising, they will revise their growth estimate downward. That is, with any regulatory lag in the pass through of higher costs, a rise in the expected inflation rate would reduce investor estimates of long-run return on common equity, and would, therefore, result in a downward revision of expected growth. In that event, simply raising the estimate of AT&T's cost of equity capital by the increase in the dividend yield would result in an overstatement of the required return.

It is our judgment that the response of investors to the rise in the expected rate of inflation has been a downward revision in expectations regarding AT&T's rate of return on common equity, implying a downward revision in its retention rate also. In support of this position, we note that Value Line lowered its prediction of 1980 earnings per share for AT&T to $7.50, and lowered its predicted 1980 dividend per share to $5.00.[1] This implies for 1980 an estimate for r of 11.60% and an estimate for b of 33.33%.

Under the present turbulent economic conditions it is extremely difficult to estimate with precision the extent

---

[1] Value Line, February 1, 1980.

to which these rates have been revised downward. If the revised figures are a 12.50% return on common equity and a 37% retention rate, then the estimated growth rate must be reduced from 5.07% to 4.63%.[1] Adding the latter figure to the current dividend yield of 10.31% results in a cost of equity capital of 14.94%. On the other hand, the rise in interest rates over the past six months may be taken as evidence that the cost of equity capital has gone up over the same time period. Hence, in some measure, this rise in intere[illegible] will lead to an upward revision in the [illegible] of return allowed by the numerous regulatory commissions that [illegible] rates for AT&T. A generous allowance for [illegible] favorable impact of increases in the allowed rates of return on inves[illegible] forecasts of the AT&T growth rate [illegible] a rise in its value from the above 4.63% to 5.25%. This latter growth rate [illegible] combined with the 10.31% dividend yield re[illegible] equity capital of 15.56%. In our judgment the AT&T cost of equity capital may well be as low as 15.0%, [illegible] unlikely to be above 15.5%, and 15.25% represents our base estimate as of March 28, 1980.

---

[1] Using this reasoning, the growth rate was adjusted downward by 69 basis points for Rochester Telephone. Ibid., Supplemental Prepared Direct Testimony, March 24, 1980.

RFS Advance Access published March 17, 2007

A Comprehensive Look at The Empirical Performance of Equity Premium Prediction*

Amit Goyal
Emory University
Goizueta Business School†

Ivo Welch
Brown University
Department of Economics
NBER‡

January 3, 2007

Abstract

Our paper comprehensively reexamines the performance of variables that have been suggested by the academic literature to be good predictors of the equity premium. We find that by and large, these models have predicted poorly both in-sample and out-of-sample for thirty years now; these models seem unstable, as diagnosed by their out-of-sample predictions and other statistics; and these models would not have helped an investor with access only to available information to profitably time the market.

*Thanks to Malcolm Baker, Ray Ball, John Campbell, John Cochrane, Francis Diebold, Ravi Jagannathan, Owen Lamont, Sydney Ludvigson, Rajnish Mehra, Michael Roberts, Jay Shanken, Samuel Thompson, Jeff Wurgler, and Yihong Xia for comments, and Todd Clark for providing us with some critical McCracken values. We especially appreciate John Campbell and Sam Thompson for challenging our earlier drafts, and iterating mutually over working papers with opposite perspectives.

†http://www.goizueta.emory.edu/agoyal. E-mail: amit_goyal@bus.emory.edu.
‡http://welch.econ.brown.edu. E-mail: ivo_welch@brown.edu.

JEL Classification: G12, G14.

1 Introduction

Attempts to predict stock market returns or the equity premium have a long tradition in finance. As early as 1920, Dow (1920) explored the role of dividend ratios. A typical specification regresses an independent lagged predictor on the stock market rate of return or, as we shall do, on the equity premium,

$$\text{Equity Premium}(t) \;=\; \gamma_0 \;+\; \gamma_1 \cdot x(t-1) + \epsilon(t) \quad . \tag{1}$$

γ_1 is interpreted as a measure of how significant x is in predicting the equity premium. The most prominent x variables explored in the literature are the dividend price ratio and dividend yield, the earnings price ratio and dividend-earnings (payout) ratio, various interest rates and spreads, the inflation rates, the book-to-market ratio, volatility, the investment-capital ratio, the consumption, wealth, and income ratio (CAY), and aggregate net or equity issuing activity.

The literature is difficult to absorb. Different papers use different techniques, variables, and time periods. Results from papers that were written years ago may change when more recent data is used. Some papers contradict the findings of others. Still, most readers are left with the impression that "prediction works"—though it is unclear exactly what works. The prevailing tone in the literature is perhaps best summarized by Lettau and Ludvigson (2001, p.842)

> It is now widely accepted that excess returns are predictable by variables such as dividend-price ratios, earnings-price ratios, dividend-earnings ratios, and an assortment of other financial indicators.

There are also a healthy number of current papers which further cement this perspective; and a large theoretical and normative literature has developed that stipulates how investors should allocate their wealth as a function of the aforementioned variables.

The goal of our own paper is to comprehensively reexamine the empirical evidence as of early 2006, evaluating each variable using the same methods (mostly, but not only, in linear models), time-periods, and estimation frequencies. The evidence suggests that most models are unstable or even spurious. Most models are no longer significant even in-sample (IS), and the few models that still are usually fail simple regression diagnostics. Most models have performed poorly for over thirty years IS. For many models, any earlier apparent statistical significance was often based exclusively on years up to *and especially on* the years of the Oil Shock of 1973-5. Most models have poor out-of-sample (OOS) performance, but not in a way that merely suggests lower power than IS tests. They predict poorly late in the sample, not early in the sample. (For many variables, we have difficulty finding robust statistical significance even when they are examined only during their most favorable contiguous OOS sub-period.) Finally, the OOS performance is not only a useful model diagnostic for the IS regressions, but also interesting in itself for an investor who had sought to use these models for market-timing. Our evidence suggests that the models would not have helped such an investor.

Therefore, although it is possible to search for, to occasionally stumble upon and then to defend some seemingly statistically significant models, we interpret our results to suggest that a healthy skepticism is appropriate when it comes to predicting the equity premium, at least as of early 2006. The models seem not robust.

Our paper now proceeds as follows. We describe our data—available at the RFS website—in Section 2 and our tests in Section 3. Section 4 explores our base case—predicting equity premia annually using OLS forecasts. In Sections 5 and 6, we predict equity premia on five-year and monthly horizons, the latter with special emphasis on the suggestions in Campbell and Thompson (2005). Section 7 tries earnings and dividend ratios with longer memory as independent variables, corrections for persistence in regressors, and encompassing model forecasts. Section 8 reviews earlier literature. Section 9 concludes.

2 Data Sources and Data Construction

Our dependent variable is always the equity premium, i.e., the total rate of return on the stock market minus the prevailing short-term interest rate.

Stock Returns: We use S&P 500 index returns from 1926 to 2005 from CRSP's month-end values. Stock returns are the continuously compounded returns on the S&P 500 index, including dividends. For yearly and longer data frequencies, we can go back as far as 1871, using data from Robert Shiller's website. For monthly frequency, we can only begin in the CRSP period, i.e., 1927.

Risk-free Rate: The risk-free rate from 1920 to 2005 is the T-bill rate. Because there was no risk-free short-term debt prior to the 1920's, we had to estimate it. Commercial paper rates for New York City are from the NBER's Macrohistory data base. These are available from 1871 to 1970. We estimated a regression from 1920 to 1971, which yielded

$$\text{T-bill Rate} \;=\; -0.004 \;+\; 0.886 \cdot \text{Commercial Paper Rate}\ , \tag{2}$$

with an R^2 of 95.7%. Therefore, we instrumented the risk-free rate from 1871 to 1919 with the predicted regression equation. The correlation for the period 1920 to 1971 between the equity premium computed using the actual T-bill rate and that computed using the predicted T-bill rate (using the commercial paper rate) is 99.8%.

The equity premium had a mean (standard deviation) of 4.85% (17.79%) over the entire sample from 1872 to 2005; 6.04% (19.17%) from 1927 to 2005; and 4.03% (15.70%) from 1965 to 2005.

Our first set of independent variables are primarily stock characteristics:

Dividends: Dividends are twelve-month moving sums of dividends paid on the S&P 500 index. The data are from Robert Shiller's website from 1871 to 1970.

Dividends from 1971 to 2005 are from the S&P Corporation. The **Dividend Price Ratio** (**d/p**) is the difference between the log of dividends and the log of prices. The **Dividend Yield** (**d/y**) is the difference between the log of dividends and the log of *lagged* prices. (See, e.g., Ball (1978), Campbell (1987), Campbell and Shiller (1988a, 1988b), Campbell and Viceira (2002), Campbell and Yogo (2006), the survey in Cochrane (1997), Fama and French (1988), Hodrick (1992), Lewellen (2004), Menzly, Santos, and Veronesi (2004), Rozeff (1984), and Shiller (1984).)

Earnings: Earnings are twelve-month moving sums of earnings on the S&P 500 index. The data are again from Robert Shiller's website from 1871 to June 2003. Earnings from June 2003 to December 2005 are our own estimates based on interpolation of quarterly earnings provided by the S&P Corporation. The **Earnings Price Ratio** (**e/p**) is the difference between the log of earnings and the log of prices. (We also consider variations, in which we explore multi-year moving averages of numerator or denominator, e.g., as in $\mathbf{e}^{10}/\mathbf{p}$, which is the moving ten-year average of earnings divided by price.) The **Dividend Payout Ratio** (**d/e**) is the difference between the log of dividends and the log of earnings. (See, e.g., Campbell and Shiller (1988a, 1998) and Lamont (1998).)

Stock Variance (**svar**): Stock Variance is computed as sum of squared daily returns on the S&P 500. G. William Schwert provided daily returns from 1871 to 1926; data from 1926 to 2005 are from CRSP. (See Guo (2006).)

Cross-Sectional Premium (**csp**): The cross-sectional beta premium measures the relative valuations of high- and low-beta stocks and is proposed in Polk, Thompson, and Vuolteenaho (2006). The **csp** data are from Samuel Thompson from May 1937 to December 2002.

Book Value: Book values from 1920 to 2005 are from Value Line's website, specifically their *Long-Term Perspective Chart* of the Dow Jones Industrial Average. The **Book to Market Ratio** (**b/m**) is the ratio of book value to market value for the Dow Jones Industrial Average. For the months from March to December, this is computed by dividing book value at the end of the previous year by the price at the end of the current month. For the months of January and February, this is computed by dividing book value at the end of two years ago by the price at the end of the current month. (See, e.g, Kothari and Shanken (1997) and Pontiff and Schall (1998).)

Corporate Issuing Activity: We entertain two measures of corporate issuing activity. **Net Equity Expansion** (**ntis**) is the ratio of twelve-month moving sums of net issues by NYSE listed stocks divided by the total end-of-year market capitalization of NYSE stocks. This dollar amount of net equity issuing activity (IPOs, SEOs, stock repurchases, less dividends) for NYSE listed stocks is computed from CRSP data as

$$\text{Net Issue}_t = \text{Mcap}_t - \text{Mcap}_{t-1} \cdot (1 + \text{vwretx}_t) , \tag{3}$$

where Mcap is the total market capitalization, and vwretx is the value weighted return (excluding dividends) on the NYSE index.[1] These data are available from 1926 to 2005. **ntis** is closely related, but not identical, to a variable proposed in Boudoukh, Michaely, Richardson, and Roberts (2005). The second measure, **Percent Equity Issuing** (**eqis**), is the ratio of equity issuing activity as a fraction of total issuing activity. This is the variable proposed in Baker and Wurgler (2000). The authors provided us with the data, except for 2005, which we added ourselves. The first equity issuing measure is relative to aggregate market cap, while the second is relative to aggregate corporate issuing.

Our next set of independent variables is interest-rate related:

Treasury Bills (**tbl**): T-bill rates from 1920 to 1933 are the *U.S. Yields On Short-Term United States Securities, Three-Six Month Treasury Notes and Certificates, Three Month Treasury* series in the NBER Macrohistory data base. T-bill rates from 1934 to 2005 are the *3-Month Treasury Bill: Secondary Market Rate* from the

[1] This calculation implicitly assumes that the delisting return is –100 percent. Using the actual delisting return, where available, or ignoring delistings altogether, has no impact on our results.

economic research data base at the Federal Reserve Bank at St. Louis (FRED). (See, e.g., Campbell (1987) and Hodrick (1992).)

Long Term Yield (lty): Our long-term government bond yield data from 1919 to 1925 is the *U.S. Yield On Long-Term United States Bonds* series in the NBER's Macrohistory data base. Yields from 1926 to 2005 are from Ibbotson's *Stocks, Bonds, Bills and Inflation Yearbook*, the same source that provided the **Long Term Rate of Returns (ltr)**. The **Term Spread (tms)** is the difference between the long term yield on government bonds and the T-bill. (See, e.g., Campbell (1987) and Fama and French (1989).)

Corporate Bond Returns: Long-term corporate bond returns from 1926 to 2005 are again from Ibbotson's *Stocks, Bonds, Bills and Inflation Yearbook*. **Corporate Bond Yields** on AAA and BAA-rated bonds from 1919 to 2005 are from FRED. The **Default Yield Spread (dfy)** is the difference between BAA and AAA-rated corporate bond *yields*. The **Default Return Spread (dfr)** is the difference between long-term corporate bond and long-term government bond *returns*. (See, e.g., Fama and French (1989) and Keim and Stambaugh (1986).)

Inflation (infl): Inflation is the *Consumer Price Index (All Urban Consumers)* from 1919 to 2005 from the Bureau of Labor Statistics. Because inflation information is released only in the following month, we wait for one month before using it in our monthly regressions. (See, e.g., Campbell and Vuolteenaho (2004), Fama (1981), Fama and Schwert (1977), and Lintner (1975).)

Like inflation, our next variable is also a common broad macroeconomic indicator.

Investment to Capital Ratio (i/k): The investment to capital ratio is the ratio of aggregate (private nonresidential fixed) investment to aggregate capital for the whole economy. This is the variable proposed in Cochrane (1991). John Cochrane kindly provided us with updated data.

Of course, many papers explore multiple variables. For example, Ang and Bekaert (2003) explore both interest rate and dividend related variables. In addition to simple univariate prediction models, we also entertain two methods that rely on multiple variables (**all** and **ms**), and two models that are rolling in their independent variable construction (**cay** and **ms**).

A "Kitchen Sink" Regression (all): This includes all the aforementioned variables. (It does not include **cay**, described below, partly due to limited data availability of **cay**.)

Consumption, wealth, income ratio (cay): Lettau and Ludvigson (2001) estimate the following equation:

$$c_t = \alpha + \beta_a \cdot a_t + \beta_y \cdot y_t + \sum_{i=-k}^{k} b_{a,i} \cdot \Delta a_{t-i} + \sum_{i=-k}^{k} b_{y,i} \cdot \Delta y_{t-i} + \epsilon_t, \quad t = k+1, \ldots, T-k, \tag{4}$$

where c is the aggregate consumption, a is the aggregate wealth, and y is the aggregate income. Using estimated coefficients from the above equation provides **cay** $\equiv \widehat{\text{cay}}_t = c_t - \hat{\beta}_a \cdot a_t - \hat{\beta}_y \cdot y_t, \quad t = 1, \ldots, T$. Note that, unlike the estimation equation, the fitting equation does not use look-ahead data. Eight leads/lags are used in quarterly estimation ($k = 8$) while two lags are used in annual estimation ($k = 2$). (For further details, see Lettau and Ludvigson (2001).) Data for **cay**'s construction are available from Martin Lettau's website at quarterly frequency from the second quarter of 1952 to the fourth quarter of 2005. Although annual data from 1948 to 2001 is also available from Martin Lettau's website, we reconstruct the data following their procedure as this allows us to expand the time-series from 1945 to 2005 (an addition of 7 observations).

Because the Lettau-Ludvigson measure of **cay** is constructed using look-ahead (in-sample) estimation regression coefficients, we also created an equivalent measure that excludes advance knowledge from the estimation equation and thus uses only prevailing data. In other words, if the current time period is 's', then we estimated equation (4) using only the data up to 's' through

$$c_t = \alpha + \beta_a^s \cdot a_t + \beta_y^s \cdot y_t + \sum_{i=-k}^{k} b_{a,i}^s \cdot \Delta a_{t-i} + \sum_{i=-k}^{k} b_{y,i}^s \cdot \Delta y_{t-i} + \epsilon_t, t = k+1, \ldots, s-k, \quad (5)$$

This measure is called **caya** ("ante") to distinguish it from the traditional variable **cayp** constructed with look-ahead bias ("post"). The superscript on the betas indicates that these are rolling estimates, i.e., a set of coefficients used in the construction of one **caya**$_S$ measure in one period.

A **model selection** approach, named "**ms**." If there are K variables, we consider 2^K models essentially consisting of all possible combinations of variables. (As with the kitchen sink model, **cay** is not a part of the **ms** selection.) Every period, we select one of these models that gives the minimum cumulative prediction errors up to time t. This method is based on Rissanen (1986) and is recommended by Bossaerts and Hillion (1999). Essentially, this method uses our criterion of minimum OOS prediction errors to choose amongst competing models *in each time period* t. This is also similar in spirit to the use of a more conventional criterion (like R^2) in Pesaran and Timmerman (1995) (who do not entertain our NULL hypothesis). This selection model also shares a certain flavor with our encompassing tests in Section 7, where we seek to find an optimal rolling combination between each model and an unconditional historical equity premium average, and with the Bayesian model selection approach in Avramov (2002).

The latter two models, **cay** and **ms**, are revised every period, which render in-sample regressions problematic. This is also why we did not include **caya** in the kitchen sink specification.

3 Empirical Procedure

Our base regression coefficients are estimated using OLS, although statistical significance is always computed from bootstrapped F-statistics (taking correlation of independent variables into account).

OOS statistics: The OOS forecast uses only the data available up to the time at which the forecast is made. Let e_{N} denote the vector of rolling OOS errors from the historical mean model and e_{A} denote the vector of rolling OOS errors from the OLS model. Our OOS statistics are computed as

$$\begin{aligned} R^2 &= 1 - \frac{\text{MSE}_{\mathsf{A}}}{\text{MSE}_{\mathsf{N}}} \quad , \quad \overline{R}^2 = R^2 - (1-R^2)\cdot\left(\frac{T-k}{T-1}\right) \quad , \\ \Delta\text{RMSE} &= \sqrt{\text{MSE}_{\mathsf{N}}} - \sqrt{\text{MSE}_{\mathsf{A}}} \quad , \\ \text{MSE-F} &= (T-h+1)\cdot\left(\frac{\text{MSE}_{\mathsf{N}} - \text{MSE}_{\mathsf{A}}}{\text{MSE}_{\mathsf{A}}}\right) \quad , \end{aligned} \tag{6}$$

where h is the degree of overlap ($h = 1$ for no overlap). MSE-F is McCracken's (2004) F-statistic. It tests for equal MSE of the unconditional forecast and the conditional forecast (i.e., ΔMSE = 0).[2] We generally do not report MSE-F statistics, but instead use their bootstrapped critical levels to provide statistical significance levels via stars in the tables.

For our encompassing tests in Section 7, we compute

$$\text{ENC} = \frac{T-h+1}{T} \cdot \frac{\sum_{t=1}^{T}\left(e_{\mathsf{N}t}^2 - e_{\mathsf{N}t}\cdot e_{\mathsf{A}t}\right)}{\text{MSE}_{\mathsf{A}}} \quad , \tag{7}$$

which is proposed by Clark and McCracken (2001). They also show that the MSE-F and ENC statistics follow non-standard distributions when testing nested models, because the asymptotic difference in squared forecast errors is exactly 0 with 0 variance under the NULL, rendering the standard distributions asymptotically invalid. Because our models are nested, we could use asymptotic critical values for MSE tests provided by McCracken, and asymptotic critical values for ENC tests provided by

[2]Our earlier drafts also entertained another performance metric, the mean absolute error difference ΔMAE. The results were similar. These drafts also described another OOS-statistic, MSE-T $= \sqrt{T+1-2\cdot h + h\cdot(h-1)/T} \cdot \left[\overline{d}/\widehat{se}\left(\overline{d}\right)\right]$, where $d_t = e_{\mathsf{N}t} - e_{\mathsf{A}t}$, and $\overline{d} = T^{-1}\cdot\sum_t^T d_t = \text{MSE}_{\mathsf{N}} - \text{MSE}_{\mathsf{A}}$ over the entire OOS period, and T is the total number of forecast observations. This is the Diebold and Mariano (1995) t-statistic modified by Harvey, Leybourne, and Newbold (1997). (We still use the latter as bounds in our plots, because we know the full distribution.) Again, the results were similar. We chose to use the MSE-F in this paper because Clark and McCracken (2001) find that MSE-F has higher power than MSE-T.

Clark and McCracken. However, because we use relatively small samples, because our independent variables are often highly serially correlated, and especially because we need critical values for our five-year *overlapping* observations (for which asymptotic critical values are not available), we obtain critical values from the bootstrap procedure described below. (The exceptions are that critical values for **caya**, **cayp**, and **all** models are not calculated using a bootstrap, and critical values for **ms** model are not calculated at all.) The NULL hypothesis is that the unconditional forecast is not inferior to the conditional forecast, so our critical values for OOS test are for a one-sided test (critical values of IS tests are, as usual, based on two-sided tests).[3]

Bootstrap: Our bootstrap follows Mark (1995) and Kilian (1999) and imposes the NULL of no predictability for calculating the critical values. In other words, the data generating process is assumed to be

$$\begin{aligned} y_{t+1} &= \alpha \qquad\qquad + u_{1t+1} \\ x_{t+1} &= \mu + \rho \cdot x_t + u_{2t+1} \quad . \end{aligned}$$

The bootstrap for calculating power assumes the data generating process is

$$\begin{aligned} y_{t+1} &= \alpha + \beta \cdot x_t + u_{1t+1} \\ x_{t+1} &= \mu + \rho \cdot x_t + u_{2t+1} \quad , \end{aligned}$$

where both β and ρ are estimated by OLS using the full sample of observations, with the residuals stored for sampling. We then generate 10,000 bootstrapped time series by drawing with replacement from the residuals. The initial observation—preceding the sample of data used to estimate the models—is selected by picking one date from the actual data at random. This bootstrap procedure not only preserves the autocorrelation structure of the predictor variable, thereby being valid under the

[3]If the regression coefficient β is small (so that explanatory power is low or the in-sample R^2 is low), it may happen that our unconditional model outperforms on OOS because of estimation error in the rolling estimates of β. In this case, ΔRMSE might be negative but still significant *because these tests are ultimately tests of whether β is equal to zero.*

Stambaugh (1999) specification, but also preserves the cross-correlation structure of the two residuals.[4]

Statistical Power: Our paper entertains both IS and OOS tests. Inoue and Kilian (2004) show that the OOS tests used in this paper are less powerful than IS tests, even though their size properties are roughly the same. Similar critiques of the OOS tests in our paper have been noted by Cochrane (2005) and Campbell and Thompson (2005). We believe this is the wrong way to look at the issue of power for two reasons:

1. It is true that under a well-specified stable underlying model, an IS OLS estimator is more efficient. Therefore, a researcher who has complete confidence in her underlying model specification (but not the underlying model parameters) should indeed rely on IS tests to establish significance—the alternative of OOS tests does have lower power. However, the point of any regression diagnostics, such as those for heteroskedasticity and autocorrelation, is always to subject otherwise seemingly successful regression models to a number of reasonable diagnostics when there is some model uncertainty. Relative to not running the diagnostic, by definition, any diagnostic that can reject the model at this stage sacrifices power *if* the specified underlying model is correct. In our forecasting regression context, OOS performance just happens to be one natural and especially useful diagnostic statistic. It can help determine whether a model is stable and well-specified, or changing over time, either suddenly or gradually.

   This also suggests why the simple power experiment performed in some of the aforementioned critiques of our own paper is wrong. It is unreasonable to propose a model if the IS performance is insignificant, regardless of its OOS performance. Reasonable (though not necessarily statistically significant) OOS performance is not a substitute, but a necessary complement for IS performance in order to establish the quality of the underlying model specification. The thought experiments and analyses in the critiques, which simply compare the power of

[4] We do not bootstrap for **cayp** because it is calculated using ex-post data; for **caya** and **ms** because these variables change each period; and for **all** because of computational burden.

OOS tests to that of IS tests, especially under their assumption of a correctly specified stable model, is therefore incorrect. The correct power experiment instead should explore whether *conditional on observed IS significance*, OOS diagnostics are reasonably powerful. We later show that they are.

Not reported in the tables, we also used the CUSUMQ test to test for model stability. Although this is a weak test, we can reject stability for all monthly models; and for all annual models except for **ntis**, **i/k**, and **cayp**, when we use data beginning in 1927. Thus, the CUSUMQ test sends the same message about the models as the findings that we shall report.

2. All of the OOS tests in our paper do not fail in the way the critics suggest. Low power OOS tests would produce relatively poor predictions early and relatively good predictions late in the sample. Instead, all of our models show the opposite behavior—good OOS performance early, bad OOS performance late.

   A simple alternative OOS estimator, which downweights early OOS predictions relative to late OOS predictions, would have more power than our unweighted OOS prediction test. Such a modified estimator would both be more powerful *and* it would show that all models explored in our paper perform even worse. (We do not use it only to keep it simple and to avoid a "cherry-picking-the-test" critique.)

