Journal of Business Finance & Accounting, 26(5) & (6), June/July 1999, 0306-686X

The Accuracy, Bias and Efficiency of Analysts' Long Run Earnings Growth Forecasts

Richard D.F. Harris*

1. INTRODUCTION

Considerable research has now been undertaken into professional analysts' forecasts of companies' earnings in respect of both their accuracy relative to the predictions of time series models of earnings, and their rationality. The evaluation of the reliability of analysts' earnings growth forecasts is an important aspect of research in accounting and finance for a number of reasons. Firstly, many empirical studies employ analysts' consensus forecasts as a proxy for the market's expectation of future earnings in order to identify the unanticipated component of earnings. The use of consensus forecasts in this way is predicated on the assumption that they are unbiased and efficient forecasts of future earnings growth. Secondly, institutional investors make considerable use of analysts' forecasts when evaluating and selecting individual shares. The quality of the forecasts that they employ therefore has important practical consequences for portfolio performance. Finally, from an academic point of view, the performance of analysts' forecasts is interesting because it sheds light on the process by which agents form expectations about key economic and financial variables.

* The author is from the Department of Economics, University of Exeter. This paper was started while the author was a Visiting Fellow at the University of Warwick, April 1997. Financial Support from the ESRC (project no. R000234759) is gratefully acknowledged. He would also like to thank George Bulkley, Alan Gregory and an anonymous referee for their helpful comments. (Paper received March 1998, revised and accepted December 1998)

Address for correspondence: Richard D.F. Harris, School of Business and Economics, University of Exeter, Exeter EX4 4PU, UK.
e-mail: R.D.F.Harris@exeter.ac.uk

Nearly all of the research to date, however, has been concerned with analysts' forecasts of quarterly and annual earnings per share.[1] While the properties of analysts' short run forecasts are undoubtedly important in their own right, it is long run expectations of earnings growth that are more relevant for security pricing (see, for instance, Brown et al., 1985). A number of papers have suggested that there is substantial mis-pricing in the stock market as a consequence of irrational long run earnings growth forecasts being incorporated into the market expectation of earnings growth (DeBondt, 1992; La Porta, 1996; Bulkley and Harris, 1997; and Dechow and Sloan, 1997). The evaluation of the performance of analysts' long run forecasts is clearly important as corroborating evidence.

This paper provides a detailed study of the accuracy, bias and efficiency of analysts' long run earnings growth forecasts for US companies. It identifies a number of characteristics of forecast earnings growth. Firstly, the accuracy of analysts' long run earnings growth forecasts is shown to be extremely low. So low, in fact, that they are inferior to the forecasts of a naïve model in which earnings are assumed to follow a martingale. Secondly, analysts' long run earnings growth forecasts are found to be significantly biased, with forecast earnings growth exceeding actual earnings growth by an average of about seven percent per annum. Thirdly, analysts' forecasts are shown to be weakly inefficient in the sense that forecast errors are correlated with the forecasts themselves. In particular, low forecasts are associated with low forecast errors, while high forecasts are associated with high forecast errors. The bias and inefficiency in analysts' long run forecasts are considerably more pronounced than in their short run and interim forecasts.

It is investigated whether analysts incorporate information about future earnings that is contained in current share prices. It is demonstrated that consistent with their short run and interim forecasts, analysts' long run earnings growth forecasts can be enhanced by assuming that each individual firm's earnings will evolve in such a way that its price-earnings ratio will converge to the current market average price-earnings ratio. Analysts therefore neglect valuable information about future earnings that is readily available at the time that their forecasts are made.

The source of analyst inaccuracy is explored by decomposing the mean square error of analysts' forecasts into two systematic components, representing the error that arises as a result of forecast bias and forecast inefficiency, and a random, unpredictable component. In principle, the systematic components of analysts' forecast errors can be eliminated by taking into account the bias and inefficiency in their forecasts. However, it is shown that the bias and inefficiency of analysts' forecasts contribute very little to their inaccuracy. Over eighty-eight percent of the mean square forecast error is random, while less than twelve percent is due to the systematic components. This is an important result for the users of analysts' forecasts since it means that correcting forecasts for their systematic errors can potentially yield only a small improvement in their accuracy.