Estimation Period: It is not clear how to choose the periods over which a regression model is estimated and subsequently evaluated. This is even more important for OOS tests. Although any choice is necessarily ad-hoc in the end, the criteria are clear. It is important to have enough initial data to get a reliable regression estimate at the start of evaluation period, and it is important to have an evaluation period that is long enough to be representative. We explore three time period specifications: the first begins OOS forecasts twenty years after data are available; the second begins OOS forecast in 1965 (or twenty years after data are available, whichever comes later); the third ignores all data prior to 1927 even in the estimation.[5] If a

[5]We also tried estimating our models only with data after World-War II, as recommended by

variable does not have complete data, some of these time-specifications can overlap. Using three different periods reflects different tradeoffs between the desire to obtain statistical power and the desire to obtain results that remain relevant today. In our graphical analysis later, we also evaluate the rolling predictive performance of variables. This analysis helps us identify periods of superior or inferior performance and can be seen as invariance to the choice of the OOS evaluation period (though not the estimation period).

4 Annual Prediction

Table 1: Annual Performance

Figure 1

Figure 2

Table 1 shows the predictive performance of the forecasting models on annual forecasting horizons. Figures 1 and 2 graph the IS and OOS performance of variables in Table 1. For the IS regressions, the performance is the cumulative squared demeaned equity premium minus the cumulative squared regression residual. For the OOS regressions, this is the cumulative squared prediction errors of the prevailing mean minus the cumulative squared prediction error of the predictive variable from the linear historical regression. Whenever a line increases, the ALTERNATIVE predicted better; whenever it decreases, the NULL predicted better. The units in the graphs are not intuitive, but the time-series pattern allows diagnosis of years with good or bad performance. Indeed, the final ΔSSE statistic in the OOS plot is sign-identical with the ΔRMSE statistic in our tables. The standard error of all the observations in the graphs is based on translating MSE-T statistic into symmetric 95% confidence intervals based on the McCracken (2004) critical values; the tables differ in using the MSE-F statistic instead.

The reader can easily adjust perspective to see how variations in starting or ending date would impact the conclusion—by shifting the graph up or down (redrawing the y=0 horizontal zero line). Indeed, a horizontal line and the right-side scale

Lewellen (2004). Some properties in some models change, especially when it comes to statistical significance and the importance of the Oil Shock for one variable, **d/p**. However, the overall conclusions of our paper remain.

indicate the equivalent zero-point for the second time period specification, in which we begin forecasts in 1965 (this is marked "Spec B Zero Val" line). The plots have also vertically shifted the IS errors, so that the IS line begins at zero on the date of our first OOS prediction. The Oil Shock recession of 1973 to 1975, as identified by the NBER, is marked by a vertical (red) bar in the figures.[6]

In addition to the figures and tables, we also summarize models' performances in small in-text summary tables, which give the IS-$\overline{R}^2$ and OOS-$\overline{R}^2$ for two time periods: the most recent 30 years and the entire sample period. The $\overline{R}^2$ for the subperiod is not the $\overline{R}^2$ for a different model estimated only over the most recent three decades, but the residual fit for the overall model over the subset of data points (e.g., computed simply as 1-SSE/SST for the last 360 residuals). The most recent three decades after the Oil Shock can help shed light on whether a model is likely to still perform well nowadays. Generally, it is easiest to understand the data by looking first at the figures, then at the in-text table, and finally at the full table.

A well-specified signal would inspire confidence in a potential investor if it had

1. both significant IS and reasonably good OOS performance over the entire sample period;
2. a generally upward drift (of course, an irregular one);
3. an upward drift which occurs not just in one short or unusual sample period—say just the two years around the Oil Shock;
4. an upward drift that remains positive over the most recent several decades—otherwise, even a reader taking the long view would have to be concerned with the possibility that the underlying model has drifted.

There are also other diagnostics that stable models should pass (heteroskedasticity, residual autocorrelation, etc.), but we do not explore them in our paper.

[6]The actual recession period was from November 1973 to March 1975. We treat both 1973 and 1975 as years of Oil Shock recession in annual prediction.

4.1 In-Sample Insignificant Models

As already mentioned, if a model has no IS performance, its OOS performance is not interesting. However, because some of the IS insignificant models are so prominent, and because it helps to understand why they may have been considered successful forecasters in past papers, we still provide some basic statistics and graph their OOS performance. The most prominent such models are the following:

Dividend Price Ratio: Figure 1 shows that there were four distinct periods for the **d/p** model, and this applies both to IS and OOS performance. **d/p** had mild underperformance from 1905 to WW-II, good performance from WW-II to 1975, neither good nor bad performance until the mid-1990s, and poor performance thereafter. The best sample period for **d/p** was from the mid 1930s to the mid 1980s. For the OOS, it was 1937 to 1984, although over half of the OOS performance was due to the Oil Shock. Moreover, the plot shows that the OOS performance of the **d/p** regression was consistently worse than the performance of its IS counterpart. The distance between the IS and OOS performance increased steadily until the Oil Shock.

Over the most recent 30 years (1976 to 2005), **d/p**'s performance is negative both IS and OOS. Over the entire period, **d/p** underperformed the prevailing mean OOS, too:

d/p	Recent 30 Years	All Years
IS $\overline{R}^2$	−4.80%	0.49%
OOS $\overline{R}^2$	−15.14%	−2.06%

Dividend Yield: Figure 1 shows that the **d/y** model's IS patterns look broadly like those of **d/p**. However, its OOS pattern was much more volatile: **d/y** predicted equity premia well during the Great Depression (1930 to 1933), the period from World War II to 1958, the Oil Shock of 1973-1975, and the market decline of 2000-2002. It had large prediction errors from 1958 to 1965 and from 1995 to 2000, and it had unremarkable performance in other years. The best OOS sample period started around 1925 and ended either in 1957 or 1975. The Oil Shock did not play an important role for **d/y**. Over the most recent 30 years, **d/y**'s performance is again negative IS and OOS. The full-sample OOS performance is also again negative:

d/y	Recent <u>30 Years</u>	All <u>Years</u>
IS $\overline{R}^2$	−5.52%	0.91%
OOS $\overline{R}^2$	−20.79%	−1.93%

Earnings Price Ratio: Figure 1 shows that **e/p** had inferior performance until WW-II, and superior performance from WW-II to the late 1970s. After the Oil Shock, it had generally non-descript performance (with the exception of the late 1990s and early 2000s). Its best sample period was 1943 to 2002. 2003 and 2004 were bad years for this model. Over the most recent 30 years, **e/p**'s performance is again negative IS and OOS. The full-sample OOS performance is negative too.

e/p	Recent <u>30 Years</u>	All <u>Years</u>
IS $\overline{R}^2$	−2.08%	1.08%
OOS $\overline{R}^2$	−5.98%	−1.78%

Table 1 shows that these three price ratios are not statistically significant IS at the 90% level. However, some disagreement in the literature can be explained by differences in the estimation period.[7]

[7]For example, the final lines in Table 1 show that **d/y** and **e/p** had positive and statistically significant IS performance at the 90% level if all data prior to 1927 is ignored. Nevertheless, Table 1 also shows that the OOS-$\overline{R}^2$ performance remains negative for both of these. Moreover, when the data begins in 1927 and the forecast begins in 1947 (another popular period choice), we find

(Data Begins in 1927)	**e/p**		**d/y**	
(Forecast Begins in 1947)	Recent <u>30 Years</u>	All <u>Years</u>	Recent <u>30 Years</u>	All <u>Years</u>
IS $\overline{R}^2$	−3.83%	3.20%	−5.20%	2.71%
OOS $\overline{R}^2$	−13.58%	3.41%	−28.05%	−16.65%

Finally, and again not reported in the table, another choice of estimation period can also make a difference. The three price models lost statistical significance over the full sample only in the 1990s. This is not because the IS-ΔRMSE has decreased further in the 1990's, but because the 1991–2005 prediction errors were more volatile, which raised the standard errors of point estimates.

Other Variables: The remaining plots in Figure 1 and the remaining IS insignificant models in Table 1 show that **d/e**, **dfy**, and **infl** essentially never had significantly positive OOS periods, and that **svar** had a huge drop in OOS performance from 1930 to 1933. Other variables (that are IS insignificant) often had good sample performance early on, ending somewhere between the Oil Shock and the mid-1980s, followed by poor performance over the most recent three decades. The plots also show that it was generally not just the late 1990s that invalidated them, unlike the case with the aforementioned price ratio models.

In sum, twelve models had insignificant in-sample full-period performance and, not surprisingly, these models generally did not offer good OOS performance.

4.2 In-Sample Significant Models

Five models were significant IS (**b/m**, **i/k**, **ntis**, **eqis**, and **all**) at least at the 10% two-sided level. Table 1 contains more details for these variables, such as the IS performance during the OOS period, and a power statistic. Together with the plots in Figure 2, this information helps the reader to judge the stability of the models—whether poor OOS performance is driven by less accurately estimated parameters (pointing to lower power), and/or by the fact that the model fails IS and/or OOS during the OOS sample period (pointing to a spurious model).

Book-market ratio: **b/m** is statistically significant at the 6% level IS. Figure 2 shows that it had excellent IS and OOS predictive performance right until the Oil Shock. Both its IS and OOS performance were poor from 1975 to 2000, and the recovery in 2000-2002 was not enough to gain back the 1997-2000 performance. Thus, the **b/m** model has negative performance over the most recent three decades, both IS and OOS.

b/m	Recent 30 Years	All Years
IS $\overline{R}^2$	−12.37%	3.20%
OOS $\overline{R}^2$	−29.31%	−1.72%

Over the entire sample period, the OOS performance is negative, too. The "IS for OOS" $\overline{R}^2$ in Table 1 shows how dependent **b/m**'s performance is on the first 20 years of the sample. The IS $\overline{R}^2$ is -7.29% for the 1965-2005 period. The comparable OOS $\overline{R}^2$ even reaches -12.71%.

As with other models, **b/m**'s lack of OOS significance is not just a matter of low test power. Table 1 shows that in the OOS prediction beginning in 1941, under the simulation of a stable model, the OOS statistic came out *statistically significantly* positive in 67%[8] of our (stable-model) simulations in which the IS regression was significant. Not reported in the table, positive performance (significant or insignificant) occurred in 78% of our simulations. A performance as negative as the observed ΔRMSE of -0.01 occurred in *none* of the simulations.

Investment-capital ratio: **i/k** is statistically significant IS at the 5% level. Figure 2 shows that, like **b/m**, it performed well only in the first half of its sample, both IS and OOS. About half of its performance, both IS and OOS, occurs during the Oil Shock. Over the most recent 30 years, **i/k** has underperformed:

i/k	Recent 30 Years	All Years
IS $\overline{R}^2$	−8.09%	6.63%
OOS $\overline{R}^2$	−18.02%	−1.77%

Corporate Issuing Activity: Recall that **ntis** measures equity issuing and repurchasing (plus dividends) relative to the price level; **eqis** measures equity issuing relative to debt issuing. Figure 2 shows that both variables had superior IS performance in the early 1930's, a part of the sample that is not part of the OOS period. **eqis** continues good performance into the late 1930's but gives back the extra gains immediately thereafter. In the OOS period, there is one stark difference between the two variables: **eqis** had superior performance during the Oil Shock, both IS and

[8]The 42% applies to draws that were not statistically significant in-sample at the 90% level. It is the equivalent of the experiment conducted in some other papers. However, because OOS performance is relevant only when the IS performance is significant, this is the wrong measure of power.

OOS. It is this performance that makes **eqis** the only variable that had statistically significant OOS performance in the annual data. In other periods, neither variable had superior performance during the OOS period.

Both variables underperformed over the most recent 30 years

	ntis		**eqis**	
	Recent 30 Years	All Years	Recent 30 Years	All Years
IS $\overline{R}^2$	−5.14%	8.15%	−10.36%	9.15%
OOS $\overline{R}^2$	−8.63%	−5.07%	−15.33%	2.04%

The plot can also help explain dueling perspectives about **eqis** between Butler, Grullon, and Weston (2005) and Baker, Taliaferro, and Wurgler (2004). One part of their disagreement is whether **eqis**'s performance is just random underperformance in sampled observations. Of course, some good years are expected to occur in any regression. Yet **eqis**'s superior performance may not have been so random, because it [a] occurred in consecutive years, and [b] in response to the Oil Shock events that are often considered to have been exogenous, unforecastable, and unusual. Butler, Grullon, and Weston also end their data in 2002, while Baker, Taliaferro, and Wurgler refer to our earlier draft and to Rapach and Wohar (2006), which end in 2003 and 1999, respectively. Our figure shows that small variations in the final year choice can make a difference in whether **eqis** turns out significant or not. In any case, both papers have good points. We agree with Butler, Grullon, and Weston that **eqis** would not have been a profitable and reliable predictor for an external investor, especially over the most recent 30 years. But we also agree with Baker, Taliaferro, and Wurgler that conceptually, it is not the OOS performance, but the IS performance that matters in the sense in which Baker and Wurgler (2000) were proposing **eqis**—not as a third-party predictor, but as documentary evidence of the fund-raising behavior of corporations. Corporations did repurchase profitably in the Great Depression and the Oil Shock era (though not in the "bubble period" collapse of 2001-2002).

all The final model with IS significance is the kitchen sink regression. It had high IS significance, but exceptionally poor OOS performance.

4.3 Time-Changing Models

caya and **ms** have no in-sample analogs, because the models themselves are constantly changing.

Consumption-Wealth-Income: Lettau and Ludvigson (2001) construct their **cay** proxy assuming that agents have some ex-post information. The experiment their study calls OOS is unusual: their representative agent still retains knowledge of the model's full-sample CAY-*construction* coefficients. It is OOS only in that the agent does not have knowledge of the *predictive* coefficient and thus has to update it on a running basis. We call the Lettau and Ludvigson (2001) variable **cayp**. We also construct **caya**, which represents a more genuine OOS experiment, in which investors are not assumed to have advance knowledge of the **cay** construction estimation coefficients.

Figure 2 shows that **cayp** had superior performance until the Oil Shock, and nondescript performance thereafter. It also benefited greatly from its performance during the Oil Shock itself.

cay	Recent 30 Years	All Years
some ex-post knowledge, **cayp** IS $\overline{R}^2$	10.52%	15.72%
some ex-post knowledge, **cayp** OOS $\overline{R}^2$	7.60%	16.78%
no advance knowledge, **caya** OOS $\overline{R}^2$	–12.39%	–4.33%

The full-sample **cayp** result confirms the findings in Lettau and Ludvigson (2001). **cayp** outperforms the benchmark OOS RMSE by 1.61% per annum. It is stable and its OOS performance is almost identical to its IS performance. In contrast to **cayp**, **caya** has had no superior OOS performance, either over the entire sample period or the most recent years. In fact, without advance knowledge, **caya** had the worst OOS $\overline{R}^2$ performance among our single variable models.

Model Selection Finally, **ms** fails with a pattern similar to earlier variables—good performance until 1976, bad performance thereafter.

ms	Recent 30 Years	All Years
IS $\overline{R}^2$	–	–
OOS $\overline{R}^2$	–43.40%	–22.50%

Conclusion: There were a number of periods with sharp stock market changes, such as the Great Depression of 1929–1933 (in which the S&P500 dropped from 24.35 at the end of 1928 to 6.89 at the end of 1932) and the "bubble period" from 1999–2001 (with its subsequent collapse). However, it is the Oil Shock recession of 1973–1975, in which the S&P500 dropped from 108.29 in October 1973 to 63.54 in September 1974—and its recovery back to 95.19 in June 1975—that stands out. Many models depend on it for their apparent forecasting ability, often both IS and OOS. (And none performs well thereafter.) Still, we caution against overreading or underreading this evidence. In favor of discounting this period, the observed source of significance seems unusual, because the important years are consecutive observations during an unusual period. (They do not appear to be merely independent draws.) In favor of not discounting this period, we do not know how one would identify these special multi-year periods ahead of time, except through a model. Thus, good prediction during such a large shock should not be automatically discounted. More importantly and less ambiguously, no model seems to have performed well since—that is, over the last thirty years.

In sum, on an annual prediction basis, there is no single variable that meets all of our four suggested investment criteria from Page 14 (IS significance, OOS performance, reliance not just on some outliers, and good positive performance over the last three decades.) Most models fail on all four criteria.

5 Five-Yearly Prediction

Table 2: Five-Yearly Frequency

Some models may predict long-term returns better than short-term returns. Unfortunately, we do not have many years to explore 5-year predictions thoroughly, and there are difficult econometric issues arising from data overlap. Therefore, we only briefly describe some preliminary and perhaps naive findings. (See, e.g., Boudoukh, Richardson, and Whitelaw (2005) and Lamoureux and Zhou (1996) for more detailed treatments.) Table 2 repeats Table 1 with 5-year returns. As before, we bootstrap all critical significance levels. This is especially important here, because the observations are overlapping and the asymptotic critical values are not available.

Table 2 shows that there are four models that are significant IS over the entire sample period: **ntis**, **d/p**, **i/k**, and **all**. **ntis** and **i/k** were also significant in the annual data (Table 1). Two more variables, **d/y** and **tms**, are IS significant if no data prior to 1927 is used.

Dividend Price Ratio: **d/p** had negative performance OOS regardless of period.

Term Spread: **tms** is significant IS only if the data begins in 1927 rather than 1921. An unreported plot shows that **tms** performed well from 1968–1979, poorly from 1979–1986, and then well again from 1986–2005. Indeed, its better years occur in the OOS period, with an IS $\overline{R}^2$ of 23.54% from 1965-2005. This was sufficient to permit it to turn in a superior OOS ΔRMSE performance of 2.77% per five-years—a meaningful difference. On the negative side, **tms** has positive OOS performance *only* if forecasting begins in 1965. Using 1927–2005 data and starting forecasts in 1947, the OOS ΔRMSE and $\overline{R}^2$ are negative.

The Kitchen Sink: **all** again turned in exceptionally poor OOS performance.

Model selection (**ms**) and **caya** again have no in-sample analogs. **ms** had the worst predictive performance observed in this paper. **caya** had good OOS performance of 2.50% per five-year period. Similarly, the investment-capital ratio, **i/k**, had both positive IS and OOS performance, and both over the most recent three decades as well as over the full sample (where it was also statistically significant).

i/k	Recent 30 Years	All Years
IS $\overline{R}^2$	30.60%	33.99%
OOS $\overline{R}^2$	28.00%	12.99%

i/k's performance is driven by its ability to predict the 2000 crash. In 1997, it had already turned negative on its 1998-2002 equity premium prediction, thus predicting the 2000 collapse, while the unconditional benchmark prediction continued with its 30% plus predictions:

Forecast made in	for years	Actual EqPm	Forecast Unc.	**i/k**	Forecast made in	for years	Actual EqPm	Forecast Unc.	**i/k**
1995	1996-2000	0.58	0.30	0.22	1998	1999-2003	−0.19	0.33	−0.09
1996	1997-2001	0.27	0.31	0.09	1999	2000-2004	−0.25	0.34	−0.07
1997	1998-2002	−0.23	0.31	−0.01	2000	2001-2005	−0.08	0.34	−0.06

This model (and perhaps **caya**) seem promising. We hesitate to endorse them further only because our inference is based on a small number of observations, and because statistical significance with overlapping multi-year returns raises a set of issues that we can only tangentially address. We hope more data will allow researchers to explore these models in more detail.

6 Monthly Prediction and Campbell-Thompson

Table 3 describes the performance of models predicting monthly equity premia. It also addresses a number of points brought up by Campbell and Thompson (2005), henceforth CT. We do not have dividend data prior to 1927, and thus no reliable equity premium data before then. This is why even our the estimation period begins only in 1927.

Table 3: Monthly and Campbell and Thompson Analysis

6.1 In-Sample Performance

Table 3 presents the performance of monthly predictions both IS and OOS. The first data column shows the IS performance when the predicted variable is logged (as in the rest of the paper). Eight out of eighteen models are in-sample significant at the 90% level, seven at the 95% level. Because CT use simple rather than log equity premia, the remaining data columns follow their convention. This generally improves the predictive power of most models, and the fourth column (by which rows are sorted) shows that three more models turn in statistically significant IS.[9]

CT argue that a reasonable investor would not have used a models to forecast a negative equity premium. Therefore, they suggest truncation of such predictions at zero. In a sense, this injects caution into the models themselves, a point we agree with. Because there were high equity premium realizations especially in the 1980s and 1990s, a time when many models were bearish, this constraint can improve performance. Of course, it also transforms formerly linear models into non-linear models, which are generally not the subject of our paper. CT do *not* truncate predictions in their in-sample regressions, but there is no reason not to do so. Therefore, the fifth column shows a revised IS $\overline{R}^2$ statistic. Some models now perform better, some perform worse.

6.2 Out-of-Sample Prediction Performance

The remaining columns explore the OOS performance. The sixth column shows that without further manipulation, **eqis** is the only model with both superior IS ($\overline{R}^2$=0.82% and 0.80%) and OOS ($\overline{R}^2$ = 0.14%) untruncated performance. The term-spread, **tms**, has OOS performance that is even better ($\overline{R}^2$ = 0.22%), but it just misses statistical

[9] Geert Bekaert pointed out to us that if returns are truly log-normal, part of their increased explanatory power could be due to the ability of these variables to forecast volatility.

significance IS at the 90% level. **infl** has marginally good OOS performance, but poor IS performance. All other models have negative IS or OOS untruncated $\overline{R}^2$.

The remaining columns show model performance when we implement the Campbell and Thompson (2005) suggestions. The seventh column describes the frequency of truncation of negative equity premium predictions. For example, **d/y**'s equity premium predictions are truncated to zero in 54.2% of all months; **csp**'s predictions are truncated in 44.7% of all months. Truncation is a very effective constraint.

CT also suggest using the unconditional model if the theory offers one coefficient sign and the estimation comes up with the opposite sign. For some variables, such as the dividend ratios, this is easy. For other models, it is not clear what the appropriate sign of the coefficient would be. In any case, this matters little in our data set. The eighth column shows that the coefficient sign constraint matters only for **dfr**, and **ltr** (and mildly for **d/e**). None of these three models has IS performance high enough to make this worthwhile to explore further.

The ninth and tenth columns, $\overline{R}^2{}_{\text{TU}}$ and $\Delta\text{RMSE}_{\text{TU}}$, show the effect of the CT truncations on OOS prediction. For many models, the performance improves. Nevertheless, the OOS $\overline{R}^2$'s remain generally much lower than their IS equivalents. Some models have positive ΔRMSE but negative OOS $\overline{R}^2$. This reflects the number of degrees of freedom: even though we have between 400 and 800 data months, the plain ΔRMSE and R^2 are often so small that the $\overline{R}^2$ turns negative. For example, even with over 400 months of data, the loss of three degrees of freedom is enough for **cay3** to render a positive ΔRMSE of 0.0088 (equivalent to an unreported unadjusted-R^2 of 0.0040) into a negative adjusted-R^2 of -0.0034.

Even after these truncations, ten of the models that had negative plain OOS $\overline{R}^2$'s still have negative CT OOS $\overline{R}^2$'s. Among the eleven IS significant models, seven (**cay3**, **ntis**, **e^{10}/p**, **b/m**, **e/p**, **d/y**, and **dfy**) have negative OOS $\overline{R}^2$ performance even after the truncation. Three of the models (**lty**, **ltr**, and **infl**) that benefit from the OOS truncation are not close to statistical significance IS, and thus can be ignored. All in all, this leaves four models that are both OOS and IS positive and significant:

csp, **eqis**, **d/p**, **tbl**, plus possibly **tms** (which is just barely not IS significant). We investigate these models further below.

6.3 OOS Utility Performance of a Trading Strategy

Like Brennan and Xia (2004), CT also propose to evaluate the OOS usefulness of models based on the certainty equivalence (CEV) measure of a trading strategy. Specifically, they posit a power-utility investor with an assumed risk-aversion parameter, γ, of three. This allows a conditional model to contribute to an investment strategy not just by increasing the mean trading performance, but also by reducing the variance. (Breen, Glosten, and Jagannathan (1989) have shown this to be a potentially important factor.)

Although the focus of our paper is on mean prediction, we know of no better procedure to judge the economic significance of forecasting models, and therefore follow their suggestion here. To prevent extreme investments, there is a 150% maximum equity investment. A positive investment weight is guaranteed by the truncation of equity premium predictions at zero.

CT show that even a small improvement in ΔRMSE by a model over the unconditional benchmark can translate into CEV gains that are ten times as large.[10] We can confirm this—and almost to a fault. **cay3** offers 6.1bp/month performance, even though it had a negative $\overline{R}^2$. Column 12 also shows that even models that have a negative OOS ΔRMSE (not just a negative $\overline{R}^2$), like **dfr**, can produce positive gains in CEV. This is because the risk-aversion parameter gamma of 3 is low enough to favor equity-tilted strategies. Put differently, some strategy CEV gains are due to the fact that the risky equity investment was a better choice than the risk-free rate in our

[10] CT show in equation (8) of their paper that the utility gain is roughly equal to OOS-R^2/γ. This magnification effect occurs only on the monthly horizon, because the difference between OOS-R^2 and the ΔRMSE scales with the square root of the forecasting horizon (for small ΔRMSE, OOS-$R^2 \approx 2\cdot\Delta\text{RMSE}/\text{StdDev(R)}$). That is, at a monthly frequency, the OOS-R^2 is about 43 times as large as ΔRMSE. On an annual prediction basis, this number drops from 43 to 12. An investor with a risk aversion of 10 would therefore consider the economic significance on annual investment horizon to be roughly the same as the ΔRMSE we consider. (We repeated the CT CEV equivalent at annual frequency to confirm this analysis.)

data. (This applies not only to strategies based on the conditional models, but also to the strategy based on the unconditional mean.) An alternative utility specification that raises the risk-aversion coefficient to 7.48 would have left an investor indifferent between the risk-free and the equity investments. Briefly considering this parameter can help judge the role of equity bias in a strategy; it does seem to matter for the **eqis** and **tms** models, as explained below.