A second decomposition is used to examine the level of aggregation at which forecast errors are made. The mean square forecast error is decomposed into the error in forecasting average earnings growth in the economy, the error in forecasting the deviation of average growth in each industry from average growth in the economy, and the error in forecasting the deviation of earnings growth for individual firms from average industry growth. It is demonstrated that the error in forecasting average earnings growth in the economy contributes relatively little to analysts' inaccuracy. Over half of total forecast error arises from the error in forecasting deviations of individual firm growth from average industry growth. The error in forecasting deviations of average industry growth from average growth in the economy is smaller, but also significant. However, there is evidence that this pattern is changing over time, with increasing accuracy at the industry level, and diminishing accuracy at the individual firm level.

Finally, it is shown that the performance of analysts' long run earnings growth forecasts varies substantially both with the characteristics of the company whose earnings are being forecast and of the forecast itself. The accuracy, bias and efficiency of analysts' forecasts is examined for sub-samples of firms partitioned by market capitalisation, price-earnings ratio, market-to-book ratio and the level of the forecast itself. The most reliable earnings growth forecasts are low forecasts issued for large companies with low price-earnings ratios and high market-to-book ratios. Again, this is of considerable practical importance since it offers users of analysts' forecasts some opportunity to discriminate between good and bad forecasts.

The organisation of this paper is as follows. The following section gives a detailed description of the data sources and the sample selection criteria. Section 3 describes the methodology used to evaluate forecast accuracy, bias and efficiency. Section 4 reports the results, while Section 5 concludes.

2. DATA

The sample is drawn from all companies listed on the New York, American and NASDAQ stock exchanges. Data on long run earnings growth expectations are taken from the Institutional Brokers Estimate System (IBES). The data item used in this paper is the 'expected EPS long run growth rate' (item 0), which has been reported by IBES since December 1981, and is defined as:

> the anticipated growth rate in earnings per share over the longer term. IBES Inc. requests that contributing firms focus on the five-year interval that begins on the first day of the current fiscal year and make their calculations based on projections of EPS before extraordinary items.

The expected long term growth rate is therefore taken to be the forecast average annual growth in earnings per share before extraordinary items, over the five year period that starts at the beginning of the current fiscal year.[2] The measure used in this paper is the median forecast calculated and reported in April of each year, *t*. The analysis was also conducted using the mean forecast, but the quantitative results are virtually identical, and the qualitative conclusions unchanged.[3]

Only December fiscal year end companies are included in the sample and so the use of the consensus forecast reported in April should ensure that the previous fiscal year's earnings are public information at the time that the individual forecasts that make up the consensus forecast are made (see Alford, Jones and Zmijewski, 1994). Restricting the sample to December fiscal year-end companies ensures that observations for a particular fiscal year span the same calendar period, thus allowing the identification of macroeconomic shocks that contemporaneously affect the earnings of all firms.

Actual growth in earnings is calculated using data on earnings per share, excluding extraordinary items, taken from the Standard and Poor's Compustat database (item EPSFX). Average annual earnings growth is computed as the average change in earnings over each five year period, from December of year $t-1$ to December of year $t+5$, scaled by earnings in December of year $t-1$. The need for five years' subsequent earnings growth data limits the sample period to the eleven years 1982–92. Data on a number of other variables are also used in the analysis. The share price and market capitalisation are both taken at the end of April of year t (Compustat items PRCCM and MKVALM). The market price-earnings ratio, used to test whether information contained in the share price is incorporated in analysts' forecasts, is computed as the price at the end of April in year t (item PRCCM) divided by earnings per share in the fiscal year ending December $t-1$ (item EPSFX). The market-to-book ratio is computed as the market value of the company in April of year t (item MKVALM) divided by the book value of the company in the fiscal year ending December of year $t-1$ (item CEQ).

There are a total of 7,660 firm-year observations that satisfy the data requirements for all the variables used in the analysis, and that have a December fiscal year-end. However, for 658 of these, earnings reported at the end of the preceding fiscal year are zero or negative. These are omitted from the sample since forecast growth has no natural interpretation when earnings in the base year are non-positive.[4] When initial earnings are close to zero, actual growth in earnings may take extreme values, resulting in outliers that have a disproportionately high degree of influence on the least squares regression results. There is no immediately obvious way to circumvent this problem without dropping some observations from the sample. The approach most commonly adopted is to omit observations for which the calculated growth rate, the forecast growth rate or the forecast error is above a certain threshold in absolute value, or for which calculated initial earnings are below a certain level. For instance, Fried and Givoly (1982) truncate observations for which forecast error exceeds 100%. Elton et al. (1984) include in their sample only those companies for which initial earnings are above 0.20 dollars per share. O'Brien (1988), in order to test the robustness of her results to outliers, also uses 0.20 dollars as a threshold value.