In order, among the IS reasonably significant models, those providing positive CEV gains were **tms** (14bp/month), **eqis** (14bp/month), **tbl** (10bp/month), **csp** (6bp/month), **cay3**(6bp/month), and **ntis** (2bp/month).

6.4 Details

We now look more closely at the set of variables with potentially appealing forecasting characteristics. **csp**, **eqis**, **tbl**, and **tms** have positive IS performance (either statistically significant or close to it), positive OOS $\overline{R}^2$ (truncated), and positive CEV gains. **cay3** and **ntis** have negative OOS $\overline{R}^2$, but very good IS performance and positive CEV gains. **d/p** has a negative CEV gain, but is positive IS and OOS $\overline{R}^2$. Thus, we describe these seven models in more detail (and with equivalent graphs): Figure 3

1. **cay3:** The best CT performer is an alternative **cay** model that also appears in Lettau and Ludvigson (2005). It predicts the equity premium not with the linear **cay**, but with all three of its highly cointegrated ingredients up to date. We name this model **cay3**. In unreported analysis, we found that the **cay** model and **cay3** models are quite different. For most of the sample period, the unrestricted predictive regression coefficients of the **cay3** model wander far off their cointegration-restricted **cay** equivalents. The model may not be as well-founded theoretically as the Lettau and Ludvigson (2001) **cay**, but if its components are known ex-ante, then **cay3** is fair game for prediction.

   Table 3 shows that **cay3** has good performance IS, but only marginal performance OOS (a positive ΔRMSE, but a negative $\overline{R}^2$). It offers good CEV gains among the models considered, an extra 6.10 bp/month. The h superscript indicates that its trading strategy requires an extra 10% more trading turnover than the unconditional model. It also reaches the maximum permitted 150% equity investment in 13.2% of all months.

A first drawback is that the **cay3** model relies on data that may not be immediately available. Its components are publicly released by the BEA about 1-2 months after the fact. Adding just one month delay to trading turns **cay3**'s performance negative:

	ΔRMSE	$\Delta\text{RMSE}_{\text{TU}}$	ΔCEV
Immediate Availability (CT)	–2.88 bp	+0.88 bp	+6.10 bp
One Month Delayed	–5.10 bp	–1.62 bp	–11.82 bp
Two Months Delayed	–5.38 bp	–1.11 bp	–9.80 bp

A second drawback is visible in Figure 3. Like **caya** and **cayp**, much of **cay3**'s performance occurs around the Oil Shock (most of its OOS performance and between one-half and one-third of its IS performance). Even IS, **cay3** has not performed well for over 30 years now:

cay3 (CT)	Recent 30 Years	All Years
IS $\overline{R}^2$	–0.30%	1.87%
OOS $\overline{R}^2$	–1.60%	–0.34%

Finally, the figure shows that many of **cay3**'s recent equity premium forecasts have been negative and therefore truncated. And, therefore, the information in its current forecasts is limited.

2. **csp:** Table 3 shows that the relative valuations of high- over low-beta stocks had good IS and truncated OOS performance, and offered a market timer 6.12 bp/month superior CEV-equivalent performance. The plot in Figure 3 shows that **csp** had good performance from September 1965 to March 1980. It underperformed by just as much from about April 1980 to October 2000. In fact, from its first OOS prediction in April 1957 to August 2001, **csp**'s total net performance

was zero even after the CT truncations, and both IS and OOS. All of **csp**'s superior OOS performance has occurred since mid-2001. Although it is commendable that it has performed well late rather than early, better performance over its first 45 years would have made us deem this variable more reliable.

The plot raises one other puzzle. The CT truncated version performs better than the plain OLS version because it truncated the **csp** predictions from July 1957 through January 1963. These CT truncations are critically responsible for its superior OOS performance, but make no difference thereafter. It is the truncation treatment of these specific 66 months that would make an investor either believe in superior positive or inferior outright negative performance for **csp** (from August 2001 to December 2005). We do not understand why the particular 66 month period from 1957 to 1963 is so crucial.

Finally, the performance during the Oil Shock recession is not important for IS performance, but it is for the OOS performance. It can practically account for its entire out-of-sample performance. Since the Oil Shock, **csp** has outperformed IS, but not OOS:

csp (CT)	Recent 30 Years	All Years
IS $\overline{R}^2$	0.33%	0.99%
OOS $\overline{R}^2$	−0.41%	0.15%

3. **ntis:** Net issuing activity had good IS performance, but a negative OOS $\overline{R}^2$. Its CEV gain is a tiny 1.53 bp/month. These 1.53 bp are likely to be offset by trading costs to turn over an additional 4.6% of the portfolio every month.[11] The strategy was very optimistic, reaching the maximum 150% investment constraint in 57.4% of all months. We do not report it in the table, but an investor with a higher 7.48 risk-aversion parameter, who would not have been so eager to highly lever herself

[11] Keim and Madhavan (1997) show that one typical roundtrip trade in large stocks for institutional investors would have conservatively cost around 38 bp from 1991–1993. Costs for other investors and earlier time-periods were higher. Futures trading costs are not easy to gauge, but a typical contract for a notional amount of $250,000 costs around $10-$30. A 20% movement in the underlying index—about the annual volatility—would correspond to $50,000, which would come to around 5 bp.

into the market, would have experienced a negative CEV with an **ntis** optimized trading strategy. Finally, the plot shows that almost all of the **csp** model's IS power derives from its performance during the Great Depression. There was really only a very short window from 1982 to 1987 when **csp** could still perform well.

4. **eqis:** Equity Issuing Activity had good IS performance, good OOS performance, and improved the CEV for an investor by a meaningful 13.67 bp/month. It, too, was an optimistic equity-aggressive strategy. With a $\gamma = 3$, trading based on this variable leads to the maximum permitted equity investment of 150% in 56% of all months. Not reported, with the higher risk-aversion coefficient of 7.48, that would leave an investor indifferent between bonds and stocks, the 13.67bp/month gain would shrink to 8.74bp/month.

   As in the annual data, Figure 3 shows that **eqis**'s performance relies heavily on the good Oil Shock years. It has not performed well in the last thirty years.

eqis (CT)	Recent 30 Years	All Years
IS $\overline{R}^2$	−0.88%	0.80%
OOS $\overline{R}^2$	−1.00%	0.30%

5. **d/p:** The dividend price ratio has good IS and OOS $\overline{R}^2$. (The OOS $\overline{R}^2$ is zero when predicting log premia.) An investor trading on **d/p** would have lost the CEV of 10bp/month. (Not reported, a more risk-averse investor might have broken even.) The plot shows that **d/p** has not performed well over the last 30 years; **d/p** has predicted negative equity premia since January 1992.

d/p (CT)	Recent 30 Years	All Years
IS $\overline{R}^2$	−0.39%	0.33%
OOS $\overline{R}^2$	−1.09%	0.17%

6. **tbl:** The short rate is insignificant IS if we forecast log premia. If we forecast unlogged premia, it is statistically significant IS at the 9% level, although this declines further if we apply the CT truncation. In its favor, **tbl**'s full-sample CT-truncated performance is statistically significant OOS, and it offers a respectable 9.53 bp/month market timing advantage. The plot shows that this is again largely Oil Shock dependent. **tbl** has offered no advantage over the last thirty years.

tbl (CT)	Recent 30 Years	All Years
IS $\overline{R}^2$	–0.41%	0.20%
OOS $\overline{R}^2$	–1.06%	0.25%

7. **tms:** The term-spread has IS significance only at the 10.1% level. (With logged returns, this drops to the 14.5% level.) Nevertheless, **tms** had solid OOS performance, either with or without the CT truncation. As a consequence, its CEV gain was a respectable 14.40 bp/month. Not reported in the table, when compared to the CEV gain of an investor with a risk-aversion coefficient of 7.48, we learn that about half of this gain comes from the fact that the term-spread was equity heavy. (It reaches its maximum of 150% equity investment in 59.3% of all months.) The figure shows that TMS performed well in the period from 1970 to the mid-1980s, that TMS has underperformed since then, and that the Oil Shock gain was greater than the overall OOS sample performance of **tms**. Thus,

tms (CT)	Recent 30 Years	All Years
IS $\overline{R}^2$	–0.19%	0.18%
OOS $\overline{R}^2$	–0.81%	0.21%

b/m, **e/p**, **e^{10}/p**, **d/y**, and **dfy** have negative OOS $\overline{R}^2$ and/or CT CEV gains, and so are not further considered. The remaining models have low or negative IS $\overline{R}^2$, and therefore should not be considered, either. Not reported, among the models that are IS insignificant, but OOS significant, none had positive performance from 1975 to today.

6.5 Comparing Findings and Perspectives

The numbers we report are slightly different from those in Campbell and Thompson (2005). In particular, they report **cay3** to have a ΔRMSE of 0.0356, more than the 0.0088 we report. This can be traced back to three equally important factors: they end their data 34 months earlier (in 2/2003), they begin their estimation one month later (1/1952), and they use an earlier version of the **cay** data from Martin Lettau's website. Differences in other variables are sometimes due to use of pre-1927 data (relying on price changes because returns are not available) for estimation though not prediction, while we exclude all pre-1927 data.

More importantly, our perspective is different from CT's. We believe that the data suggests not only that these models are not good enough for actual investing, but also that the models are not stable. Therefore, by and large, we consider even their IS significance to be dubious. Because they fail stability diagnostics, we would recommend against their continued use. Still, we can agree with some points CT raise:

1. One can reasonably truncate the models' predictions.
2. On shorter horizons, even a small predictive ΔRMSE difference can gain a risk-averse investor good CEV gains.
3. OOS performance should not be used for primary analysis.

We draw different conclusions from this last point. We view OOS performance not as a substitute but as a necessary complement to IS performance. We consider it to be an important regression *diagnostic*, and *if and only if* the model is significant IS. Consequently, we disagree with the CT analysis of the statistical power of OOS tests. In our view, because the OOS power matters only if the IS regression is statistically significant, the power of the OOS tests is conditional and thus much higher than suggested in CT, Cochrane (2005), and elsewhere. Of course, any additional diagnostic test can only reject a model—if an author is sure that the linear specification is correct, then not running the OOS test surely remains more powerful.

In judging the usefulness of these models, our paper attaches more importance than CT to the following facts:

1. Most models are not IS significant. That is, many variables in the academic literature no longer have IS significance (even at the 90% level). It is our perspective that this disqualifies them as forecasters for researchers without strong priors.

2. After three decades of poor performance, often even IS, one should further doubt the stability of most prediction models.

3. Even after the CT truncation, many models earn negative CEV gains.

4. What we call OOS performance is not truly OOS, because it still relies on the same data that was used to establish the models. (This is especially applicable to **eqis** and **csp**, which were only recently proposed.)

5. For practical use, an investor would have had to have known *ex-ante* which of the models would have held up, and that none of the models had superior performance over the last three decades—in our opinion because the models are unstable.

We believe it is now best left to the reader to concur either with our or CT's perspective. (The data is posted on the website.)

7 Alternative Specifications

We now explore some other models and specifications which have been proposed as improvements over the simple regression specifications.

7.1 Longer-Memory Dividend and Earnings Ratios

Table 4 considers dividend-price ratios, earnings-price ratios, and dividend-earnings ratios with memory (which simply means that we consider sums of multiple year dividends or earnings in these ratios). The table is an excerpt from a complete set of 1-year, 5-year, and 10-year dividend-price ratios, earnings-price ratios, and dividend-earnings ratios. (That is, we tried all 90 possible model combinations.) The table contains *all* 27 IS significant specifications from our monthly regressions that begin forecasting in 1965, and from our annual and five-yearly forecasts that begin forecasting either in 1902 or 1965.

Table 4: Long-Memory Ratios

Even though there were more combinations of dividend-earnings ratios than either dividend-price or earnings-price ratios, not a single dividend-earnings ratio turned out IS statistically significant. The reader can also see that out of our 27 IS significant models, only 5 had OOS positive and statistically significant performance. (For 2 of these models, the OOS significance is modest, not even reaching the 95% significance level.) Unreported graphs show that none of these performed well over the last 3 decades. (We also leave it to the readers to decide whether they believe that real-world investors would have been able to choose the right five models for prediction, and to get out right after the Oil Shock.)

7.2 Different Estimation Methods To Improve Power For Nonstationary Independent Variables

Stambaugh (1999) shows that predictive coefficients in small samples are biased if the independent variable is close to a random walk. Many of our variables have autoregressive coefficients above 0.5 on monthly frequency. Goyal and Welch (2003) show that **d/p** and **d/y**'s auto-correlations are not stable but themselves increase over the sample period, and similar patterns occur with other variables in our study. (The exceptions are **ntis**, **ltr**, and **dfy**.) Our previously reported statistics took stable positive autoregressive coefficients into account, because we bootstrapped for significance levels mimicking the IS autocorrelation of each independent variable.

However, one can use this information itself to also design more powerful tests. Compared to the plain OLS techniques in our preceding tables, the Stambaugh coefficient correction is a more powerful test in non-asymptotic samples. There is also information that the autocorrelation is not constant for the dividend ratios, which we are ignoring in our current paper. Goyal and Welch (2003) use rolling dividend-price ratio and dividend-growth autocorrelation estimates as instruments in their return predictions. This is model specific, and thus can only apply to one model, the dividend price ratio (**d/p**). In contrast, Lewellen (2004) and Campbell and Yogo (2006) introduce two further statistical corrections, extending Stambaugh (1999) and assuming different boundary behavior. This subsection, therefore, explores equity premium forecasts using these corrected coefficients.

Table 5: Stambaugh and Lewellen Estimation Corrections for Non-stationary Independent Variables

In Table 5, we predict with Stambaugh and Lewellen corrected coefficients. Both methods break the link between $\overline{R}^2$ (which is maximized by OLS) and statistical significance. The Lewellen coefficient is often dramatically different from the OLS coefficients, resulting in negative $\overline{R}^2$, even among its IS significant variable estimations. However, it is also tremendously powerful. Given our bootstrapped critical rejection levels under the NULL hypothesis, this technique is able to identify eight (rather than just three) ALTERNATIVE models as different from the NULL. In six of them, it even imputes significance in each and every one of our 10,000 bootstraps!

Unfortunately, neither the Stambaugh nor the Lewellen technique manage$\boxed{\text{s}}$ to improve OOS prediction. Of all models, only the **e/p** ratio in the Lewellen specification seems to perform better with a positive ΔRMSE. However, like other variables, it has not performed particularly well over the most recent 30 years—even though it has non-negative OOS ΔRMSE (but not $\overline{R}^2$) performance over the last three decades.

e/p (Lewellen)	Recent 30 Years	All Years
IS $\overline{R}^2$	–0.16%	0.02%
OOS $\overline{R}^2$	–0.08%	–0.01%

7.3 Encompassing Tests

Our next tests use encompassing predictions. A standard encompassing test is a hybrid of ex-ante OOS predictions and an ex-post optimal convex combination of unconditional forecast and conditional forecast. A parameter λ gives the ex-post weight on the conditional forecast for the optimal forecast that minimizes the ex-post MSE. The ENC statistic in equation (7) can be regarded as a test statistic for λ. If λ is between 0 and 1, we can think of the combination model as a "shrinkage" estimator. It produces an optimal combination OOS forecast error, which we denote $\Delta\text{RMSE}^{\star}$. However, investors would not have known the optimal *ex-post* λ. This means that they would have computed λ based on the best predictive up-to-date combination of the two OOS model (NULL and ALTERNATIVE), and then would have used this λ to forecast one month ahead. We denote the relative OOS forecast error of this rolling λ procedure as $\Delta\text{RMSE}^{\star r}$.[12]

Table 6: Encompassing Tests

Table 6 shows the results of encompassing forecast estimates. Panel A predicts annual equity premia. Necessarily, all ex-post λ combinations have positive $\Delta\text{RMSE}^{\star}$— but almost all rolling λ combinations have negative $\Delta\text{RMSE}^{\star r}$. The exceptions are **d/e** and **cayp** (with OOS knowledge). In some but not all specifications, this also applies to **dfy**, **all**, and **caya**. **d/e**, **dfy**, and **all** can immediately be excluded, because their optimal λ is negative. This leaves **caya**. Again, not reported, **caya** could not outperform over the most recent three decades. In the monthly rolling encompassing tests (not reported), only **svar** and **d/e** (in one specification) are positive, neither with a positive λ.

In sum, "learned shrinking" does not improve any of our models to the point where we would expect them to outperform.

[12]For the first three observations, we presume perfect optimal foresight, resulting in the minimum ΔRMSE. This tilts the rolling statistic slightly in favor of superior performance. The results remain the same if we use reasonable variations.

8 Other Literature

Our paper is not the first to explore or to be critical of equity premium predictions. Many bits and pieces of evidence we report have surfaced elsewhere, and some authors working with the data may already know which models work, and when and why—but this is not easy to systematically determine for a reader of this literature. There is also a publication bias in favor of significant results—non-findings are often deemed less interesting. Thus, the general literature tenet has remained that the empirical evidence and professional consensus is generally supportive of predictability. This is why we believe that it is important for us to review models in a comprehensive fashion—variable-wise, horizon-wise, and time-wise—and to bring all variables up-to-date. The updating is necessary to shed light on post-Oil Shock behavior and explain some otherwise startling disagreements in the literature.

There are many other papers that have critiqued predictive regressions. In the context of dividend ratios, see, e.g., Goetzmann and Jorion (1993) and Ang and Bekaert (2003). A number of papers have also documented low in-sample power (e.g., see Goetzmann and Jorion (1993), Nelson and Kim (1993), and Valkanov (2003)). We must apologize to everyone whose paper we omit to cite here—the literature is simply too voluminous to cover fully.

The papers that explore model instability and/or OOS tests have the closest kinship to our own. The possibility that the underlying model has changed (often through regime shifts) has also been explored in such papers as Heaton and Lucas (2000), Jagannathan, McGrattan, and Scherbina (2000), Bansal, Tauchen, and Zhou (2004), and Kim, Morley, and Nelson (2005), and Lettau and Van Nieuwerburgh (2005). Interestingly, Kim, Morley, and Nelson (2005) cannot find any structural univariate break post WW-II. Bossaerts and Hillion (1999) suggest one particular kind of change in the underlying model—a disconnect between IS and OOS predictability because investors themselves are learning about the economy.

Again, many of the earlier OOS tests have focused on the dividend ratios.

- Fama and French (1988) interpret the OOS performance of dividend ratios to have been a success. Our paper comes to the opposite conclusion primarily because we have access to a longer sample period.
- Bossaerts and Hillion (1999) interpret the OOS performance of the dividend yield (not dividend price ratio) to be a failure, too. However, they rely on a larger cross-section of 14 (correlated) countries and not on a long OOS time period (1990–1995). Because this was a period when the dividend-yield was known to have performed poorly, the findings were difficult to generalize.
- Ang and Bekaert (2003) similarly explore the dividend yield in a more rigorous structural model. They, too, find poor OOS predictability for the dividend yield.
- Goyal and Welch (2003) explore the OOS performance of the dividend ratios in greater detail on annual horizons. (Our current paper has much overlap in perspective, but little overlap in implementation.)

Lettau and Ludvigson (2001) run rolling OOS regressions—but not in the same spirit as our paper: the construction of their CAY variable itself relies on ex-post coefficient knowledge. This thought experiment applies to a representative investor, who knows the full-sample estimation coefficients for CAY, but does not know the full-sample predictive coefficients. This is *not* the experiment our own paper pursues. (Lettau and Ludvigson also do not explore their model's stability, or note its performance since 1975.) Some tests are hybrids between IS and OOS tests (as are our encompassing tests). For example, Fisher and Statman (2005) explore mechanical rules based on P/E and dividend-yield ratios, which are based on pre-specified numerical cutoff values. None works robustly across countries.

Most of the above papers have focused on a relatively small number of models. There are at least three studies in which authors seek to explore more comprehensive sets of variables:

- Pesaran and Timmerman (1995) (and others) have pointed out that our profession has snooped data (and methods) in search of models that seem to predict the equity premium in the same single U.S. or OECD data history. They explore model selection in great detail, exploring dividend-yield, earnings-price ratios, interest rates, and money in $2^9 = 512$ model variations. Their data series is monthly, begins in 1954 and ends (by necessity) twelve years ago in 1992. They conclude that investors could have succeeded, especially in the volatile periods of the 1970s (i.e., the Oil Shock). But they do not entertain the historical equity premium mean as a NULL hypothesis, which makes it difficult to compare their results to our own. Our paper shows that the Oil Shock experience generally is almost unique in making many predictive variables seem to outperform. Still, even including the two-year Oil Shock period in the sample, the overall OOS performance of our ALTERNATIVE models is typically poor.

- Ferson, Sarkissian, and Simin (2003) explore spurious regressions and data mining in the presence of serially correlated independent variables. They suggest increasing the critical t-value of the in-sample regression. The paper concludes that "many of the regressions in the literature, based on individual predictor variables, may be spurious." Torous and Valkanov (2000) disagree with Ferson, Sarkissian, and Simin. They find that a low signal-noise ratio of many predictive variables makes a spurious relation between returns and persistent predictive variables unlikely and, at the same time, would lead to no out-of-sample forecasting power.
- An independent study, Rapach and Wohar (2006), is perhaps closest to our paper. It is also fairly recent, fairly comprehensive, and explores out-of-sample performance for a number of variables. We come to many similar conclusions. Their study ends in 1999, while our data end in 2005—a fairly dramatic five years. Moreover, our study focuses more on diagnosis of weaknesses, rather than just on detection.[13]

[13] Another study by Guo (2006) finds that **svar** has OOS predictive power. However, Guo uses post WW-II sample period and downweights the fourth quarter of 1987 in calculating stock variance. We check that this is why he can find significance where we find none. In the pre-WW2 period, there are many more quarters that have even higher stock variance than the fourth quarter of 1987. If we use a longer sample period, Guo's results also disappear regardless of whether we downweight the highest observation or not.

9 Conclusion

Findings: Our paper systematically investigates the IS and OOS performance of (mostly) linear regressions that predict the equity premium with prominent variables from earlier academic research. Our analysis can be regarded as conservative because we do not even conduct a true OOS test—we select variables from previously published papers and include the very same data that were used to establish the models in the first place. We also ignore the question of how a researcher or investor would have known which among the many models we considered would ultimately have worked.

There is one model for which we feel judgment should be reserved (**eqis**), and some models that deserve more investigation on very-long term frequencies (5 years). None of the remaining models seems to have worked well. To draw this conclusion, our paper relies not only on the printed tables in this final version, but on a much larger set of tables that explored combinations of modified data definitions, data frequencies, time periods, econometric specifications, etc).[14] Our findings are not driven by a few outlier years. Our findings do not disappear if we use different definitions and corrections for the time-series properties of the independent variable. Our findings do not arise because our tests have weak power (which would have manifested itself mostly in poor early predictions). Our findings hold up if we apply statistical corrections, data driven model selection, and encompassing tests.

Instead, our view based on this evidence is now that most models seem unstable or even spurious. Our plots help diagnose when they performed well or poorly, both in-sample and out-of-sample. They shine light on the two most interesting subperiods, the 1973-75 Oil Shock, and the most recent thiry years, 1975 to today. (And we strongly suggest that future papers proposing equity premium predictive models include similar plots.) If we exclude the Oil Shock, most models perform even worse—many were statistically significant in the past only because of the stellar

[14] The tables in this paper have been distilled from a larger set of tables, which are available from our website—and on which we sometimes draw in our text description of results.

model performance during these contiguous unusual years. One can only imagine whether our profession would have been equally comfortable rationalizing away these years "as unusual" if they had been the main negative and not the main positive influence.

As of the end of 2005, most models have lost statistical significance, both in-sample and out-of-sample. Out-of-sample, most models not only fail to beat the unconditional benchmark (the prevailing mean) in a statistically or economically significant manner, but underperform it outright. If we focus on the most recent decades, i.e., the period after 1975, we find that no model had superior performance OOS and few had acceptable performance IS. With 30 years of poor performance, believing in a model today would require strong priors that the model is well specified and that the underlying model has not changed.

Of course, even today, researchers can cherry-pick models—intentionally or unintentionally. Still, this does not seem to be an easy task. It is rare that a choice of sample start, data frequency, and method leads to robust superior statistical performance in-sample. Again, to ignore OOS tests even as a diagnostic, a researcher would have to have supreme confidence that the underlying model is stable. Despite extensive search, we were unsuccessful in identifying any models on annual or shorter frequency that systematically had both good in-sample and out-of-sample performance, at least in the period from 1975 to 2005—although more search might eventually produce one. To place faith in a model, we would want to see genuine superior and stable IS and OOS performance in years after the model identification. Switching perspective from a researcher to an investor, we believe the evidence suggests that none of the academic models we reexamine warrants a strong investment endorsement today. By assuming that the equity premium was "like it always has been," an investor would have done just as well.

Directions: An academic researcher could explore more variables and/or more sophisticated models (e.g., through structural shifts or Kalman filters). Alternatively, one could predict disaggregated returns, for example, the returns on value-stocks and the returns on growth stocks. The former could respond more strongly to dividends, while the latter could respond more strongly to book-market factors. However, such explorations aggravate the problems arising from (collective) specification search. Some of these models are bound to work both IS or OOS by pure chance. At the very least, researchers should wait for more new OOS data to become available in order to accumulate faith in such new variables or more sophisticated models.

Having stated the obvious, there are promising directions. We are looking forward to accumulating more data. Lettau and Van Nieuwerburgh (2005) model structural change not based on the forecasting regression, but based on mean shifts in the dependent variables. This reduces (but does not eliminate) snooping bias. Another promising method relies on theory—an argument along the line of Cochrane's (2005) observation that the dividend yield must predict future returns eventually if it fails to predict dividend growth.[15]

Broader Implications: Our paper is simple, but we believe its implications are not. The belief that the state variables which we explored in our paper can predict stock returns and/or equity premia is not only widely held, but the basis for two entire literatures: one literature on how these state variables predict the equity premium and one literature on how smart investors should use these state variables in better portfolio allocations. This is not to argue that an investor would not update his estimate of the equity premium as more equity premium realizations come in. Updating will necessarily induce time-varying opportunity sets (see Xia

[15]We do not agree with all of Cochrane's (2005) conclusions. He has strong priors, placing full faith in a stationary specification of the underlying model—even though Goyal and Welch (2003) have documented dramatic increases in the autocorrelation of dividend growth. Therefore, he does not consider whether changes in the model over the last 30 years could lead one to the conclusion that dividend ratios do not predict *as of 2006*. He also draws a stark dichotomy between a NULL (no return prediction, but dividend growth prediction) and an ALTERNATIVE (no dividend growth prediction, but return prediction). He evaluates both hypotheses separately for dividend growth and return predictability. He then proceeds under unconditional confidence in the ALTERNATIVE to show that if dividend growth rates are truly unpredictable, then dividend ratios increase in significance to conventional levels. With residual doubts about the ALTERNATIVE, this conclusion could change.

(2001) and Lewellen and Shanken (2002)). Instead, our paper suggests only that the profession has yet to find some variable that has meaningful and robust empirical equity premium forecasting power, both IS and OOS. We hope that the simplicity of our approach strengthens the credibility of our evidence.

Website Data Sources

Robert Shiller's Website: `http://aida.econ.yale.edu/~shiller/data.htm`.

NBER Macrohistory Data Base:
`http://www.nber.org/databases/macrohistory/contents/chapter13.html`.

FRED: `http://research.stlouisfed.org/fred2/categories/22`.

Value-Line: `http://www.valueline.com/pdf/valueline_2005.pdf`.

Bureau of Labor Statistics Webpage: `http://www.bls.gov/cpi/`

Martin Lettau's Webpage: (**cay**), `http://pages.stern.nyu.edu/~mlettau/`.

William Schwert's Webpage: (**svar**), `http://schwert.ssb.rochester.edu/`.

Jeff Wurgler's Webpage: (**eqis**), `http://pages.stern.nyu.edu/~jwurgler/`

Figure 1: Annual Performance of In-Sample Insignificant Predictors

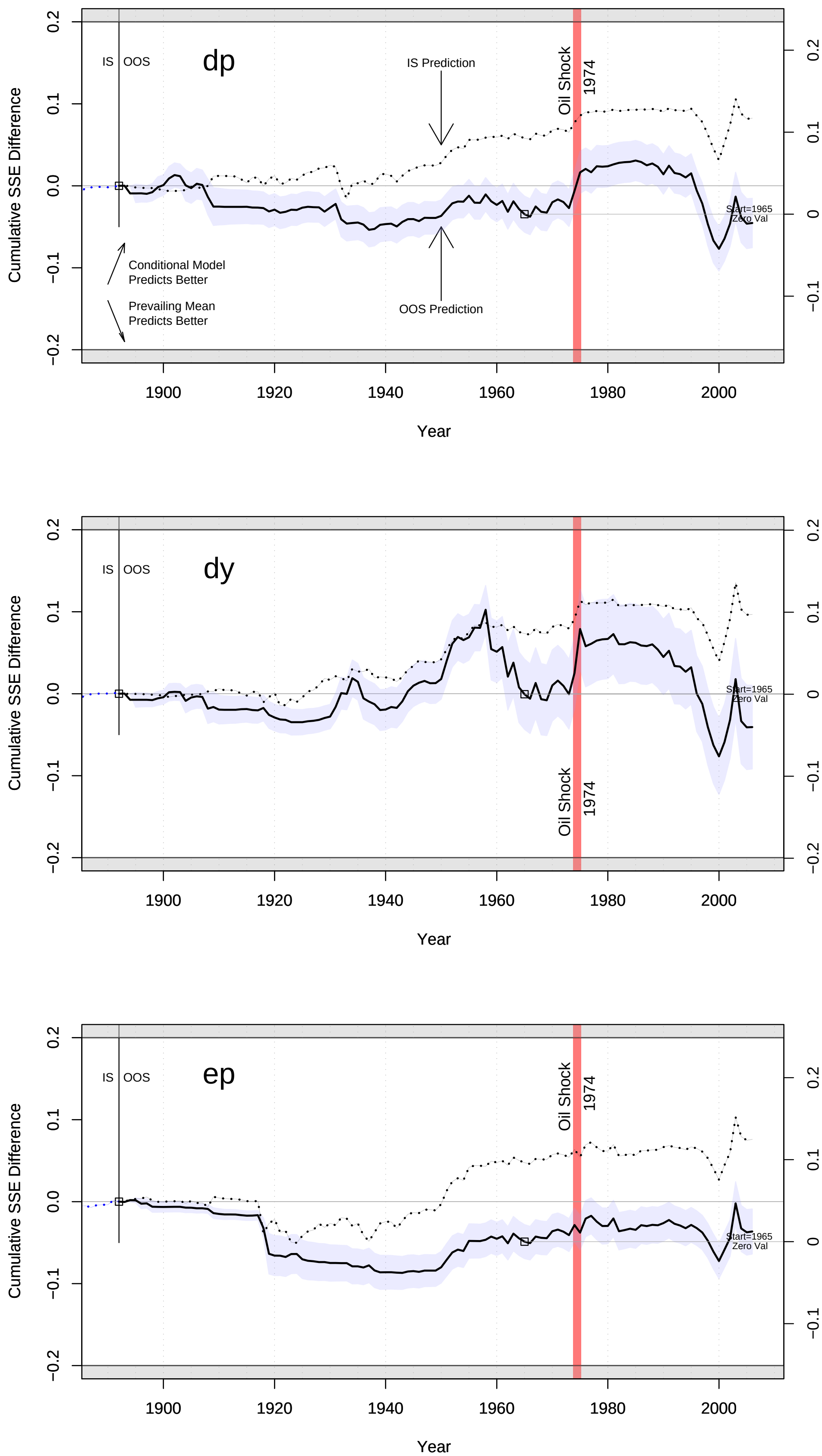


Figure 1: continued

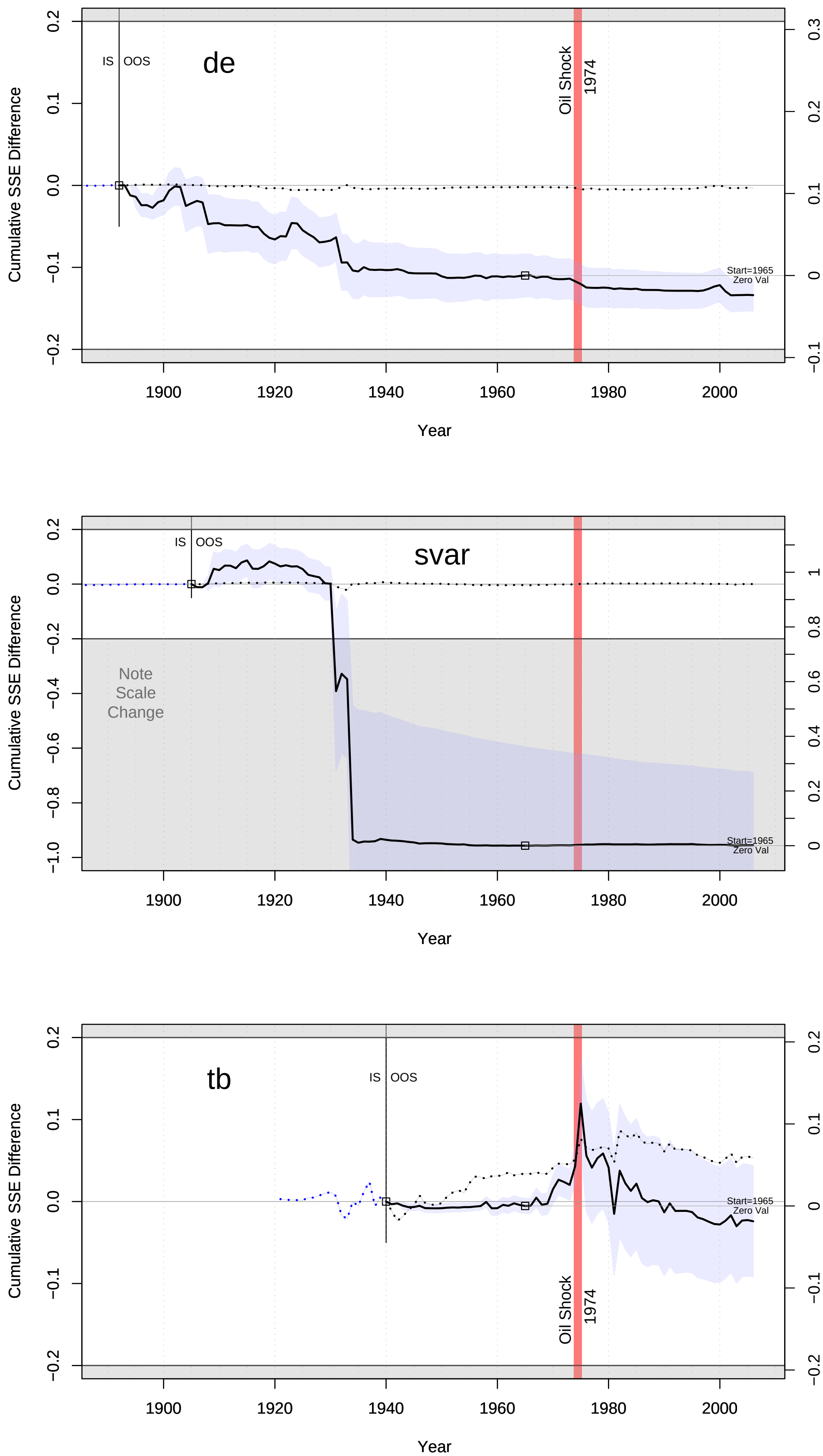


Figure 1: continued

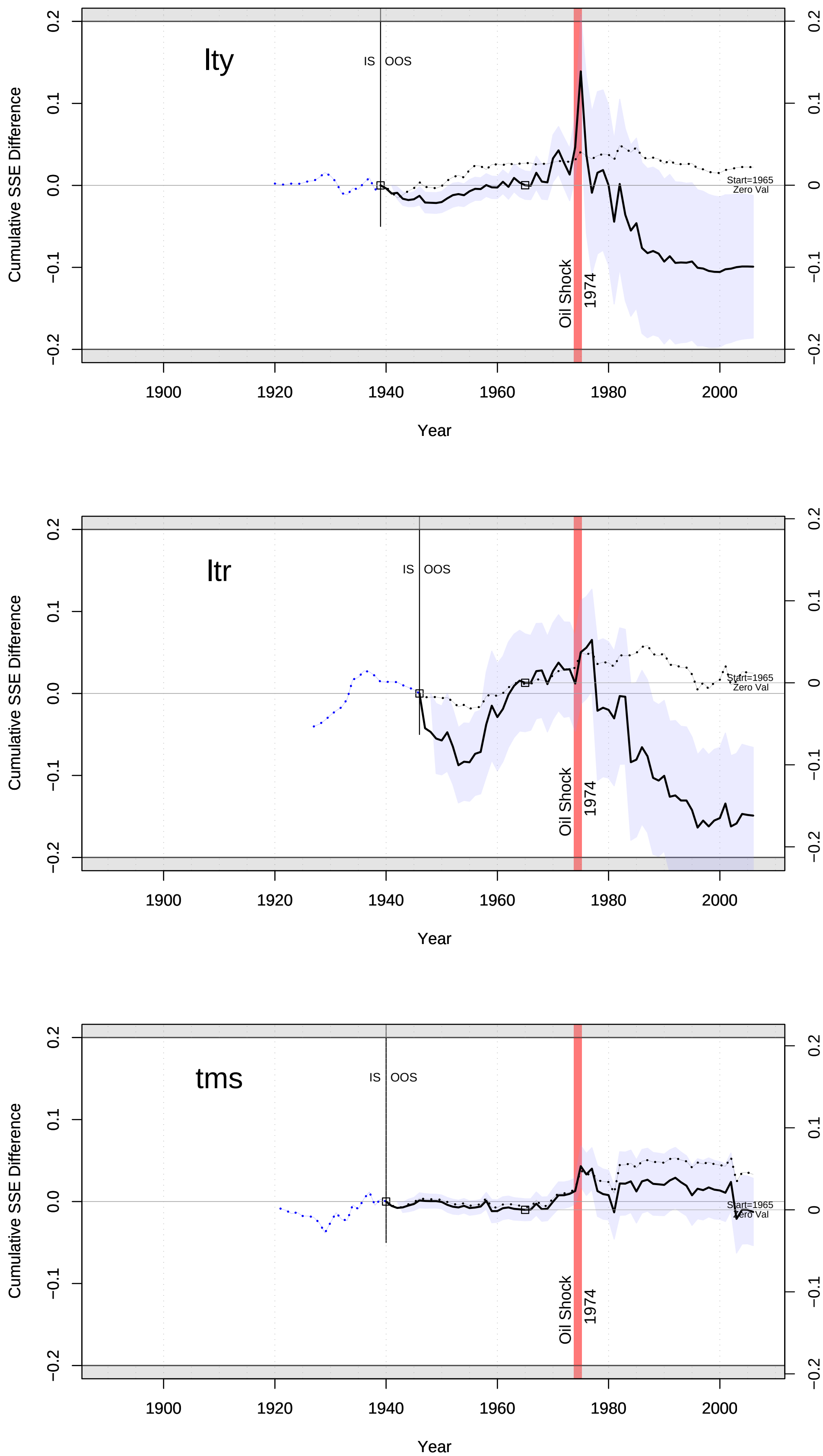
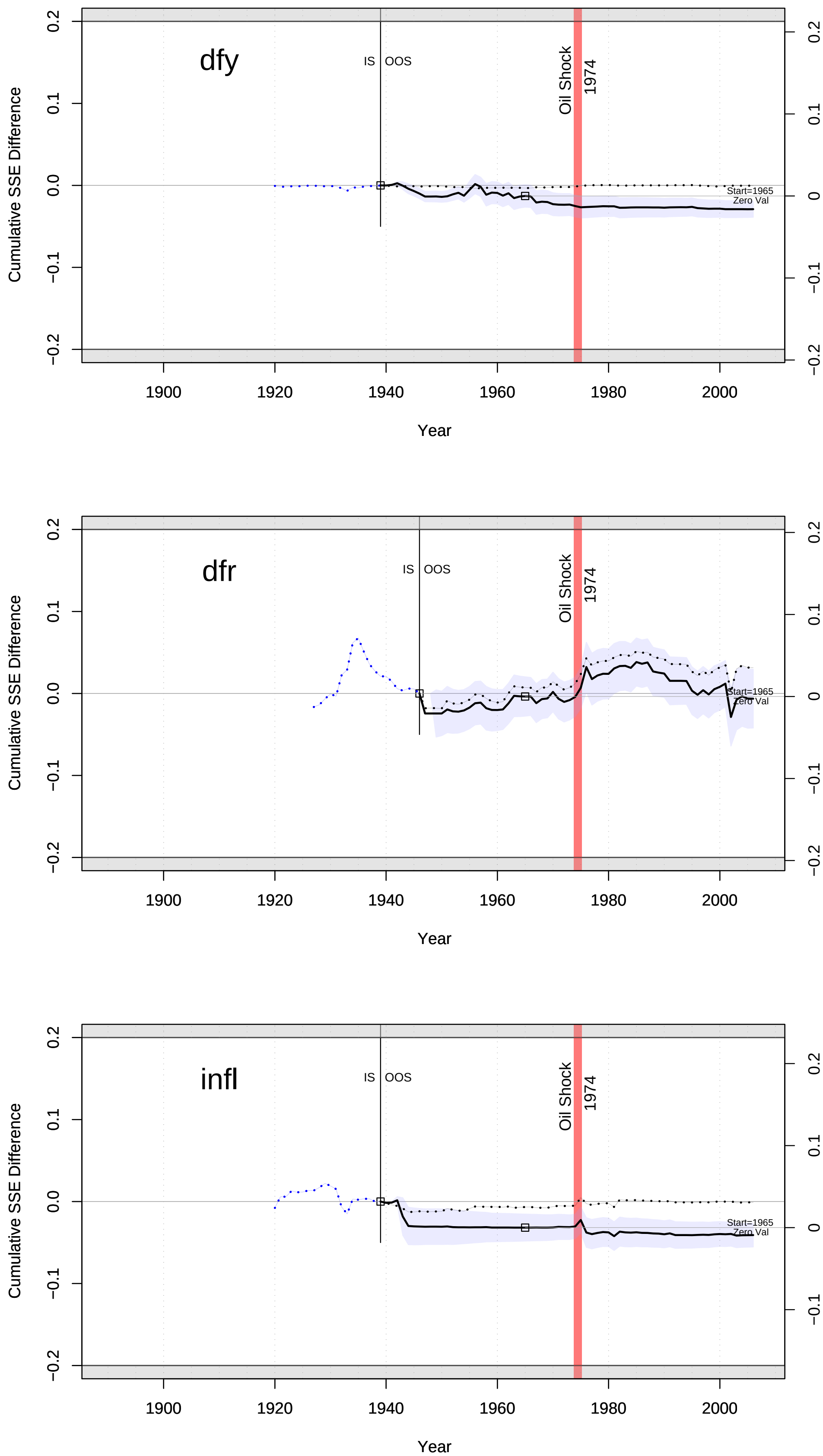
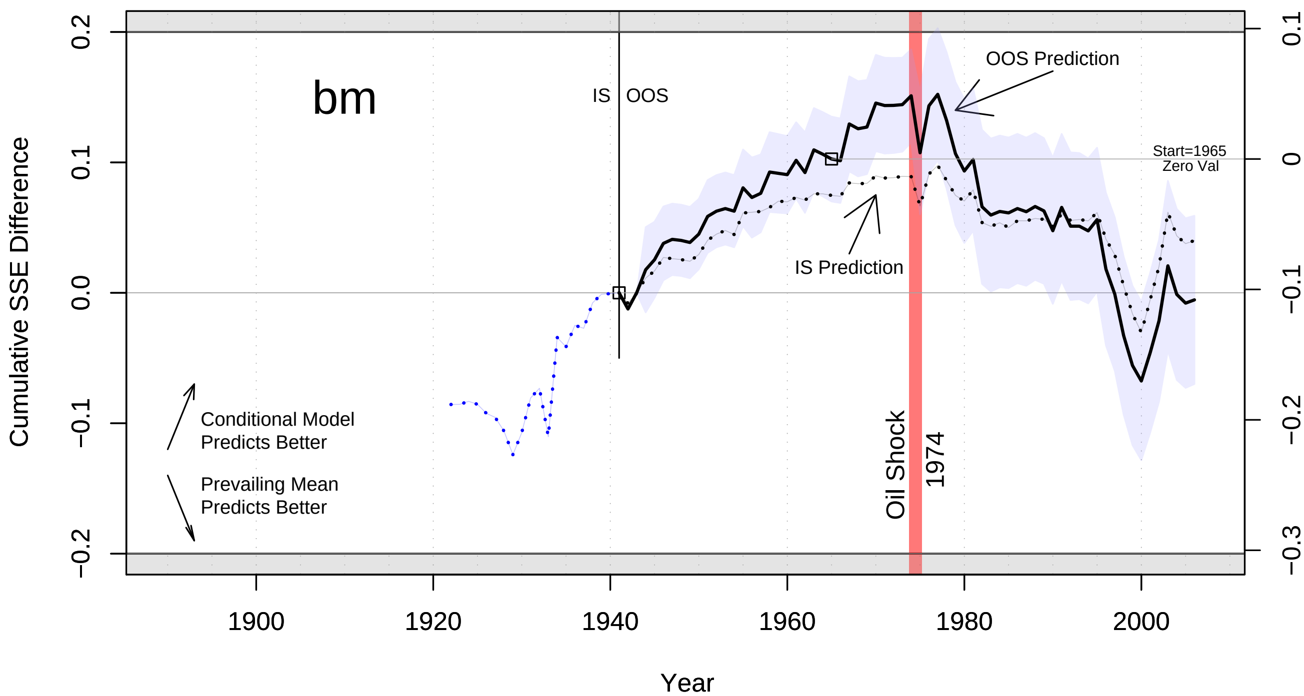


Figure 1: continued



Explanation: These figures plot the IS and OOS performance of annual predictive regressions. Specifically, these are the cumulative squared prediction errors of the NULL minus the cumulative squared prediction error of the ALTERNATIVE. The ALTERNATIVE is a model that relies on predictive variables noted in each graph. The NULL is the prevailing equity premium mean for the OOS graph, and the full-period equity premium mean for the IS graph. The IS prediction relative performance is dotted (and usually above), the OOS prediction relative perfomance is solid. An increase in a line indicates better performance of the named model; a decrease in a line indicates better performance of the NULL. The blue band is the equivalent of 95% two-sided levels, based on MSE-T critical values from McCracken (2004). (MSE-T is the Diebold and Mariano (1995) t-statistic modified by Harvey, Leybourne, and Newbold (1998)). The right axis shifts the zero point to 1965. The Oil Shock is marked by a red vertical line.