Capstaff et al. (1995) omit observations for which forecast earnings growth or forecast error exceeds 100%, while Capstaff et al. (1998) exclude companies for which forecast earnings growth or actual earnings growth exceeds 100%. In this paper, all observations for which actual earnings growth or forecast earnings growth exceeds 100% in absolute value are omitted from the analysis, reducing the sample by a further 336 firm-year observations. The final pooled sample comprises 6,666 firm-year observations.[5]

3. METHODOLOGY

(i) Forecast Accuracy

The metric used to evaluate forecast performance is the forecast error, defined as the difference between actual and forecast earnings growth:

$$fe_{it} = g_{it} - g_{it}^{f} \tag{1}$$

where fe_{it} is the forecast error for firm i corresponding to the forecast made at date t, g_{it} is actual earnings growth over the five year forecast period and g_{it}^{f} is forecast five year earnings growth. Forecast accuracy is evaluated using the mean square forecast error, which is computed in each year t as:

$$\mathrm{MSFE}_t = \frac{1}{N}\sum_{i=1}^{N}(g_{it} - g_{it}^{f})^2. \tag{2}$$

The mean square forecast error for the pooled sample is computed over all firms and years. The mean square forecast error was chosen in preference to the mean absolute forecast error to maintain consistency with the subsequent analysis which uses the former measure rather than the latter. However, it should be noted that the use of the mean square forecast error is consistent with a quadratic loss function of risk averse economic agents (see Theil, 1964; and Mincer and Zarnovitz, 1969). It can be reported that the conclusions drawn about forecast accuracy are not sensitive to the choice of measure.

As a benchmark against which to compare the accuracy of analysts' long run forecasts, the performance of two 'naïve' forecasts is also considered. The first is the forecast generated by a martingale model of earnings, in which expected earnings growth is zero. The second is the forecast generated by a sub-martingale model, in which expected earnings is equal to a drift parameter that is identical for all firms. In each forecast year, the common drift parameter is set equal to the average growth rate in earnings over all firms, over the previous five year period.[6] This choice of naïve forecasts is motivated by the early evidence on the time series properties of earnings, which suggests that annual earnings follow a random walk, or a random walk with drift (see, for instance, Brooks and Buckmaster, 1976; or Foster, 1977). Although more recent evidence finds that annual earnings may have a mean reverting component (see Ramakrishnan and Thomas, 1992), the martingale and sub-martingale models of earnings nevertheless provide simple alternative models that are approximately consistent with the reported evidence.

(ii) Forecast Bias

In order for a forecast to be unbiased, the unconditional expectation of the forecast error must be zero. If the average forecast error is greater than zero then analysts are systematically over-pessimistic (since their forecasts are on average exceeded) while if the average forecast error is less than zero analysts are systematically over-optimistic (since their forecasts are on average unfulfilled). Unbiasedness is tested using the mean forecast error, which is computed in each year t as:

$$\mathrm{MFE}_t = \frac{1}{N}\sum_{i=1}^{N}(g_{it} - g_{it}^{f}). \tag{3}$$

The mean forecast error for the pooled sample is computed over all firms and years. The hypothesis that the mean forecast error is zero is tested using the standard error of the mean forecast error across all firms and years for the pooled sample, and across all firms for each of the annual samples.

(iii) Forecast Efficiency

A forecast is efficient if it optimally reflects currently available information, and is therefore associated with a forecast error that is unpredictable. If a forecast is strongly efficient, the forecast error is uncorrelated with the entire information set at time *t*. Strong efficiency is a stringent condition, and so more usually forecasts are instead tested for weak efficiency, which requires that the forecast error is uncorrelated with the forecast itself (see Nordhaus, 1987). Weak efficiency is tested by estimating the following regression:

$$g_{it} = \alpha + \beta g_{it}^{f} + v_{it}. \tag{4}$$

Under the null hypothesis that analysts' forecasts are weakly efficient, the intercept, α, should be zero, while the slope coefficient, β, should be unity. If β is significantly different from one then conditioning on the forecast itself, the forecast error is predictable.[7] If β is significantly less than one then analysts' forecasts are too extreme, in the sense that high forecasts are associated with high forecast errors, while low forecasts are associated with low forecast errors. If β is significantly greater than one then forecasts are too compressed.