Figure 2: Annual Performance of Predictors That Are Not In-Sample Significant



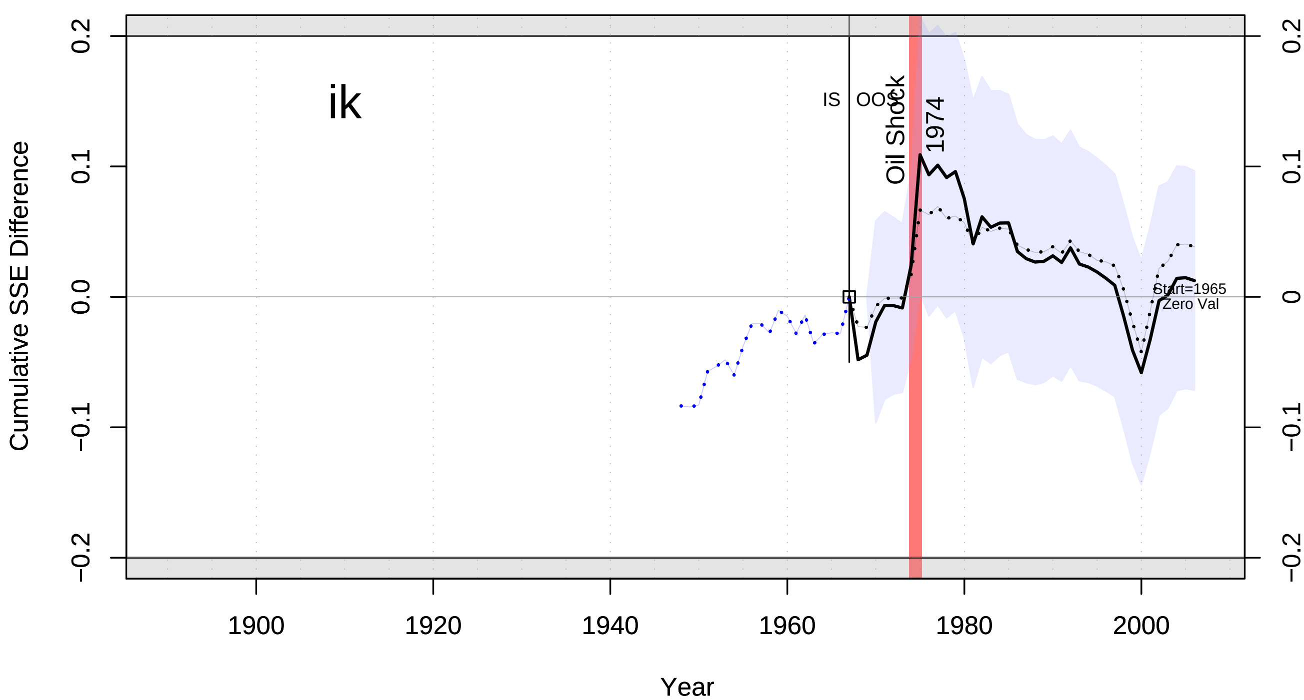
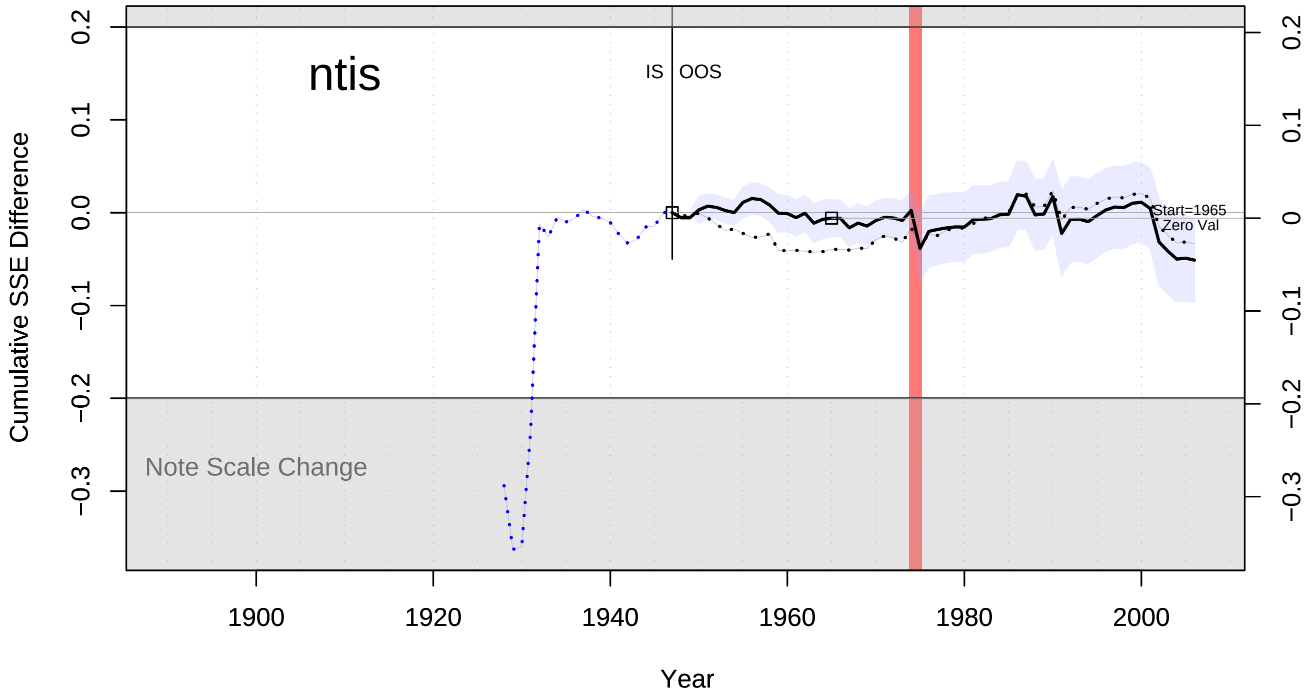


Figure 2: continued



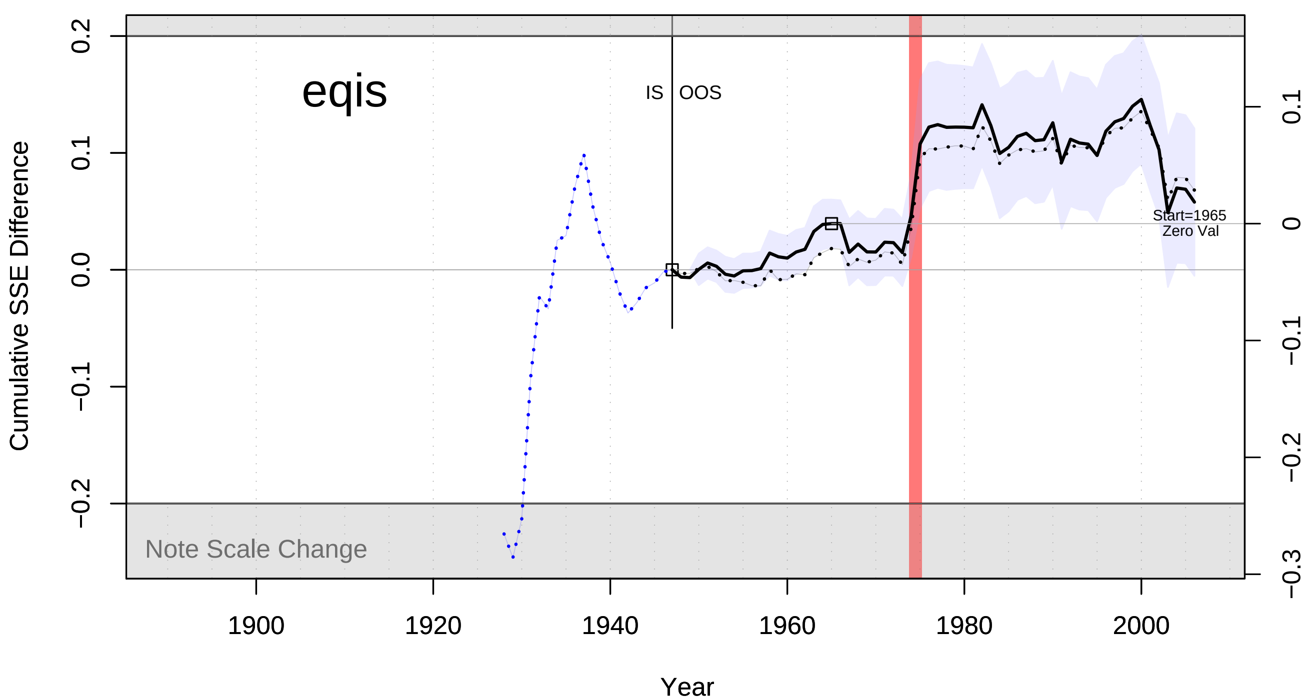
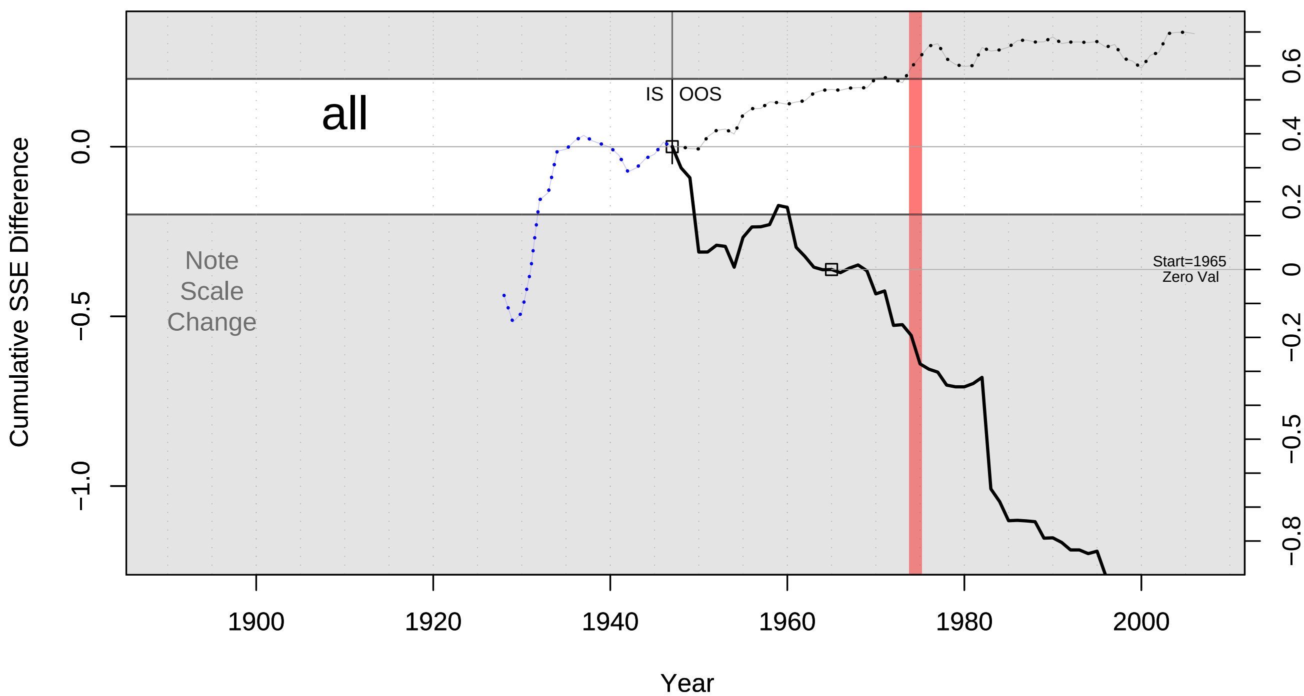


Figure 2: continued



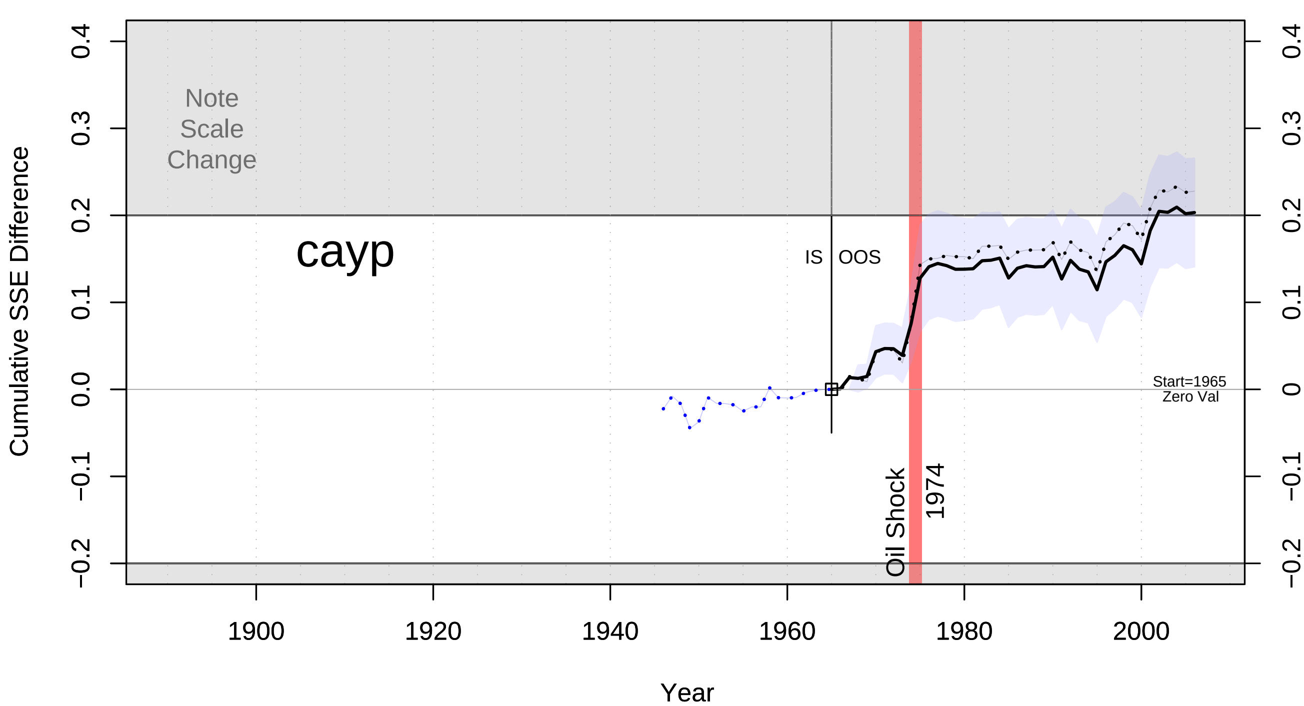
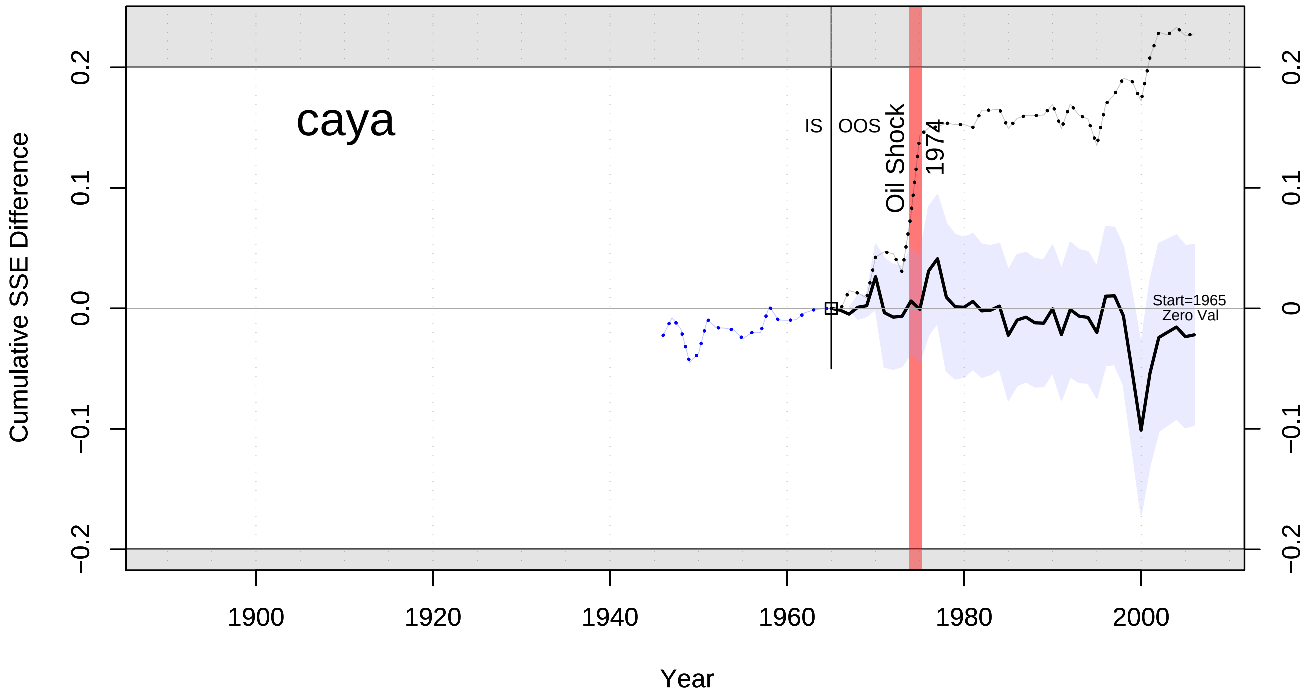
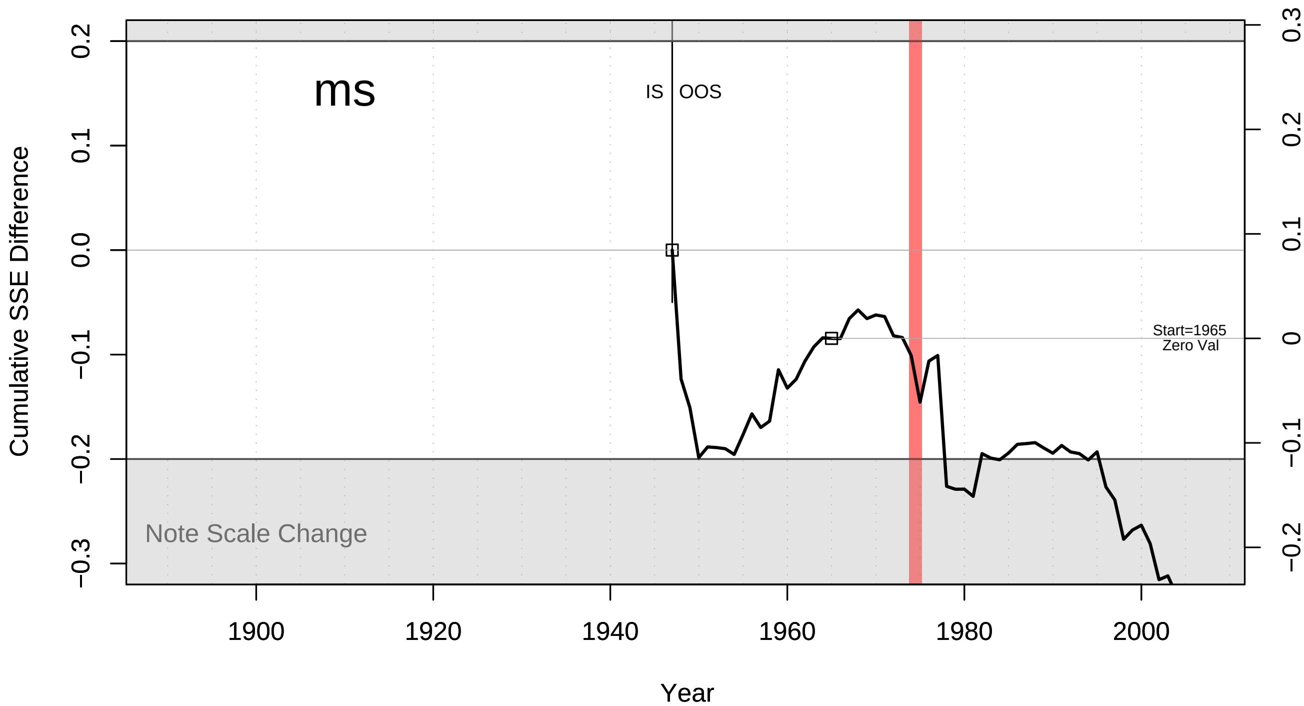


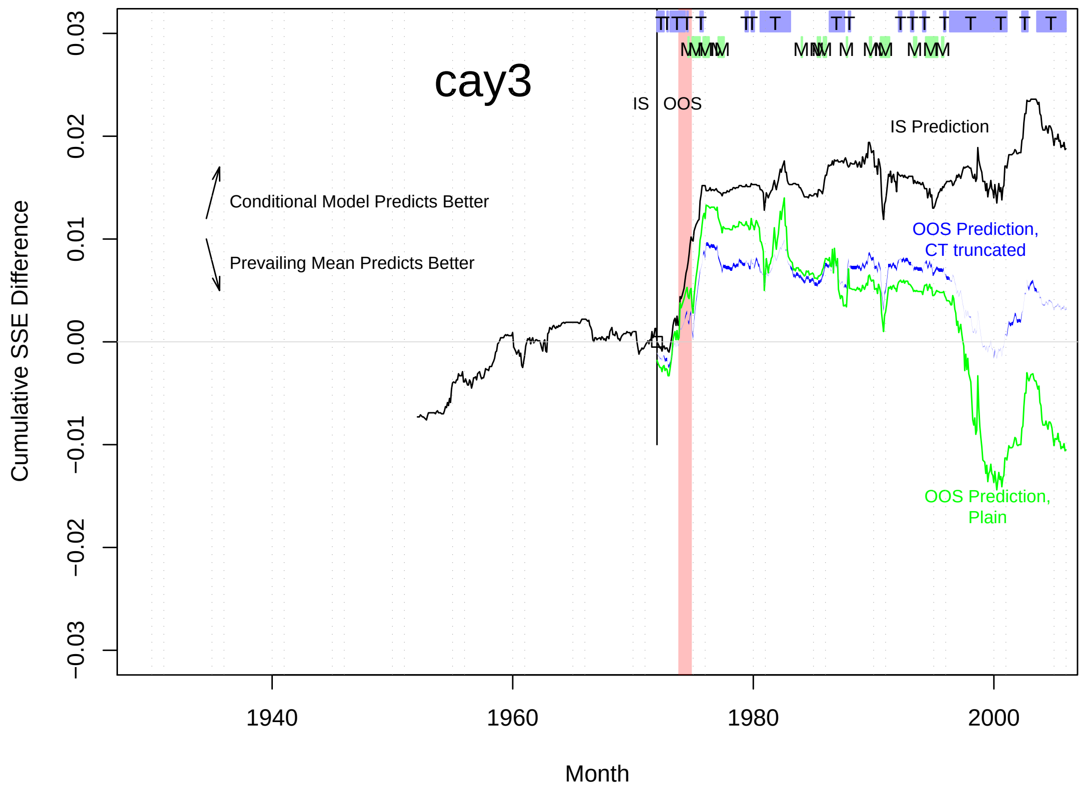
Figure 2: continued





Explanation: See Figure 1.

Figure 3: Monthly Performance of In-Sample Significant Predictors



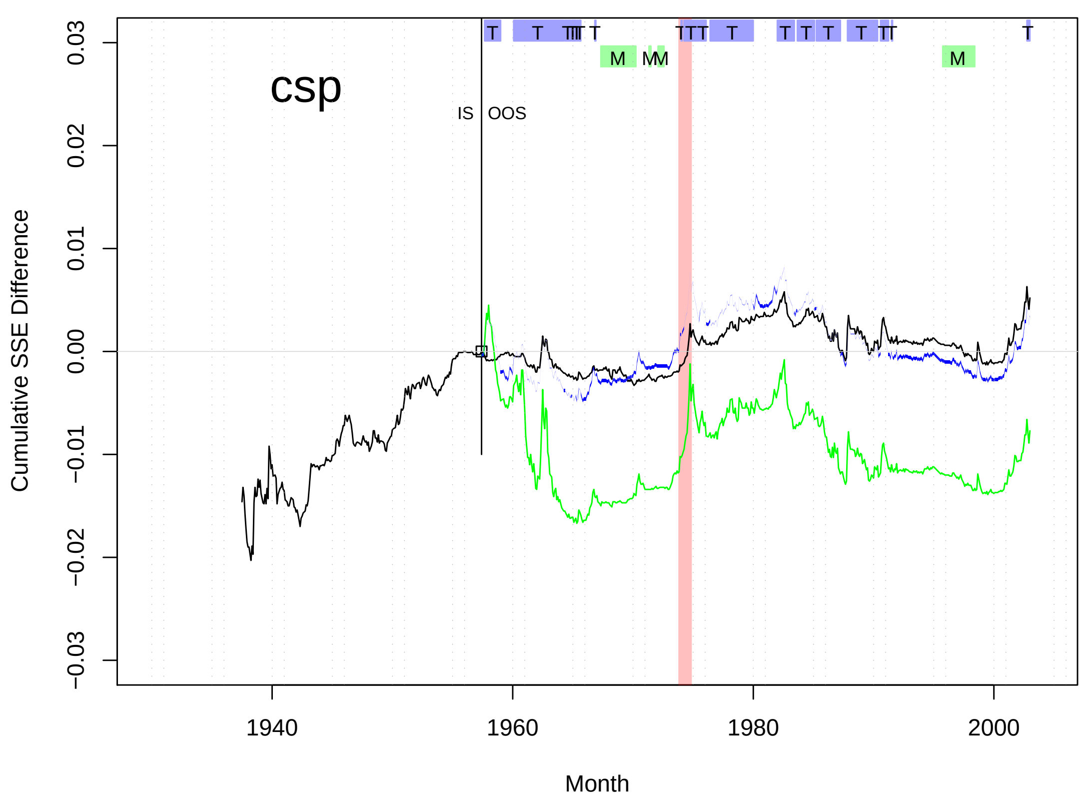
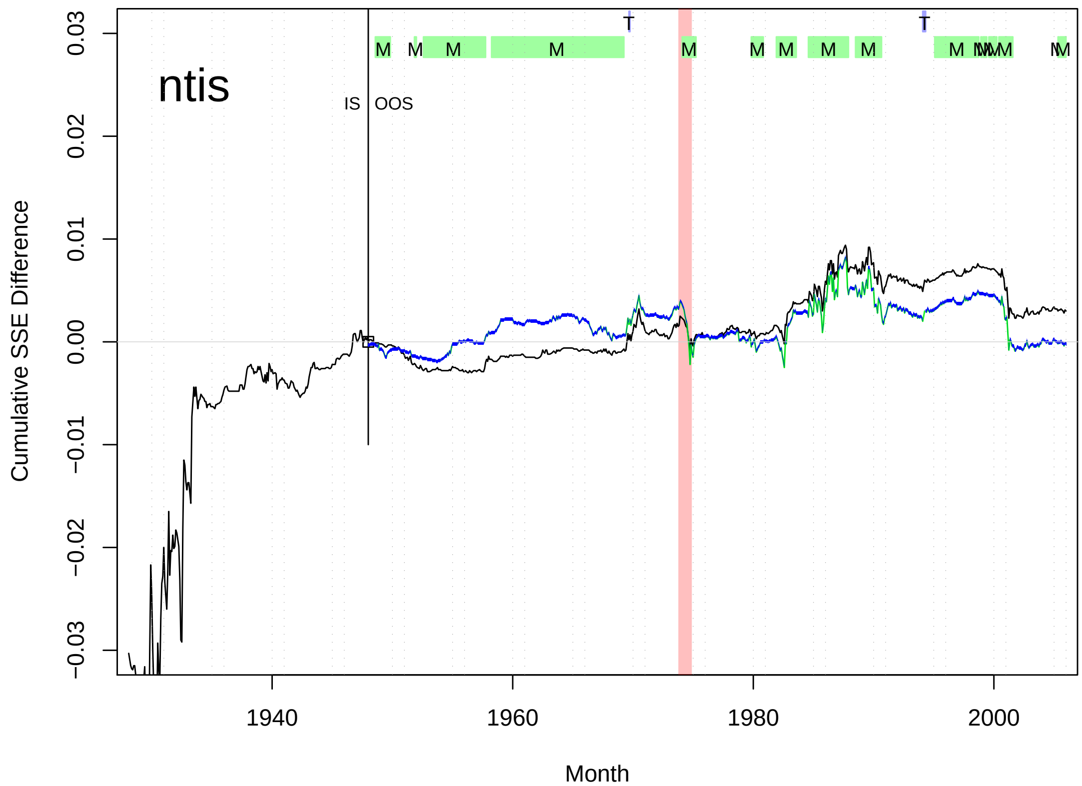


Figure 3: continued



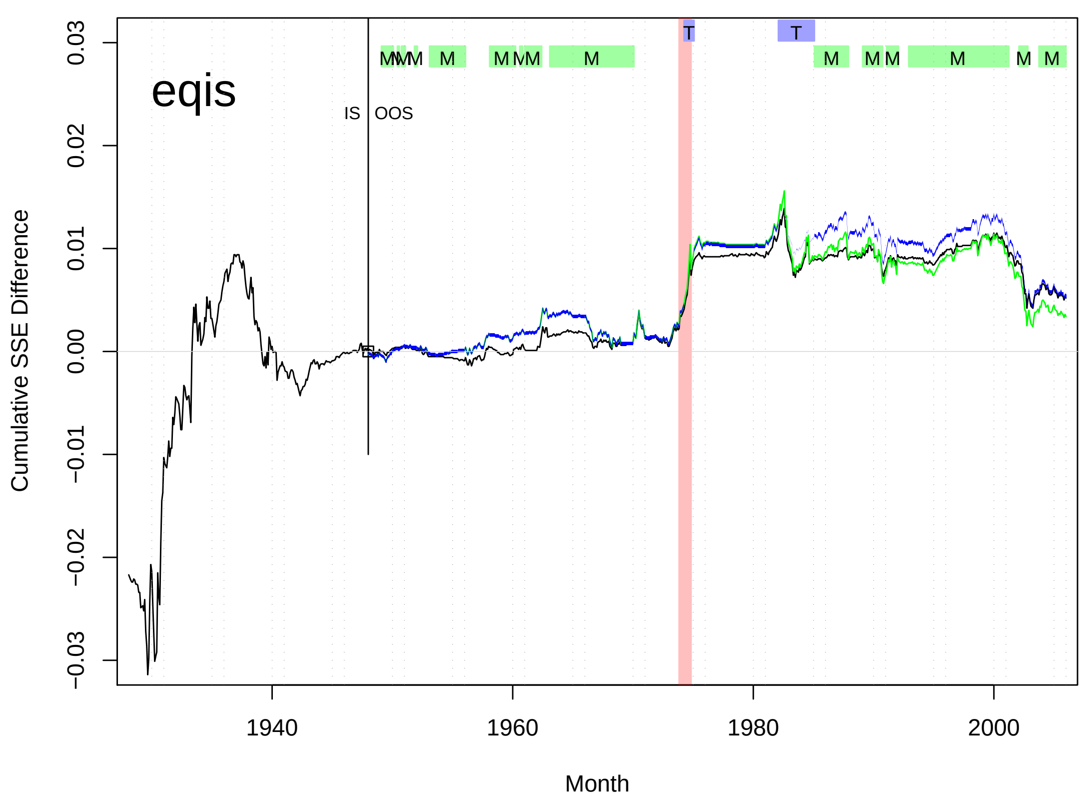
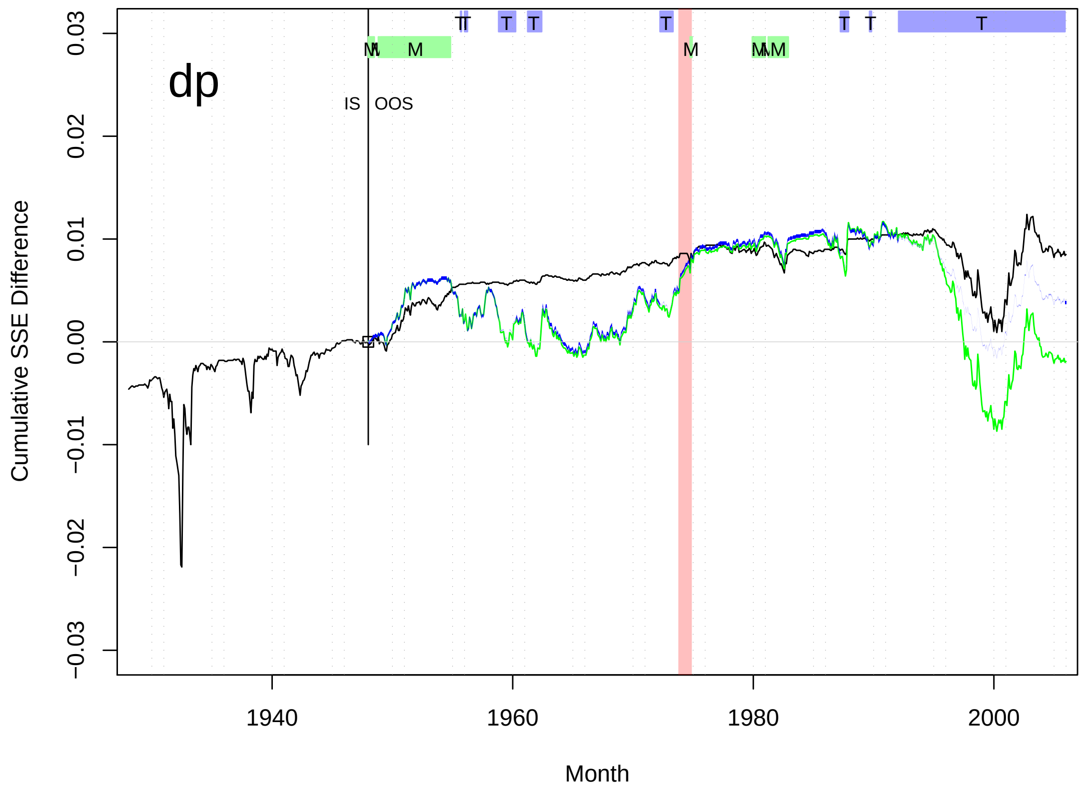


Figure 3: continued



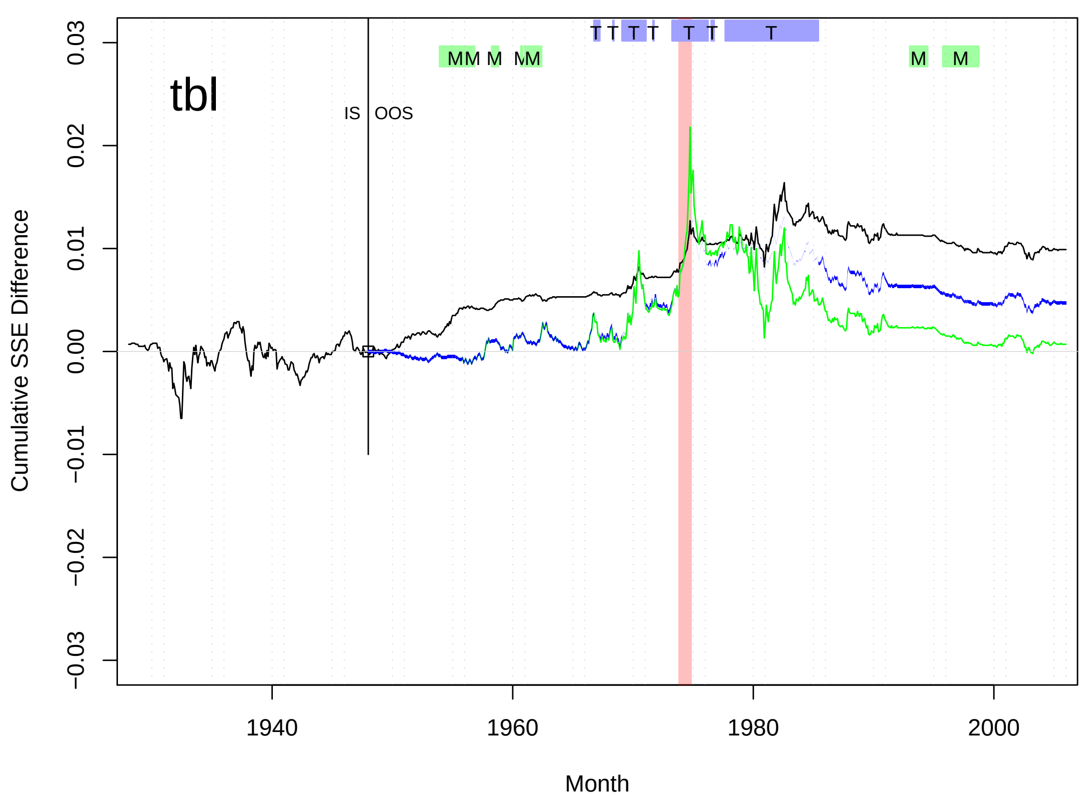
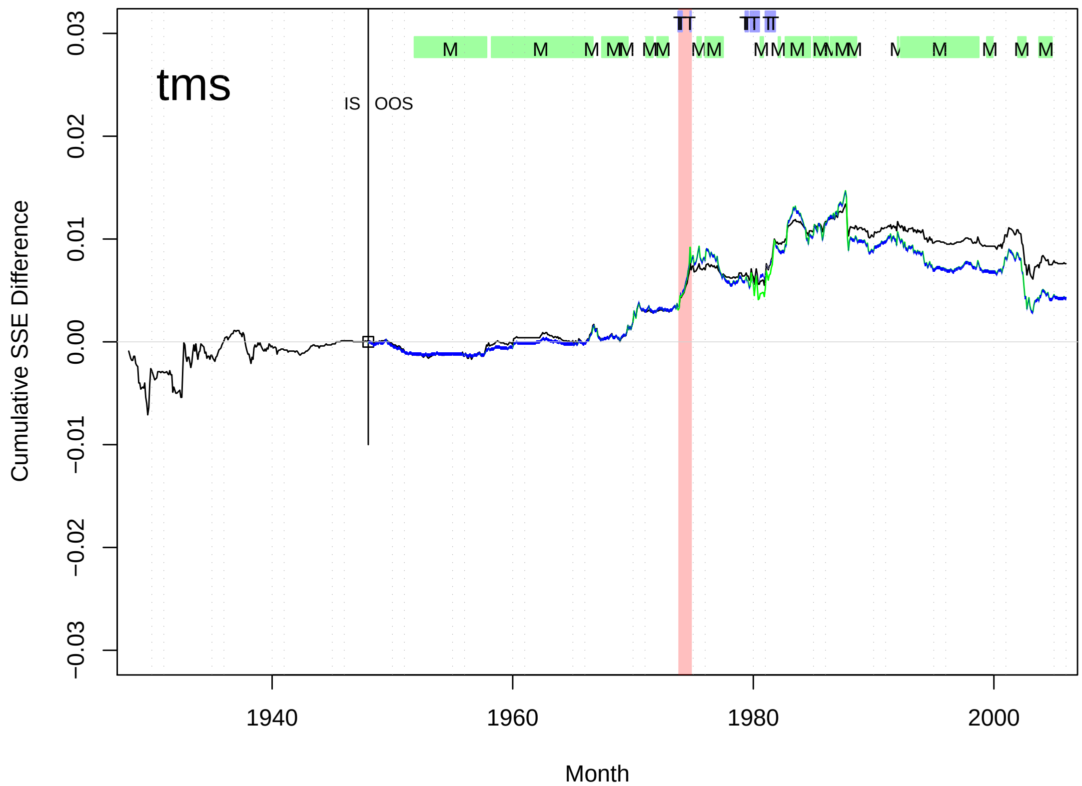


Figure 3: continued



Explanation: These figures are the analogs of Figures 1 and 2, plotting the IS and OOS performance of the named model. However, they use monthly data. The IS performance is in black. The Campbell-Thompson (2005) (CT) OOS model performance is plotted in blue, the plain OOS model performance is plotted in green. The top bars ("T") indicate truncation of the equity prediction at 0, inducing the CT investor to hold the risk-free security. (This also lightens the shade of blue in the CT line.) The lower bars ("M") indicate when the CT risk-averse investor would purchase equities worth 150% of his wealth, the maximum permitted. The Oil Shock (Nov 1973 to Mar 1975) is marked by a red vertical line.

Table 1: Forecasts at Annual Frequency

This table presents statistics on forecast errors in-sample (IS) and out-of-sample (OOS) for log equity premium forecasts at annual frequency (both in the forecasting equation and forecast). Variables are explained in Section 2. Stock returns are price changes, including dividends, of the S&P500. All numbers are in percent per year, except except $\overline{R}^2$ and power which are simple percentages. A star next to IS-$\overline{R}^2$ denotes significance of the in-sample regression as measured by F-statistics (critical values of which are obtained empirically from bootstrapped distributions). The column 'IS for OOS' gives the IS-$\overline{R}^2$ for the OOS period. ΔRMSE is the RMSE (root mean square error) difference between the unconditional forecast and the conditional forecast for the same sample/forecast period. Positive numbers signify superior out-of-sample conditional forecast. The OOS-$\overline{R}^2$ is defined in equation (6). A star next to OOS-$\overline{R}^2$ is based on significance of MSE-F statistic by McCracken (2004), which tests for equal MSE of the unconditional forecast and the conditional forecast. One-sided critical values of MSE statistics are obtained empirically from bootstrapped distributions, except for **caya** and **all** models where they are obtained from McCracken (2004). Critical values for the **ms** model are not calculated. Power is calculated as the fraction of draws where the simulated ΔRMSE is greater than the empirically calculated 95% critical value. The two numbers under the power column are for all simulations and for those simulations in which the in-sample estimate was significant at the 95% level. Significance levels at 90%, 95%, and 99% are denoted by one, two, and three stars, respectively.

	Variable	Data	Full Sample: IS $\overline{R}^2$	Forecasts begin 20 years after sample: IS for OOS $\overline{R}^2$	OOS $\overline{R}^2$	OOS ΔRMSE	OOS Power	Forecasts begin 1965: IS for OOS $\overline{R}^2$	OOS $\overline{R}^2$	OOS ΔRMSE	OOS Power	1927–2005 Sample: IS $\overline{R}^2$
Full Sample, Not Significant IS												
dfy	Default Yield Spread	1919–2005	−1.18		−3.29	−0.14			−4.15	−0.12		−1.31
infl	Inflation	1919–2005	−1.00		−4.07	−0.20			−3.56	−0.08		−0.99
svar	Stock Variance	1885–2005	−0.76		−27.14	−2.33			−2.44	+0.01		−1.32
d/e	Dividend Payout Ratio	1872–2005	−0.75		−4.33	−0.31			−4.99	−0.18		−1.24
lty	Long Term Yield	1919–2005	−0.63		−7.72	−0.47			−12.57	−0.76		−0.94
tms	Term Spread	1920–2005	0.16		−2.42	−0.07			−2.96	−0.03		0.89
tbl	T-Bill Rate	1920–2005	0.34		−3.37	−0.14			−4.90	−0.18		0.15
dfr	Default Return Spread	1926–2005	0.40		−2.16	−0.03			−2.82	−0.02		0.32
d/p	Dividend Price Ratio	1872–2005	0.49		−2.06	−0.11			−3.69	−0.09		1.67
d/y	Dividend Yield	1872–2005	0.91		−1.93	−0.10			−6.68	−0.31		**2.71***
ltr	Long Term Return	1926–2005	0.99		−11.79	−0.76			−18.38	−1.18		0.92
e/p	Earning Price Ratio	1872–2005	1.08		−1.78	−0.08			−1.10	+0.11		**3.20***
Full Sample, Significant IS												
b/m	Book to Market	1921–2005	**3.20***	1.13	−1.72	−0.01	42 (67)	−7.29	−12.71	−0.77	40 (61)	**4.14***
i/k	Invstmnt Capital Ratio	1947–2005	**6.63****	−0.25	−1.77	+0.07	47 (77)		same			same
ntis	Net Equity Expansion	1927–2005	**8.15*****	−4.21	−5.07	−0.26	57 (78)	0.96	−6.79	−0.32	53 (72)	same
eqis	Pct Equity Issuing	1927–2005	**9.15*****	2.81	**2.04****	+0.30	72 (85)	3.64	−1.00	+0.12	66 (77)	same
all	Kitchen Sink	1927–2005	**13.81****	2.62	−139.03	−5.97	- (-)	−20.91	−176.18	−6.19	- (-)	same
Full Sample, No IS Equivalent (caya, ms) or Ex-Post Information (cayp)												
cayp	Cnsmptn, Wlth, Incme	1945–2005	**15.72*****	20.70	**16.78*****	+1.61	- (-)		same			same
caya	Cnsmptn, Wlth, Incme	1945–2005	-	-	−4.33	−0.14	- (-)		same			same
ms	Model Selection	1927–2005	-	-	−22.50	−1.69	- (-)	-	−23.71	−1.79	- (-)	same
1927-2005 Sample, Significant IS												Full Sample
d/y	Dividend Yield	1927–2005	**2.71***					−0.35	−6.44	−0.30	30 (71)	0.91
e/p	Earning Price Ratio	1927–2005	**3.20***					−0.94	−3.15	−0.05	39 (64)	1.08
b/m	Book to Market	1927–2005	**4.14***					−8.65	−19.46	−1.26	45 (64)	**3.20***

Table 2: Forecasts at 5-year Frequency

This table is identical to Table 1, except that we predict overlapping 5-yearly equity premia, rather than annual equity premia.

			Full Sample									1927–2005 Sample
				Forecasts begin 20 years after sample				Forecasts begin 1965				
			IS	IS for	OOS			IS for	OOS			IS
	Variable	Data	$\overline{R}^2$	OOS $\overline{R}^2$	$\overline{R}^2$	ΔRMSE	Power	OOS $\overline{R}^2$	$\overline{R}^2$	ΔRMSE	Power	$\overline{R}^2$
Full Sample, Not Significant IS												
ltr	Long Term Return	1926–2005	−1.36		−7.40	−1.10			−18.92	−2.72		−1.39
dfr	Default Return Spread	1926–2005	−1.36		−5.71	−0.77			−4.01	−0.25		−1.36
infl	Inflation	1919–2005	−1.21		−11.25	−1.70			−7.34	−0.85		−1.21
lty	Long Term Yield	1919–2005	−0.15		−122.13	−17.41			−72.47	−10.96		−0.30
svar	Stock Variance	1885–2005	0.33		−79.33	−13.31			−2.37	+0.03		−0.84
d/e	Dividend Payout Ratio	1872–2005	0.66		−4.87	−0.76			−0.64	+0.31		1.64
dfy	Default Yield Spread	1919–2005	3.54		−59.33	−9.18			**4.97***	+1.38		0.94
tbl	T-Bill Rate	1920–2005	3.83		−17.66	−2.78			−30.19	−4.70		4.91
d/y	Dividend Yield	1872–2005	6.04		−4.45	−0.68			−16.84	−2.19		**14.99***
e/p	Earning Price Ratio	1872–2005	6.24		−1.04	−0.03			−3.03	−0.07		**14.96***
tms	Term Spread	1920–2005	7.84		−26.52	−4.24			**10.46****	+2.44		**12.47***
eqis	Pct Equity Issuing	1927–2005	9.50		−2.35	−0.11			−5.75	−0.56		same
b/m	Book to Market	1921–2005	10.78		−13.06	−2.03			−46.34	−7.17		13.93
Full Sample, Significant IS												
ntis	Net Equity Expansion	1927–2005	**6.59***	−8.28	−3.46	−0.32	21 (70)	1.49	−13.77	−1.92	21 (67)	same
d/p	Dividend Price Ratio	1872–2005	**10.24***	14.35	−1.19*	−0.06	21 (69)	8.30	−26.09	−3.54	20 (51)	**21.24****
i/k	Invstmnt Capital Ratio	1947–2005	**33.99*****	27.42	**12.99****	+3.39	22 (78)		same			same
all	Kitchen Sink	1927–2005	**41.48*****	43.29	−499.83	−45.47	- (-)	19.75	−442.08	−34.19	- (-)	same
Full Sample, No IS Equivalent (caya, ms) or Ex-Post Information (cayp)												
cayp	Cnsmptn, Wlth, Incme	1945–2005	**36.05*****	63.11	**30.35*****	+7.50	- (-)		same			same
caya	Cnsmptn, Wlth, Incme	1945–2005	-	-	**9.10*****	+2.50	- (-)		same			same
ms	Model Selection	1927–2005	-	-	−14465.67	−408.06	- (-)	-	−122.89	−18.03	- (-)	same
1927-2005 Sample, Significant IS												Full Sample
tms	Term Spread	1927–2005	**12.47***					23.24	**12.59****	+2.77	11 (65)	7.84
e/p	Earning Price Ratio	1927–2005	**14.96***					−4.04	−15.33	−2.18	28 (65)	6.24
d/y	Dividend Yield	1927–2005	**14.99***					6.16	−9.47	−1.19	22 (72)	6.04
d/p	Dividend Price Ratio	1927–2005	**21.24****					4.28	−12.69	−1.74	29 (61)	**10.24***

Table 3: Forecasts at Monthly Frequency using Campbell and Thompson (2005) procedure

Refer to Table 1 for basic explanations. This table presents statistics on forecast errors in-sample (IS) and out-of-sample (OOS) for equity premium forecasts at the monthly frequency (both in the forecasting equation and forecast). Variables are explained in Section 2. The data period is December 1927 to December 2004, except for **csp** (May 1937 to December 2002) and **cay**3 (December 1951 to December 2004). Critical values of all statistics are obtained empirically from bootstrapped distributions, except for **cay**3 model where they are obtained from McCracken (2004). The resulting significance levels at 90%, 95%, and 99% are denoted by one, two, and three stars, respectively. They are two-sided for IS model significance, and one-sided for OOS superior model performance. The first data column is the IS $\overline{R}^2$ when returns are logged, as they are in our other tables. The remaining columns are based on predicting simple returns for correspondence with Campbell and Thompson (2005). Certainty Equivalence (CEV) gains are based on the utility of an optimizer with a risk-aversion coefficient of $\gamma = 3$ who trades based on unconditional forecast and conditional forecast. Equity positions are winsorized at 150% ($w = w_{\max}$). At this risk-aversion, the base CEV are 82bp for a market-timer based on the unconditional forecast, 79bp for the market, and 40bp for the risk-free rate. "T" means "truncated" to avoid a negative equity premium prediction. "U" means unconditional, that is, to avoid a forecast that is based on a coefficient that is inverse to what the theory predicts. A superscript h denotes high trading turnover of about 10%/month more than the trading strategy based on unconditional forecasts.

		Log Returns	Simple Returns									
			IS		OOS	Campbell and Thompson (2005) OOS						
	Variable	IS $\overline{R}^2$	$\overline{R}^2$	$\overline{R}^2$ T	$\overline{R}^2$	Frcst= T	Frcst= U	$\overline{R}^2$ TU	ΔRMSE TU	$w = w_{\max}$	ΔCEV	Fig
d/e	Dividend Payout Ratio	0.02	−0.10	−0.10	−0.70	0.0	7.9	−0.69	−0.0114	57.7	-0.01	
svar	Stock Variance	−0.09	−0.07	−0.07	−0.79	0.0	0.0	−0.79	−0.0134	35.4	-0.04	
dfr	Default Return Spread	−0.02	−0.07	−0.08	−0.37	0.0	20.9	−0.29	−0.0030	44.9	0.01	
lty	Long Term Yield	−0.03	0.02	0.02	−0.80	34.1	0.0	**0.26****	+0.0085	19.5	0.06	
ltr	Long Term Return	0.04	0.07	0.08	−0.63	3.0	38.2	**0.11****	+0.0053	51.2^h	0.06	
infl	Inflation	−0.01	0.14	−0.05	**0.01***	1.3	0.0	**0.07****	+0.0045	43.5^h	0.04	
tms	Term Spread	0.12	0.18	0.20	**0.22****	3.7	0.0	**0.21****	+0.0073	59.3	0.14	F3.G
tbl	T-Bill Rate	0.10	**0.20***	0.15	−0.08*	23.1	0.0	**0.25****	+0.0081	16.4	0.10	F3.F
dfy	Default Yield Spread	−0.06	**0.28***	0.28	−0.56	4.0	0.0	−0.49	−0.0071	27.3	-0.08	
d/p	Dividend Price Ratio	0.12	**0.33***	0.29	−0.30	32.3	0.0	**0.17***	+0.0066	16.1	-0.10	F3.E
d/y	Dividend Yield	**0.22***	**0.47****	0.45	−1.12	54.2	0.0	−0.04*	+0.0023	16.4	-0.14	
e/p	Earning Price Ratio	**0.51****	**0.54****	0.45	−1.04	18.1	0.0	−1.03	−0.0183	34.4	-0.04	
eqis	Pct Equity Issuing	**0.82*****	**0.80*****	0.59	**0.14****	6.7	0.0	**0.30*****	+0.0093	55.8	0.14	F3.D
b/m	Book to Market	**0.45****	**0.81*****	0.88	−3.28	44.3	0.0	−2.23	−0.0432	31.3	-0.22	
e^{10}**/p**	Earning(10Y) Price Ratio	**0.46****	**0.86*****	0.96	−2.21	52.4	0.0	−0.48	−0.0071	15.4	-0.13	
csp	Cross-Sectional Prem	**0.92*****	**0.99*****	0.93	−0.94	44.7	0.0	**0.15****	+0.0072	13.5	0.06	F3.B
ntis	Net Equity Expansion	**0.94*****	**1.02*****	0.88	−0.16	0.4	0.0	−0.16	−0.0003	57.4	0.02	F3.C
cay3	Cnsmptn, Wlth, Incme	**1.88*****	**1.87*****	1.57	−2.05	44.7	0.0	−0.34*	+0.0088	13.2	0.06	F3.A

Table 4: Significant Forecasts Using Various d/p, e/p, and d/e Ratios

Refer to Table 1 for basic explanations. The table reports only those combinations of **d/p e/p** and **d/e** that were found to predict equity premia significantly in-sample. This table presents statistics on forecast errors in-sample (IS) and out-of-sample (OOS) for excess stock return forecasts at various frequencies. Variables are explained in Section 2. All ΔRMSE numbers are in percent per frequency corresponding to the column entitled 'Freq'. The 'Freq' column also gives the first year of forecast. A star next to OOS-$\overline{R}^2$ is based on the MSE-F-statistic by McCracken (2004), which tests for equal MSE of the unconditional forecast and the conditional forecast. One-sided critical values of MSE statistics are obtained empirically from bootstrapped distributions. Significance levels at 90%, 95%, and 99% are denoted by one, two, and three stars, respectively.