(iv) The Incremental Information Content of Price-Earnings Based Forecasts

A stronger form of forecast efficiency can be tested by examining whether analysts' forecasts incorporate particular sources of publicly available information. One such source of information is the current share price. In an efficient market, the share price is the present discounted value of all rationally expected future economic earnings of the company, and hence it should reflect, *inter alia*, the market's expectation of long run earnings growth. To extract the information about future earnings embodied in the share price, some assumption must be made about the company's cost of equity, or risk. The simplest assumption is that all companies face the same constant cost of equity in the long run, so that the earnings of each company evolve in such a way that its price-earnings ratio converges to the current market average price-earnings ratio. The earnings growth forecast that is implicit in this assumption can then be used to supplement the analysts' earnings growth forecast in the following regression:

$$g_{it} = \alpha + \beta g_{it}^{f} + \gamma g_{it}^{p} + v_{it}, \tag{5}$$

where

$$g_{it}^{p} = \frac{p_{it}/pe_{mt} - e_{it}}{e_{it}},\, pe_{mt} = \frac{1}{N}\sum_{i=1}^{N}\frac{p_{it}}{e_{it}}$$

and p_{it} is the share price of firm i at time t. If analysts incorporate all information contained in the current share price, the coefficient, γ, should be zero (see Capstaff et al., 1995 and 1998). Naturally, the assumption that all firms have the same long run price-earnings ratio is a strong simplification, and a superior forecast would almost certainly be obtained by assuming that price-earnings ratios differ between industries. Nevertheless, the assumption of a single market-wide long run price-earnings ratio has been shown to forecast earnings growth over shorter horizons (see, for instance, Ou and Penman, 1989).

(v) Forecast Error Decomposition

In order to analyse the source of analysts' forecast errors, two decompositions of the mean square forecast error are used. The first decomposes the mean square forecast error into systematic and unsystematic components. The systematic component is further divided into a component due to forecast bias and a component due to forecast inefficiency. In each year t, the decomposition of the MSFE is given by:

$$\text{MSFE}_t = \frac{1}{N_t}\sum_{i=1}^{N_t}(g_{it} - g_{it}^{f})^2 = (\overline{g}_t - \overline{g}_t^{f})^2 + (1-\beta_t)^2\sigma_{g^f t}^2 + (1-\rho_t^2)\sigma_{gt}^2 \tag{6}$$

where N_t is the sample size in year t, $\overline{g}_t$ and $\overline{g}_t^{f}$ are the average values of g_{it} and g_{it}^{f}, β_t is the slope coefficient from regression (4), above, ρ_t is the correlation coefficient between g_{it} and g_{it}^{f}, and $\sigma_{g^f t}^2$ and σ_{gt}^2 are the variances of g_{it} and g_{it}^{f}. The first term in the decomposition gives the error that is due to the inability of analysts to forecast earnings growth for the whole sample. When computed over all years, it is therefore a measure of the error that is due to forecast bias. The second term captures the error that is due to forecast inefficiency. Together, these two terms capture the systematic error in analysts' forecasts. In contrast, the third term captures the component of the error that is purely random. This decomposition is particularly useful since it reveals to what extent forecasts can be improved through 'optimal linear correction' procedures (see Mincer and Zarnovitz, 1969; and Theil, 1966). For instance, if the main component of mean square error is systematic, rather than random, then assuming that the data generating process for both the actual data and the forecast data remains constant, the accuracy of analysts' forecasts can be substantially improved by using the predicted values from regression (4), above, rather than the forecasts themselves. The extent to which this reduces the inaccuracy of the forecasts depends upon the fraction of the mean square forecast error that is due to the systematic component.

The second decomposition breaks the mean square forecast error into economy, industry and firm components. The decomposition of the MSFE is given each year t by:

$$\begin{aligned} \mathrm{MSFE}_t &= \frac{1}{N}\sum_{i=1}^{N_t}(g_{it} - g_{it}^f)^2 \\ &= (\overline{g}_t - \overline{g}_t^f)^2 + \frac{1}{N_t}\sum_{j=1}^{J_t} N_{jt}[(\overline{g}_{jt} - \overline{g}_t) - (\overline{g}_{jt}^f - \overline{g}_t^f)]^2 \qquad (7) \\ &\quad + \frac{1}{N_t}\sum_{i=1}^{Nt}[(g_{it} - \overline{g}_{jt}) - (g_{it}^f - \overline{g}_{jt}^f)]^2, \end{aligned}$$