				IS	OOS	
	Variable	Data	Freq	$\overline{R}^2$	$\overline{R}^2$	ΔRMSE
e/p	Earning(1Y) Price Ratio	1927–2005	M 1965–	**0.54****	−1.20	−0.02
e^5/p	Earning(5Y) Price Ratio	1927–2005	M 1965–	**0.32***	−0.60	−0.01
e^{10}/p	Earning(10Y) Price Ratio	1927–2005	M 1965–	**0.49****	−0.83	−0.01
e^3/p	Earning(3Y) Price Ratio	1882–2005	A 1902–	**2.53****	−1.05*	−0.01
e^5/p	Earning(5Y) Price Ratio	1882–2005	A 1902–	**2.88****	−0.52*	+0.04
e^{10}/p	Earning(10Y) Price Ratio	1882–2005	A 1902–	**4.89****	**2.12****	+0.30
d^3/p	Dividend(3Y) Price Ratio	1882–2005	A 1902–	**1.85***	−1.53	−0.05
d^5/p	Dividend(5Y) Price Ratio	1882–2005	A 1902–	**2.48***	−0.54*	+0.04
d^{10}/p	Dividend(10Y) Price Ratio	1882–2005	A 1902–	**2.11***	−1.07*	−0.01
e^3/p	Earning(3Y) Price Ratio	1882–2005	A 1965–	**2.53****	−3.41	−0.06
e^5/p	Earning(5Y) Price Ratio	1882–2005	A 1965–	**2.88****	−5.01	−0.19
e^{10}/p	Earning(10Y) Price Ratio	1882–2005	A 1965–	**4.89****	−11.45	−0.66
d^3/p	Dividend(3Y) Price Ratio	1882–2005	A 1965–	**1.85***	−6.55	−0.30
d^5/p	Dividend(5Y) Price Ratio	1882–2005	A 1965–	**2.48***	−8.79	−0.47
d^{10}/p	Dividend(10Y) Price Ratio	1882–2005	A 1965–	**2.11***	−8.32	−0.43
e^3/p	Earning(3Y) Price Ratio	1882–2005	5Y 1902–	**11.35***	**3.46****	+0.89
e^5/p	Earning(5Y) Price Ratio	1882–2005	5Y 1902–	**16.16****	**4.76****	+1.16
e^{10}/p	Earning(10Y) Price Ratio	1882–2005	5Y 1902–	**16.47****	−2.85*	−0.37
d/p	Dividend(1Y) Price Ratio	1882–2005	5Y 1902–	**12.30***	−0.66*	+0.06
d^3/p	Dividend(3Y) Price Ratio	1882–2005	5Y 1902–	**13.11***	−2.02*	−0.21
d^5/p	Dividend(5Y) Price Ratio	1882–2005	5Y 1902–	**13.75***	−3.85*	−0.57
e^3/p	Earning(3Y) Price Ratio	1882–2005	5Y 1965–	**11.35***	−12.55	−1.56
e^5/p	Earning(5Y) Price Ratio	1882–2005	5Y 1965–	**16.16****	−21.16	−2.85
e^{10}/p	Earning(10Y) Price Ratio	1882–2005	5Y 1965–	**16.47****	−25.65	−3.51
d/p	Dividend(1Y) Price Ratio	1882–2005	5Y 1965–	**12.30***	−29.33	−4.03
d^3/p	Dividend(3Y) Price Ratio	1882–2005	5Y 1965–	**13.11***	−28.11	−3.86
d^5/p	Dividend(5Y) Price Ratio	1882–2005	5Y 1965–	**13.75***	−30.71	−4.23

Table 5: Forecasts at Monthly Frequency with Alternative Procedures and Total Returns

Refer to Table 1 for basic explanations.Columns under the heading 'OLS' are unadjusted betas, columns under the heading 'Stambaugh' correct for betas following Stambaugh (1999), and columns under the heading 'Lewellen' correct for betas following Lewellen (2004). ρ under the column OLS gives the autoregressive coefficient of the variable over the entire sample period (the variables are sorted in descending order of ρ).

			OLS				Stambaugh			Lewellen		
			IS		OOS		IS	OOS		IS	OOS	
	Variable	Data	ρ	$\overline{R}^2$	$\overline{R}^2$	Power	$\overline{R}^2$	$\overline{R}^2$	Power	$\overline{R}^2$	$\overline{R}^2$	Power
d/e	Dividend Payout Ratio	192701–200512	0.9989	0.01	−2.02	15 **(70)**	0.01	−2.11	15 **(69)**	0.01	−2.05	15 **(69)**
lty	Long Term Yield	192701–200512	0.9963	−0.01	−1.15	9 **(68)**	−0.01	−1.71	9 **(68)**	−0.01	−1.05	10 **(67)**
d/y	Dividend Yield	192701–200512	0.9929	**0.25***	−0.40	33 **(71)**	**0.25***	−0.36	33 **(71)**	**0.25***	−0.26	32 **(72)**
d/p	Dividend Price Ratio	192701–200512	0.9927	0.15	−0.15	29 **(56)**	0.05	−0.31	26 **(69)**	−0.15**	−1.03	3 **(3)**
tbl	T-Bill Rate	192701–200512	0.9922	0.11	−0.18	19 **(69)**	0.11	−0.33	20 **(69)**	0.11	−0.27	20 **(68)**
e/p	Earning Price Ratio	192701–200512	0.9879	**0.54****	−1.21	56 **(64)**	**0.48****	−0.54	59 **(73)**	**0.02*****	−0.01**	41 **(41)**
b/m	Book to Market	192701–200512	0.9843	**0.40****	−2.45	48 **(65)**	**0.36****	−1.61	48 **(71)**	−0.14**	−0.31	19 **(19)**
csp	Cross-Sectional Prem	193705–200212	0.9788	**0.92*****	**0.70*****	65 **(80)**	**0.92*****	**0.70*****	65 **(80)**	**0.91*****	**0.71*****	65 **(81)**
dfy	Default Yield Spread	192701–200512	0.9763	−0.07	−0.14	9 **(59)**	−0.07	−0.33	8 **(59)**	−0.15	−0.71	5 **(43)**
ntis	Net Equity Expansion	192701–200512	0.9680	**0.75*****	−0.28	59 **(76)**	**0.75*****	−0.29	59 **(76)**	**0.74*****	−0.38	58 **(74)**
tms	Term Spread	192701–200512	0.9566	0.07	**0.09***	21 **(66)**	0.07	**0.07***	20 **(65)**	0.07	**0.06***	20 **(65)**
svar	Stock Variance	192701–200512	0.6008	−0.08	−0.34	7 **(53)**	−0.08	−0.34	7 **(53)**	−1.66**	−0.63	7 **(7)**
infl	Inflation	192701–200512	0.5513	−0.00	−0.07	14 **(62)**	−0.00	−0.07	14 **(62)**	−0.03	−0.13	15 **(56)**
ltr	Long Term Return	192701–200512	0.0532	0.04	−0.49	18 **(62)**	0.04	−0.48	18 **(62)**	−1.55***	−6.41	12 **(12)**
dfr	Default Return Spread	192701–200512	-0.1996	−0.02	−0.30	12 **(61)**	−0.02	−0.30	12 **(61)**	−1.32*	−2.64	10 **(38)**

Table 6: Encompassing Tests

This table presents statistics on encompassing tests for excess stock return forecasts at various frequencies. Variables are explained in Section 2. All numbers are in percent per frequency corresponding to the panel. λ gives the ex-post weight on the conditional forecast for the optimal forecast that minimizes the MSE. ENC is the test statistic proposed by Clark and McCracken (2001) for a test of forecast encompassing. One-sided critical values of ENC statistic are obtained empirically from bootstrapped distributions, except for **caya**, **cayp**, and **all** models where they are obtained from Clark and McCracken (2001). Critical values for **ms** model are not calculated. **cayp** uses ex-post information. ΔRMSE* is the RMSE difference between the unconditional forecast and the optimal forecast for the same sample/forecast period. ΔRMSE^{*r} is the RMSE difference between the unconditional forecast and the optimal forecast for the same sample/forecast period using rolling estimates of λ. Significance levels at 90%, 95%, and 99% are denoted by one, two, and three stars, respectively.

Panel A: Annual Data

		Estimation: OOS Forecast:	All Data After 20 years					All Data After 1965				After 1927 After 1965				
		Data	$\overline{R}^2$	λ	ENC	ΔRMSE*	ΔRMSE^{*r}	λ	ENC	ΔRMSE*	ΔRMSE^{*r}	$\overline{R}^2$	λ	ENC	ΔRMSE*	ΔRMSE^{*r}
d/p	Dividend Price Ratio	1872–2005	0.49	0.21	0.48	+0.0084	−0.2583	0.40	**0.87***	+0.0664	−0.4989	1.67	0.54	**2.19****	+0.2297	−0.3539
d/y	Dividend Yield	1872–2005	0.91	0.38	1.94	+0.0614	−0.5713	0.30	**1.24***	+0.0749	−0.5389	**2.71***	0.41	**3.24****	+0.2662	−0.2858
e/p	Earning Price Ratio	1872–2005	1.08	0.22	0.40	+0.0074	−0.2266	0.66	**1.21****	+0.1508	−0.4845	**3.20***	0.48	**2.51****	+0.2346	−0.4049
d/e	Dividend Payout Ratio	1872–2005	−0.75	-1.73	−1.46	+0.2135	+0.0960	-8.46	−0.45	+0.7545	+0.2858	−1.24	-4.57	−1.25	+1.2308	+0.7796
svar	Stock Variance	1885–2005	−0.76	-0.42	−4.74	+0.2387	−0.6475	2.07	0.03	+0.0134	−0.5937	−1.32	-16.73	−0.18	+0.5906	+0.4490
b/m	Book to Market	1921–2005	**3.20***	0.49	**4.16****	+0.2532	−0.0575	0.20	**1.27***	+0.0559	−0.7885	**4.14***	0.18	**1.67***	+0.0689	−0.4821
ntis	Net Equity Expansion	1927–2005	**8.15*****	0.31	1.46	+0.0619	−0.2708	0.31	**1.30***	+0.0805	−0.9310	**8.15*****	0.31	**1.30***	+0.0805	−0.9310
eqis	Pct Equity Issuing	1927–2005	**9.15*****	0.67	**4.45****	+0.3917	−0.0564	0.56	**3.12****	+0.3342	−0.7106	**9.15*****	0.56	**3.12****	+0.3342	−0.7106
tbl	T-Bill Rate	1920–2005	0.34	0.39	**2.14***	+0.1031	−1.2425	0.41	**2.16****	+0.1790	−1.3058	0.15	0.33	**2.72****	+0.1863	−0.5619
lty	Long Term Yield	1919–2005	−0.63	0.29	**2.67***	+0.0971	−0.7012	0.28	**2.39****	+0.1447	−0.9358	−0.94	0.25	**2.39****	+0.1317	−0.5682
ltr	Long Term Return	1926–2005	0.99	0.31	**4.55****	+0.2077	−0.1412	0.24	**2.45****	+0.1300	−8.4290	0.92	0.25	**2.44****	+0.1348	−8.7284
tms	Term Spread	1920–2005	0.16	0.38	0.93	+0.0433	−1.0292	0.47	**1.07***	+0.0977	−0.8750	0.89	0.50	**1.95****	+0.1880	−0.5375
dfy	Default Yield Spread	1919–2005	−1.18	-2.62	−0.48	+0.1503	−0.9718	-10.65	−0.30	+0.6395	+0.4999	−1.31	-11.91	−0.24	+0.5677	+0.4496
dfr	Default Return Spread	1926–2005	0.40	0.44	0.87	+0.0501	−0.3698	0.47	0.78	+0.0710	−0.3808	0.32	0.48	0.74	+0.0692	−0.3877
infl	Inflation	1919–2005	−1.00	-2.46	−0.68	+0.2019	−0.4520	-1.48	−0.15	+0.0429	−15.1368	−0.99	-3.12	−0.86	+0.5541	−0.4697
i/k	Invstmnt Capital Ratio	1947–2005	**6.63****	0.53	**3.01****	+0.3330	−0.1950	0.53	**3.01****	+0.3330	−0.1950	**6.63****	0.53	**3.01****	+0.3330	−0.1950
cayp	Cnsmptn, Wlth, Incme	1945–2005	**15.72*****	1.34	**7.62*****	+1.7225	+0.3315	1.34	**7.62*****	+1.7225	+0.3315	**15.72*****	1.34	**7.62*****	+1.7225	+0.3315
all	Kitchen Sink	1927–2005	**13.81****	0.13	4.86	+0.1607	+0.0160	-0.07	−1.26	+0.0342	−0.4666	**13.81****	-0.07	−1.26	+0.0342	−0.4666
caya	Cnsmptn, Wlth, Incme	1945–2005	-	0.45	**3.39****	+0.3117	−0.3185	0.45	**3.39****	+0.3117	−0.3185	-	0.45	**3.39****	+0.3117	−0.3185
ms	Model Selection	1927–2005	-	0.24	4.82	+0.1870	+0.0739	0.07	0.59	+0.0094	−1.1268	-	0.07	0.59	+0.0094	−1.1268

Panel B: Monthly Data

		OOS Forecast:	After 194701					After 196501			
		Data	$\overline{R}^2$	λ	ENC	ΔRMSE*	ΔRMSE*^{r}	λ	ENC	ΔRMSE*	ΔRMSE*^{r}
d/p	Dividend Price Ratio	192701–200512	0.15	0.53	**4.14****	+0.0065	−0.0134	0.53	**2.67****	+0.0063	−0.0109
d/y	Dividend Yield	192701–200512	**0.25***	0.43	**6.53*****	+0.0083	−0.0115	0.45	**3.90****	+0.0078	−0.0084
e/p	Earning Price Ratio	192701–200512	**0.54****	0.35	**9.27*****	+0.0097	−0.0135	0.28	**3.08****	+0.0039	−0.0172
d/e	Dividend Payout Ratio	192701–200512	0.01	-0.02	−0.22	+0.0000	−0.0146	-1.12	−3.01	+0.0152	+0.0003
svar	Stock Variance	192701–200512	−0.08	-12.30	−0.47	+0.0172	+0.0046	-12.93	−0.32	+0.0184	+0.0060
csp	Cross-Sectional Prem	193705–200212	**0.92*****	0.38	**6.21*****	+0.0093	−0.0138	0.82	**5.50*****	+0.0219	−0.0007
b/m	Book to Market	192701–200512	**0.40****	0.18	**3.04****	+0.0016	−0.0416	0.07	0.89	+0.0003	−0.0260
ntis	Net Equity Expansion	192701–200512	**0.75*****	0.60	**4.28****	+0.0075	−0.0055	0.47	**2.77****	+0.0058	−0.0180
tbl	T-Bill Rate	192701–200512	0.11	0.50	**5.47****	+0.0081	−0.0222	0.51	**4.86*****	+0.0110	−0.0218
lty	Long Term Yield	192701–200512	−0.01	0.35	**7.57*****	+0.0079	−0.0084	0.35	**5.47*****	+0.0086	−0.0161
ltr	Long Term Return	192701–200512	0.04	-0.15	−0.77	+0.0003	−0.0129	0.30	**1.02***	+0.0014	−0.0234
tms	Term Spread	192701–200512	0.07	0.68	**2.51****	+0.0050	−0.0311	0.73	**2.37****	+0.0076	−0.0538
dfy	Default Yield Spread	192701–200512	−0.07	-1.04	−0.27	+0.0008	−0.0070	2.15	0.20	+0.0019	−0.0197
dfr	Default Return Spread	192701–200512	−0.02	-0.85	−0.72	+0.0018	−0.0134	-0.03	−0.01	+0.0000	−0.0221
infl	Inflation	192701–200512	−0.00	1.01	0.69	+0.0021	−0.0114	1.19	0.58	+0.0030	−0.0541
all	Kitchen Sink	192701–200512	**1.98*****	0.05	4.39	+0.0008	−0.0150	0.14	**5.88****	+0.0040	−0.0366
ms	Model Selection	192701–200512	–	0.09	1.51	+0.0004	−0.0232	0.14	1.39	+0.0009	−0.0245

References

Ang, A., and G. Bekaert, 2003, "Stock Return Predictability: Is it There?," Working Paper, Columbia University.

Avramov, D., 2002, "Stock Return Predictability and Model Uncertainty," *Journal of Financial Economics*, 64(3), 423–458.

Baker, M., R. Taliaferro, and J. Wurgler, 2004, "Pseudo Market Timing and Predictive Regressions," NBER Working paper No. 10823.

Baker, M., and J. Wurgler, 2000, "The Equity Share in New Issues and Aggregate Stock Returns," *Journal of Finance*, 55(5), 2219–2257.

Ball, R., 1978, "Anomalies in Relationship Between Securities' Yields and Yield-Surrogates," *Journal of Financial Economics*, 6(2/3), 103–126.

Bansal, R., G. Tauchen, and H. Zhou, 2004, "Regime-Shifts, Risk Premiums in the Term Structure, and the Business Cycle," *Journal of Business and Economic Statistics*, 22(4), 396–409.

Bossaerts, P., and P. Hillion, 1999, "Implementing Statistical Criteria to Select Return Forecasting Models: What Do We Learn?," *Review of Financial Studies*, 12(2), 405–428.

Boudoukh, J., R. Michaely, M. P. Richardson, and M. R. Roberts, 2005, "On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing," Working Paper, forthcoming in *Journal of Finance*.

Boudoukh, J., M. Richardson, and R. F. Whitelaw, 2005, "The Myth of Long-Horizon Predictability," NBER Working Paper No. 11841.

Breen, W., L. R. Glosten, and R. Jagannathan, 1989, "Economic Significance of Predictable Variations in Stock Index Returns," *Journal of Finance*, 64(5), 1177–1189.

Brennan, M. J., and Y. Xia, 2004, "Persistence, Predictability, and Portfolio Planning," Working Paper, UCLA and Wharton.

Butler, A. W., G. Grullon, and J. Weston, 2005, "Can Managers Forecast Aggregate Market Returns?," *Journal of Finance*, 60(2), 963–986.

Campbell, J. Y., 1987, "Stock returns and the Term Structure," *Journal of Financial Economics*, 18(2), 373–399.

Campbell, J. Y., and R. J. Shiller, 1988a, "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43(3), 661–676.

——— , 1988b, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1(3), 195–227.

——— , 1998, "Valuation Ratios and the Long-Run Stock Market Outlook," *Journal of Portfolio Management*, 24(2), 11–26.

Campbell, J. Y., and S. B. Thompson, 2005, "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?," Working Paper, Harvard University.

Campbell, J. Y., and L. M. Viceira, 2002, *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors.* Oxford University Press, Oxford, UK.

Campbell, J. Y., and T. Vuolteenaho, 2004, "Inflation Illusion and Stock Prices," *American Economic Review*, 94(2), 19–23.

Campbell, J. Y., and M. Yogo, 2006, "Efficient Tests of Stock Return Predictability," Working Paper, forthcoming in *Journal of Financial Economics*.

Clark, T. E., and M. W. McCracken, 2001, "Tests of Forecast Accuracy and Encompassing for Nested Models," *Journal of Econometrics*, 105, 85–110.

Cochrane, J. H., 1991, "Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations," *Journal of Finance*, 46(1), 209–237.

——— , 1997, "Where is the Market Going? Uncertain Facts and Novel Theories," *Federal Reserve Bank of Chicago - Economic Perspectives*, 21(6), 3–37.

——— , 2005, "The Dog That Did Not Bark: A Defense of Return Predictability," Working Paper, University of Chicago.

Diebold, F. X., and R. S. Mariano, 1995, "Comparing Predictive Accuracy," *Journal of Business & Economic Statistics*, 13(3), 253–263.

Dow, C. H., 1920, "Scientific Stock Speculation," The Magazine of Wall Street.

Fama, E. F., 1981, "Stock Returns, Real Activity, Inflation, and Money," *American Economic Review*, 71(4), 545–565.

Fama, E. F., and K. R. French, 1988, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics*, 22(1), 3–25.

——— , 1989, "Business Conditions and Expected Returns on Stocks and Bonds," *Journal of Financial Economics*, 25(1), 23–49.

Fama, E. F., and G. W. Schwert, 1977, "Asset Returns and Inflation," *Journal of Financial Economics*, 5(2), 115–146.

Ferson, W. E., S. Sarkissian, and T. T. S. Simin, 2003, "Spurious Regressions in Financial Economics," *Journal of Finance*, 58(4), 1393–1413.

Fisher, K. L., and M. Statman, 2005, "Market Timing at Home and Abroad," Working Paper, Santa Clara University.

Goetzmann, W. N., and P. Jorion, 1993, "Testing the Predictive Power of Dividend Yields," *Journal of Finance*, 48(2), 663–679.

Goyal, A., and I. Welch, 2003, "Predicting the Equity Premium with Dividend Ratios," *Management Science*, 49(5), 639–654.

Guo, H., 2006, "On the Out-of-Sample Predictability of Stock Market Returns," *Journal of Business*, 79(2), 645–670.

Harvey, D. I., S. J. Leybourne, and P. Newbold, 1997, "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13(2), 281–291.

Heaton, J., and D. J. Lucas, 2000, "Stock Prices and Fundamentals," in *NBER Macroeconomics Annual 1999*, ed. by B. S. Bernanke, and J. Rotemberg. MIT Press/NBER, Cambridge, MA, vol. 14, pp. 213–242.

Hodrick, R. J., 1992, "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5(3), 257–286.

Inoue, A., and L. Kilian, 2004, "In-Sample or Out-of-Sample Tests of Predictability: Which One Should We Use?," *Econometric Reviews*, 23(4), 371–402.

Jagannathan, R., E. R. McGrattan, and A. Scherbina, 2000, "The Declining U.S. Equity Premium," *Federal Reserve Bank of Minneapolis Quarterly Review*, 24(4), 3–19.

Keim, D., and A. Madhavan, 1997, "Transaction Costs and Investment Style: An Inter-Exchange Analysis of Institutional Equity Trades," *Journal of Financial Economics*, 46(3), 265–292.

Keim, D. B., and R. F. Stambaugh, 1986, "Predicting Returns in the Stock and Bond Markets," *Journal of Financial Economics*, 17(2), 357–390.

Kilian, L., 1999, "Exchange Rates and Monetary Fundamentals: What Do We Learn From Long–Horizon Regressions?," *Journal of Applied Econometrics*, 14(5), 491–510.

Kim, C.-J., J. C. Morley, and C. R. Nelson, 2005, "The Structural Break in the Equity Premium," *Journal of Business and Economic Statistics*, 23(2), 181–191.

Kothari, S., and J. Shanken, 1997, "Book-to-market, Dividend Yield, and Expected Market Returns: A Time-Series Analysis," *Journal of Financial Economics*, 44(2), 169–203.

Lamont, O., 1998, "Earnings and Expected Returns," *Journal of Finance*, 53(5), 1563–1587.

Lamoureux, C. G., and G. Zhou, 1996, "Temporary Components of Stock Returns: What Do the Data Tell Us?," *Review of Financial Studies*, 9(4), 1033–1059.

Lettau, M., and S. Ludvigson, 2001, "Consumption, Aggregate Wealth, and Expected Stock Returns," *Journal of Finance*, 56(3), 815–849.

——— , 2005, "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, 76(3), 583–626.

Lettau, M., and S. Van Nieuwerburgh, 2005, "Reconciling the Return Predictability Evidence," Working Paper, NYU.

Lewellen, J., 2004, "Predicting Returns with Financial Ratios," *Journal of Financial Economics*, 74(2), 209–235.

Lewellen, J., and J. Shanken, 2002, "Learning, Asset-Pricing Tests, and Market Efficiency," *Journal of Finance*, 57(3), 1113–1145.

Lintner, J., 1975, "Inflation and Security Returns," *Journal of Finance*, 30(2), 259–280.

Mark, N. C., 1995, "Exchange Rates and Fundamentals: Evidence on Long–Horizon Predictability," *American Economic Review*, 85(1), 201–218.

McCracken, M. W., 2004, "Asymptotics for Out-of-Sample Tests of Causality," Working Paper, University of Missouri-Columbia.

Menzly, L., T. Santos, and P. Veronesi, 2004, "Understanding Predictability," *Journal of Political Economy*, 112(1), 1–47.