where J_t is the number of industries in the sample, N_{jt} is the number of firms in industry j, $\overline{g}_{jt}$ and $\overline{g}_{jt}^f$ are the average values of g_{it} and g_{it}^f in industry j. The decomposition has the following interpretation. As before, the first term measures the error that is due to analysts' inability to forecast the average growth for the whole sample, which in this context may be interpreted as their inability to forecast earnings growth for the economy. The second term measures the error that is due to an inability to forecast the deviation of average growth in an industry from average growth in the economy. The third term measures the error that is due to an inability to forecast deviation of individual firm growth from average growth in its industry. The decomposition for the pooled sample is computed by taking the weighted average of the decomposition for the annual samples, with weights proportional to the sample size each year. Such a decomposition is useful because it reveals the level of aggregation at which forecast errors are made, and may reflect the particular approach used to generate earnings growth forecasts (see Elton, Gruber and Gultekin, 1984). In the present study, each industry is defined by a two digit SIC code. This yields a total of 56 industries, with an average of about twelve firms in each industry. The use of three digit SIC codes yields a large number of industries that comprise only a single firm. In these cases, the firm-specific error and industry specific error are not separately identifiable, and are reflected in the third component of the decomposition. The effect of using two digit, rather than three digit SIC codes is therefore to increase the firm specific error and reduce the industry specific error.

For both decompositions, it is convenient to express each term as a percentage of the total mean square forecast error. For the pooled samples, the mean square forecast error components are averaged over the individual years, with weights proportional to the sample size each year.

(vi) The Performance of Analysts' Forecasts Conditional on Firm and Forecast Characteristics

In order to explore possible heterogeneity in the performance of analysts' long run earnings growth forecasts, the sample is partitioned by various characteristics of the firm whose earnings are being forecast and of the forecast itself. Specifically, the sample is split into equally sized quintiles on the basis of market capitalisation, market-to-book ratio, price-earnings ratio and the level of the forecast itself. Forecast accuracy, bias and efficiency is then examined for each sub-sample. Forecast accuracy is measured by the mean square forecast error given by (2), forecast bias is measured by the mean forecast error given by (3), while forecast efficiency is measured by the estimated slope parameter in regression (4).

In order to identify the marginal effects of each of the firm and forecast characteristics on forecast accuracy, bias and weak form efficiency, the following regressions are estimated:

$$(g_{it} - g_{it}^f)^2 = \alpha + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it}, \qquad (10)$$

$$g_{it} - g_{it}^f = \alpha + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it} \qquad (11)$$

and

$$(g_{it}^f - \overline{g}_t^f)[(g_{it} - \overline{g}_t) - (g_{it}^f - \overline{g}_t^f)] = \alpha + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it}, \tag{12}$$

where $\ln m_{it}$ is the natural logarithm of the market capitalisation of firm i at the beginning of the forecast period, mb_{it} is the market-to-book ratio and pe_{it} is the price-earnings ratio. The dependent variables in the three regressions are the summands in (a) the mean square forecast error, (b) the mean forecast error and (c) the estimated covariance between $(g_{it} - g_{it}^f)$ and g_{it}^f.[8]

(vii) Estimation Procedure

In order to allow for time specific market wide shocks, each of the regression equations (4), (5), (9), (10), (11) and (12) is estimated by OLS, including fixed time effects. However, inference based on OLS estimates of the variance-covariance matrix of the disturbance term may be misleading since both heteroscedasticity and cross-sectional correlation are likely to be present in the data. One potential solution is to use GLS, in which the heteroscedasticity and cross-section correlation are parameterised and estimated. However, in the present case, GLS is infeasible since the number of cross-section observations is large relative to the number of time series observations. This paper employs instead the non-parametric approach of Froot (1989), which is robust to both contemporaneous correlation and heteroscedasticity. This involves partitioning the data by a two digit SIC code and assuming that the intra-industry correlation is zero. This then allows the consistent estimation of the parameter covariance matrix. The Froot estimator is modified using the Newey-West (1987) procedure in order to allow for the serial correlation in the regression error term that is induced by the use of overlapping data.

4. RESULTS

(i) Forecast Accuracy

Panel A of Table 1 reports the mean square forecast error, given by (2), for the pooled sample and for each individual year. It also reports the mean square forecast errors for the naïve forecasts of the martingale model, where forecast earnings growth is zero, and the sub-martingale model, where forecast earnings growth is the historical economy wide average earnings growth rate.