Nelson, C. R., and M. J. Kim, 1993, "Predictable Stock Returns: The Role of Small Sample Bias," *Journal of Finance*, 48(2), 641–661.

Pesaran, H. M., and A. Timmerman, 1995, "Predictability of Stock Returns: Robustness and Economic Significance," *Journal of Finance*, 50(4), 1201–1228.

Polk, C., S. Thompson, and T. Vuolteenaho, 2006, "Cross-Sectional Forecasts of the Equity Premium," Working Paper, forthcoming in *Journal of Financial Economics*.

Pontiff, J., and L. D. Schall, 1998, "Book-to-Market Ratios as Predictors of Market Returns," *Journal of Financial Economics*, 49(2), 141–160.

Rapach, D. E., and M. E. Wohar, 2006, "In-Sample vs. Out-of-Sample Tests of Stock Return Predictability in the Context of Data Mining," *Journal of Empirical Finance*, 13(2), 231–247.

Rissanen, J., 1986, "Order Estimation by Accumulated Prediction Errors," *Journal of Applied Probability*, 23A, 55–61.

Rozeff, M. S., 1984, "Dividend Yields Are Equity Risk Premiums," *Journal of Portfolio Management*, 11(1), 68–75.

Shiller, R. J., 1984, "Stock Prices and Social Dynamics," *Brookings Papers on Economic Activity*, 2, 457–498.

Stambaugh, R. F., 1999, "Predictive Regressions," *Journal of Financial Economics*, 54(3), 375–421.

Torous, W., and R. Valkanov, 2000, "Boundaries of Predictability: Noisy Predictive Regressions," Working Paper, UCLA.

Valkanov, R., 2003, "Long-Horizon Regressions: Theoretical Results and Applications," *Journal of Financial Economics*, 68(2), 201–232.

Xia, Y., 2001, "Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation," *Journal of Finance*, 56(1), 205–246.

Equity Risk Premium: 2006 Update

by Roger J. Grabowski, ASA

Are you aware of recent research questioning the use of those realized equity premiums as an estimate of the equity risk premium (ERP)? [1,2] Or do you simply choose to ignore the research?

ERP is a forward-looking concept. ERP is an expectation as of the valuation date for which no "market quotes" are observable. While you can observe premiums realized over time by referring to historical data, such calculated premiums serve only as estimates for the *expected* ERP. If we are to truly mimic the market, then our goal should be to estimate the true expected ERP as of the valuation date. To do that you need to look beyond the realized premiums.

While there is no one universally accepted standard for estimating ERP, you need to be aware of recent research and not blindly continue using the historical realized equity premiums reported in the *SBBI Yearbook*. The methods used can be broadly categorized into one of two approaches: the Realized Return or *ex post* approach and the Forward-looking or *ex ante* approach.

Ex Post Approach

The realized return approach employs the premium that investors have, on the average, realized over some historical holding period (historical realized premium). The underlying theory is that the past provides an indicator of how the market will behave in the future, and investors' expectations are influenced by the historical performance of the market. If periodic (say, monthly) returns are serially independent (i.e., not correlated) and if expected returns are stable through time, the arithmetic average of historical returns provides an unbiased estimate of expected future returns. A more indirect justification for use of the historical approach is the contention that, for whatever reason, securities in the past have been priced in such a way as to earn the returns observed. By using the historical realized premium in applying the income approach to valuation (i.e., in the discounted cash flow valuation method), one may, to some extent, replicate this level of pricing.

Academics often formulate their research in terms of the equity risk premium relative to Treasury bills. But the variability of Treasury bill returns is such that one can hardly consider them riskless. Further we are generally valuing closely held businesses. Those investments are generally thought of as long-term and long-term government bonds are the benchmark security we use in developing discount rates. Therefore, in this article we have reported the research results in terms of the premium over long-term government bonds in calculating the historical realized premium.[3]

In applying the realized return method, the analyst selects the number of years of historical return data to include in the average. One school of thought holds that the future is best estimated using a very long horizon of past returns. Another school of thought holds that the future is best measured by the (relatively) recent past. These differences in opinion result in disagreement as to the number of years to include in the average.

While the *SBBI Yearbook*[4] contains summaries of returns on U.S. stocks and bonds derived from data accumulated by the Center for Research in Security Prices (CRSP) at the University of Chicago since 1926, good stock market data is available back to 1871, and less reliable data is available from various sources back to the end of the eighteenth century. Data for yields on government bonds is also available for these periods.[5] Exhibit 1 displays realized average annual premiums of

---

[1] Readers interested in more detailed information on the ERP issue are invited to attend the American Society of Appraisers' Center for Advanced Business Valuation Studies Cost of Capital course and to read Grabowski and King, Chapter 1, "Equity Risk Premium" in *The Handbook of Business Valuation and Intellectual Property Analysis*, (McGraw-Hill, 2004);"Equity Risk Premium: What Valuation Consultants Need to Know About Current Research" *Valuation Strategies* (Sept/Oct 2003); "Equity Risk Premium: What Valuation Consultants Need to Know About Current Research – 2005 Update" *Valuation Strategies* (Sept/Oct 2005); "Equity Risk Premium – What is the Current Evidence", *Business Valuation Review* (Fall 2005)

[2] The equity risk premium (ERP) (sometimes referred to as the market risk premium) is defined as the extra return (over the expected yield on government securities) that investors expect to receive from an investment in a diversified portfolio of common stocks. ERP = Rm - Rf where Rm is the expected return on a fully diversified portfolio of equity securities and Rf is the rate of return expected on an investment free of default risk.

[3] In applying the ERP in, say, the CAPM, one must use the return on a risk-free security with a term (maturity) consistent with the benchmark security used in developing the ERP. For example, this article measures ERP in terms of the premium over that of long-term government bonds. In CAPM, ke = Rf + (Beta × ERP). The Rf used as of the valuation date should be the yield on a long-term government bond because the data cited herein has been developed comparing equity returns to the income return (i.e., the yield promised at issue date) of long-term government bonds.

[4] *Stocks, Bonds, Bills and Inflation Valuation Edition 2006 Yearbook* (Ibbotson Associates, 2006)

[5] See Fisher and Lorie, "Rates of Return on Investments in Common Stocks," 37–1 *Journal of Business* (1964); Wilson and Jones, "A Comparison of Annual Stock Market Returns: 1871–1925 with 1926–1985," 60–2 *Journal of Business* 1 (1987); Schwert, "Indexes of Common Stock Returns from 1802 to 1987," 60–3 Journal of Business 239 (1990); Ibbotson and Brinson, *Global Investing* (McGraw-Hill, 1993); Wilson and Jones, "An Analysis of the S&P 500 Index and Cowles's Extensions: Price Indexes and Stock Returns, 1870–1999," 75–3 *Journal of Business* 505 (2002); Wright, "Measures of Stock Market Value and Returns for the US Nonfinancial Corporate Sector, 1900–2000," working paper, 2/1/02.; Goetzmann, Ibbotson and Peng, "A new historical database for the NYSE 1815 to 1925: Performance and Predictability", *Journal of Financial Markets* 4 (2001) 1–32; Dimson, Marsh and Staunton, *Triumph of the Optimists: 101Years of Global Investment Returns* (Princeton University Press, 2002) with annual updates of their Global Returns database for seventeen countries including the U.S. available at www.ibbotson.com.

Exhibit 1
Historical Realized Equity Risk Premiums: Stock Market Returns vs. Treasury Bonds (Income Returns)

Period	Arithmetic (%)	Geometric (%)
20 years (since 1986)	6.4	5.1
30 years (since 1976)	6.0	4.9
40 years (since 1966)	4.2	2.9
50 years (since 1956)	5.0	3.8
80 years (since 1926)	7.1	5.2
106 years (since 1900)	6.7	4.9
134 years (since 1872)	5.9	4.3
208 years (since 1798)	5.1	3.6

stock market returns (relative to the income return on long-term government bonds) for alternative periods through 2005.

The historical realized premium is measured by comparing the stock market returns realized during the period to the income return on bonds. While the stock market return is not known when investing at the beginning of the period, the rate of interest promised on a long-term government bond is known in terms of the yield to maturity. Therefore, analysts measure the stock market returns realized over the expected returns on bonds. An investor makes a decision to invest in the stock market today by comparing the expected return from that investment to the return on a benchmark security (in this case the long-term government bond) given the rate of return today on that benchmark security. The realized return approach is based on the expectation that history will repeat itself and such a premium return will again be realized (on the average) in the future.

Selection of the Observation Period

The historical realized premium derived from realized returns is sensitive to the period chosen for the average. For example, if one includes in the average only observed premiums in the immediate past period, that *ex post* premium may be the inverse of the *ex ante* estimate analysts are looking to develop. Almost all practitioners who use historical data focus on a longer-run view of historical returns. But selection of the period over which to measure those returns is key.

The selection of 1926 as a starting point is a happenstance of the arbitrary selection of that date by the founders of the CRSP database. The average calculated using 1926 return data as a beginning point may be too heavily influenced by the unusually low interest rates during the 1930s to mid-1950s. Some observers have suggested that the period, which includes the 1930s, 1940s, and the immediate post-World War II boom period may have exhibited an unusually high average

Exhibit 2
Historical Realized Returns: Relative Volatility of Stock Returns to Bond Returns

	Realized Equity Risk premiums over Treasury Bond Income Returns Nominal (i.e., without inflation removed)	
	1926–1957	1958–2005
Arithmetic averages (%)	9.5	5.4
Geometric average (%)	6.6	4.2
Standard Deviations		
Stock Market annual returns (%)	24.8	16.7
Long-term Treasury Bond		
Income Returns (%)	0.5	2.4
Total Returns (%)	4.9	11.0
Ratio of Equity to Bond		
Total Return Volatility	5.1	1.5

Source: Ibbotson Associates' data; calculations by author.

realized return premium. If we disaggregate the 80 years reported in the *SBBI Yearbook* into two sub-periods, the first covering the periods before and after the mid-1950s, we get the following comparative figures for stock and bond returns as shown in Exhibit 2.

The period since the mid-1950s has been characterized by a more stable stock market and a more volatile bond market compared to the earlier period. Interest rates have become more volatile in the later period.[6] The effect is amplified in the volatility of bond total returns.[7] This data indicates that the *relative* risk of stocks versus bonds is lower today which indicates that the equity risk premium is likely lower today. Thus, the historical arithmetic average realized premium reported in the *SBBI Yearbook* as measured from 1926 likely overstates expected returns as of 2006.

If the average expected return on stocks has changed through time, averages of realized returns using the longest available data become questionable. A short-run horizon may give a better estimate if changes in economic conditions have created a different expected return environment than that of more remote past periods. For example, why not use the average realized return over the past 20-year period? A drawback of using averages over shorter periods is that they are susceptible to large errors in measuring the true ERP due to high volatility of annual stock returns. Also, the average of the realized

[6] As reflected in Ibbotson Associates' Long-term Treasure Bond Income Return statistics.
[7] As reflected in Ibbotson Associates' Long-term Treasure Bond Total Returns which include the capital gains and losses associated with interest rate fluctuations.

premiums over the past 20 years may overstate today's expected returns due to the general downward movement of interest rates since 1981.

Even using long-term observations, the volatility of annual stock returns is high. For example, the standard deviation of the realized average return for the entire 80-year period 1926–2005 is approximately 20%. Even assuming that the 80-year average gives an unbiased estimate, a 95% confidence interval for the unobserved true ERP still spans a range of approximately 2.7% to 11.5%.

Which Average—Arithmetic or Geometric?

Realized return premiums measured using geometric (compound) averages are always less than those using the arithmetic average. The choice between which average to use remains a matter of disagreement among practitioners. The arithmetic average receives the most support in the literature,[8] other authors recommend a geometric average,[9] and still others support something in between.[10] The use of the arithmetic average relies on the assumption that (1) market returns are serially independent (not correlated) and (2) the distribution of market returns is stable (not time-varying). Under these assumptions, an arithmetic average gives an unbiased estimate of expected future returns. Empirical studies generally indicate a fairly low degree of serial correlation, supporting use of the arithmetic average. Moreover, the more observations, the more accurate the estimate will be.

But even if one agrees that stock returns are serially independent, the arithmetic average of one-year realized premiums may not be the best estimate of future premiums. Textbook models of stock returns (e.g., CAPM) are generally single period models that estimate returns over unspecified investment horizons. As the investment horizon increases, the arithmetic average of realized premiums decreases asymptotically to the geometric average of the entire realized premium series. As a result, some recommend using the mid-point of the arithmetic average of one-year realized premiums and the geometric average of the entire realized premium series as the best estimate of the future premiums when one is using historical realized premiums as the basis for their future ERP estimate.[11]

Expected ERP versus Realized Equity Premiums

Much has recently been written comparing the realized returns as reported in sources such as the *SBBI Yearbook* with the ERP that must have been expected by investors given the underlying economics of publicly traded companies (i.e., expected growth in earnings or expected growth in dividends) and the underlying economics of the economy (i.e., expected growth in Gross Domestic Product). Such studies conclude that investors could not have expected as large an ERP as the equity premiums actually realized.

Roger Ibbotson and Peng Chen report on their study of estimated forward looking long-term sustainable equity returns and expected ERPs.[12] They first analyzed historical equity returns by decomposing returns into factors including inflation, earnings, dividends, price-to-earnings ratio, dividend-payout ratio, book value, return on equity, and gross domestic product per capita. They forecast what could have been expected as an ERP through "supply side" models built from historical data. In the most recent update to this study reported in the *SBBI Yearbook*, Ibbotson Associates determined that the long-term ERP that could have been expected given the underlying economics was approximately 6.3% on an arithmetic basis (4.2% on a geometric basis) compared to the historical realized risk premium of 7.1% on an arithmetic basis (5.2% on a geometric basis). The greater-than-expected historical realized equity returns were caused by an unexpected increase in market multiples relative to economic fundamentals (i.e., decline in the discount rates).

What caused the decline in discount rates that led to the unexpected capital gain? The marginal income tax rate declined (the marginal tax rate on corporate distributions averaged 43% in the 1955–1962 period and averaged only 17% in the 1987–2000 period), and equity investments could not be held "tax free" in 1962. By 2000 however, equity investment could be held "tax deferred" in defined benefit and contribution pension plans and in individual retirement accounts. The decrease in income tax rates on corporate distributions and the inflow of retirement plan investment capital into equity

[8] E.g., Kaplan, "Why the Expected Rate of Return is an Arithmetic Average," 14–3 Business Valuation Review 126, (September 1995); *Stocks, Bonds, Bills and Inflation Valuation Edition 2005 Yearbook*, (Ibbotson Associates, 2005) pp 75–77; Kritzman, "What Practitioners Need to Know About Future Value," 50–3 *Financial Analysts Journal* 12 (May/June 1994); Bodie, Kane, and Marcus, *Investments* (Richard D. Irwin, Inc.,1989) p. 720.

[9] E.g., Damodaran, *Investment Valuation, 2nd ed*. (John Wiley & Sons, Inc., 2002) p. 161.

[10] Copeland, Koller and Murrin, *Valuation: Measuring and Managing the Value of Companies*, *3rd ed*. (John Wiley & Sons, Inc., 2000) p. 218; Koller, Goedhart and Wessels, *Valuation: Measuring and Managing the Value of Companies, 4th ed.* (John Wiley & Sons, Inc., 2005), p. 299–302; Cornell, *The Equity Risk Premium* (John Wiley & Sons, Inc.,1999) p. 36; Julius, "Market Returns in Rolling Multi-Year Holding Periods: An Alternative Interpretation to Ibbotson Data," 15–2 *Business Valuation Review* 57 (June 1996).

[11] Note 10, *supra*.

[12] Ibbotson and Chen, "Long-Run Stock Returns, Participating in the Real Economy," 591 *Financial Analysts Journal* 88 (January/February 2003) updated in *Stocks, Bonds, Bills and Inflation, Valuation Edition 2006 Yearbook* (Ibbotson Associates, 2006) p 98.

investments combined to lower discount rates and increase market multiples relative to economic fundamentals.[13]

Assuming that investors did not expect such changes, the true ERP during this period has been less than the historical realized premium calculated as the arithmetic average of excess returns realized since 1926. Further, assuming that the likelihood of changes in such factors being repeated are remote and investors do not expect another such decline in discount rates, the true ERP as of today can also be expected to be less than the historical realized premium.

Ex Ante Approaches

Merrill Lynch publishes "bottom-up" expected return estimates for the S&P 500 stock index derived from averaging return estimates for stocks in the S&P 500. While Merrill Lynch does not cover every company in the S&P 500 index, it does cover a high percentage of the companies as measured in market value terms. Merrill Lynch uses a multi-stage dividend discount model (DDM) to calculate expected returns for several hundred companies using projections from its own securities analysts. The resulting data is published monthly in the Merrill Lynch publication *Quantitative Profiles.* The Merrill Lynch expected return estimates have indicated an implied ERP ranging from 3% to 7% in recent years (approximately 6.6% at the end of 2005), with an average over the last 15 years of approximately 4.6%.[14]

Graham and Harvey report the results from a series of surveys of chief financial officers of U.S. corporations conducted from mid-2000 to the end of 2005. They report that the range of ERP given a ten-year investment horizon was 3.6% to 4.7% (premium over ten-year Treasury bonds). The most recent survey reports an ERP given a ten-year investment horizon was 4.7% on an arithmetic average basis (2.4% on a geometric average basis).[15]

Elroy Dimson, Paul Marsh and Mike Staunton studied the realized equity returns and historical equity premiums for 17 countries (including the U.S.) from 1900 to the end of 2005.[16]

These authors report that the historical equity premiums have been 6.5% on an arithmetic basis (4.6% on a geometric basis) for the U.S. (in excess of the total return on bonds) and 5.2% on an arithmetic basis (4.0% on a geometric basis) for the total of the 17 countries.

They observe larger equity returns earned in the second half of the 20th century compared to the first half due to (1) corporate cash flows growing faster than investors anticipated fueled by rapid technological change and unprecedented growth in productivity and efficiency, (2) transaction and monitoring costs falling over the course of the century, (3) inflation rates generally declining over the final two decades of the century and the resulting increase in real interest rates, and (4) required rates of return on equity declining due to diminished business and investment risks. They conclude that the observed increase in the overall price-to-dividend ratio during the century is attributable to the long-term decrease in the required risk premium and that the decrease will not continue into the future. The authors note that:

> Further adjustments should almost certainly be made to historical risk premiums to reflect long-term changes in capital market conditions. Since, in most countries corporate cash flows historically exceeded investors' expectations, a further downward adjustment is in order.

They conclude that a downward adjustment in the expected ERP compared to the historical equity premiums due to the increase in price/dividend ratio is reasonable. Further, they conclude that a further downward adjustment in the expected ERP of approximately 50 to 100 basis points is plausible if one assumes that the current level of dividend yield will continue (versus the greater historical average yield).

Removing the historical increase in the price/dividend ratio and adjusting the historical average dividend yield to today's dividend yield results in an expected equity premium (relative to bonds) of approximately 4.8% - 5.3% on an arithmetic basis (2.8% - 3.3% on a geometric basis) for the U.S. and 3.5% - 4.0% on an arithmetic

[13] McGrattan and Prescott, "Is the Market Overvalued?" *Federal Reserve Bank of Minneapolis Quarterly Review* (24,2000) and "Taxes, Regulations and Asset Prices," Federal Reserve Bank of Minneapolis working paper 610 (July, 2001).

[14] Use of analyst projections leads one to the literature on analyst projection bias (i.e., are analyst forecasts overly optimistic?). For example, see Ramnath, Rock and Stone, "*Value Line* and *I/B/E/S* earnings forecasts", working paper (Nov 2001). Those authors reports the results of projected earnings amounts, rather than growth rates (they use the *I/B/E/S* longterm growth rate to project the EPS four years into the future, and compares this with the actual EPS four years in the future. The results indicate that *I/B/E/S* mean forecast error in year 4 EPS is negative. This can be translated into a preliminary typical growth rate adjustment for say a projected 15% growth rate follows: ((1.15^4)(1-.0545))^.25 −1 = 13.4%, implying a ratio of actual to forecast of .134/.15 = .89. This would imply that equity risk premium forecasts using analyst forecasts are biased high. See also, Bonini, Zanetti and Bianchini, "Target Price Accuracy in Equity Research", working paper (Jan 2006).

[15] Graham and Harvey, "Expectations of Equity Risk Premia, Volatility and Asymmetry from a Corporate Finance Perspective," National Bureau of Economic Research working paper, December 2001, updated quarterly by *Duke CFO Outlook Survey* (www.cfosurvey.org); "The Equity Risk Premium in January 2006: Evidence from the Global CFO Outlook Survey", Dec 19, 2005.

[16] Dimson, Marsh and Staunton, "Global Evidence on the Equity Premium," 15–4 The *Journal of Applied Corporate Finance* (Summer 2003); "The Worldwide Equity Premium: A Smaller Puzzle", April 7, 2006; *The Global Investment Returns Yearbook 2006* (ABN-AMRO/ London Business School, 2006)

basis (2.4% - 2.9% on a geometric basis) for a world index (denominated in U.S. dollars for 17 countries).[17]

The *SBBI Yearbook* reports on an update to the work authored by Roger Ibbotson and Peng Chen, forecasting ERP based on the contribution of earnings growth to price to earnings ratio growth and on growth in per capital gross domestic product (a "supply side" approach).[18] They remove the increase in historical returns due to the overall increase in price-to-earnings ratio from 1926 to 2005 resulting in an estimate of ERP at the end of 2005 of approximately 6.3% on an arithmetic basis (4.2% on a geometric basis).

William Goetzmann and Roger Ibbotson commenting on the supply side approach of estimating expected risk premiums note:

> These forecasts tend to give somewhat lower forecasts than historical risk premiums, primarily because part of the total returns of the stock market have come from price-earnings ratio expansion. This expansion is not predicted to continue indefinitely,and should logically be removed from the expected risk premium.[19]

Tim Koller, Marc Goedhart, and David Wessels conclude on their assessment of the research and evidence:

> Although many in the finance profession disagree about how to measure the (ERP), we believe 4.5 to 5.5 percent is the appropriate range.[20]

Conclusion

Estimating the ERP is one of the most important issues when you estimate the cost of capital of the subject business. One needs to consider a variety of alternative sources including examining realized returns over various periods and employing forward-looking estimates such as those implied from projections of future prices, dividends, and earnings.

What is a reasonable estimate of ERP in 2006? While giving consideration to long-run historical arithmetic averages realized returns, this author concludes that the post-1925 historical arithmetic average of one-year realized premiums as reported in the *SBBI Yearbook* results in an expected ERP estimate that is too high. I come to that conclusion based on the works of various researchers (e.g., Dimson, Marsh and Staunton, Goetzmann and Ibbotson) and current market expectations (e.g., survey of chief financial officers).

Some appraisers express dismay over the necessity of considering a forward ERP since that would require changing their current "cookbook" practice of relying exclusively on the post-1925 historical arithmetic average of one-year realized premiums reported in the *SBBI Yearbook* as their estimate of the ERP. My reply – valuation is a forward-looking concept, not an exercise in mechanical application of formulas. Correct valuation requires applying value drivers reflected in today's market pricing. Our role is to mimic the market. In the experience of this author, one often cannot match current market pricing for equities using the post-1925 historical arithmetic average of one-year realized premiums as the basis for developing discount rates. The entire appraisal process is based on applying reasoned judgment to the evidence derived from economic, financial and other information and arriving at a well reasoned opinion of value. Estimating the ERP is no different. I challenge all appraisers to look at the evidence.

After considering the evidence, any reasonable long-term estimate of the *normal* ERP as of 2006 should be in the range of 3.5% to 6%.[21]

Roger Grabowski is a Managing Director of Duff & Phelps LLC in Chicago, Il. This author wants to thank Ryan Brown and David Turney of Duff and Phelps and my former colleague, David King, for their assistance. But I accept full responsibility for the final form of the paper. Moreover, this work should not be construed as representing the official organization position of any organization.

---

[17] Based on this author's converting premium over total returns on bonds as reported by Dimson, Marsh and Staunton, removing the impact of the growth in price-dividend ratios from the geometric average historical premium, reducing the historical average dividend yield to a current dividend yield and converting to an approximate arithmetic average.

One method of converting the geometric average into an arithmetic average is to assume the returns are independently log-normally distributed over time. Then the arithmetic and geometric averages approximately follow the relationship: Arithmetic average of returns for the period = Geometric average of returns for the period + (variance of returns for the period/2).

[18] Note 12, *supra*; Ibbotson, "Equity Risk Premium Forum," AIMR, 11/8/01, pp. 100–104, 108.

[19] Goetzmann and Ibbotson, "History and the Equity Risk Premium", Yale ICF Working Paper No. 05–04 (April 2005), p 8.

[20] Note 10, *supra*: Koller et al., p 306.

[21] Where in this range is the current ERP? Research has shown that ERP is cyclical during the business cycle. When the economy is near or in recession (and reflected in relatively recent low returns on stocks), the *conditional* ERP is more likely at the higher end of the range. When the economy improves (with expectations of improvements reflected in higher stock returns), the *conditional* ERP moves toward the mid-point of the range. When the economy is near its peak (and reflected in relatively recent high stock returns), the *conditional* ERP is more likely at the lower end of the range. This author will let the reader decide where his valuation date lies in the business cycle.