The accuracy of analysts' long run earnings growth forecasts is extremely low. In the pooled sample, the mean square forecast error for analysts is 7.15%. For the martingale model, the mean square error is 6.63%, while for the sub-martingale model, it is marginally lower at 6.60%. On average, therefore, a superior forecast of long run earnings growth for individual companies can be obtained simply by assuming that average annual earnings growth will be zero. This is a strong indictment of the accuracy of analysts' long run forecasts, and in view of the additional information available to analysts, is surprising. It also contrasts with the evidence for shorter horizon forecasts where analysts appear to have some advantage over time series models. Furthermore, the alternative models used here are relatively simple. If in fact earnings are stationary, then it is likely that a yet superior forecast could be obtained from an estimated time series model for each firm, and so the relative inferiority of analysts' forecasts is probably understated here.

Turning to the annual samples, the martingale model generates superior forecasts in seven out of eleven years, while the sub-martingale model generates forecasts that are superior to analysts' forecast in nine of the eleven years, and superior to the forecasts of the martingale model in ten out of eleven years. This suggests that one can improve on the zero growth forecast of the martingale model by using the historical economy average earnings growth rate to predict subsequent growth for individual firms. However, the improvement is only marginal, reflecting both considerable variation in average earnings growth between years and considerable dispersion in earnings growth rates across the economy. The time-series pattern of forecast errors suggests that analyst inferiority is not caused by just one or two outlying years. Nor does it suggest that there is any improvement in the accuracy of analysts' forecasts over the sample period, either relative to the forecasts of the martingale and sub-martingale models, or in absolute terms. The (unweighted) average mean square forecast error for the first five years in the sample is 7.02%, while in the last five years it is 7.28%. This is in contrast with evidence reported elsewhere that analyst accuracy has increased over time (see Brown, 1997).

(ii) Forecast Bias

Panel B of Table 1 reports the mean forecast error for analysts' forecasts of long run earnings growth, given by (3), and its standard error. In the pooled sample, the mean forecast error is negative indicating that analysts' long run earnings growth forecasts are over-optimistic. The mean forecast error is very significant both in statistical and economic terms. On average, forecast growth exceeds actual growth by about seven percent per annum. Over-optimism in long run earnings growth forecasts is consistent with evidence reported for analysts' shorter horizon earnings forecasts (see, for instance, Fried and Givoly, 1982; Brown et al., 1985; and O'Brien, 1988). It is also consistent with international evidence on analysts short run and interim forecasts (see Capstaff et al., 1995 and 1998).

The mean forecast error is also negative in each individual year, and significantly negative in all but the last, ranging from 1.50% to 11.82% per annum. This is in contrast with analysts' shorter horizon forecasts where the direction of the reported bias displays considerable year to year variation (see, for instance, Givoly, 1985). It is again notable that the degree of over-optimism has not diminished significantly over time. The (unweighted) mean forecast error for the first five years of the sample is −6.99%, while for the last five years it is −7.20%. It is of course possible that the last year in the sample, where the mean forecast error is less than two percent, marks the start of a reduction in analyst over-optimism. Whether this is borne out by future studies will be of considerable interest.

(iii) Forecast Efficiency

Panel A of Table 2 presents the results of regression (4). The efficiency condition is very strongly rejected for analysts' long run earnings growth forecasts. In the pooled sample, $\hat{\beta}$ is significantly less than unity and at 0.20, only marginally greater than zero. This is a considerably stronger rejection of efficiency than found by other authors for shorter horizon forecasts. For instance,

Table 1

Forecast Accuracy and Forecast Bias

	Panel A: Forecast Accuracy			**Panel B: Forecast Bias**	
	MSFE of Analysts	*MSFE of Martingale*	*MSFE of Sub-martingale*	*MFE of Analysts*	*Standard Error*
Pooled sample	7.15	6.63	6.60	−7.33	(0.31)
1982	7.34	5.15	6.41	−11.39	(1.01)
1983	6.88	7.01	6.51	−5.48	(1.20)
1984	6.75	7.14	6.40	−4.01	(1.12)
1985	7.19	6.67	6.29	−6.61	(1.08)
1986	6.92	6.47	6.24	−7.44	(1.08)
1987	6.95	5.77	5.75	−10.78	(0.99)
1988	7.38	6.32	6.40	−10.20	(1.00)
1989	6.99	5.22	5.71	−11.82	(0.91)
1990	5.69	5.20	4.95	−7.40	(0.85)
1991	7.58	7.78	7.60	−5.04	(0.99)
1992	8.78	9.62	9.78	−1.50	(1.10)

Notes:
Panel A reports the mean square forecast error for analysts' forecasts and the forecasts of two naïve models.

The MSFE of analysts forecasts is calculated each year as $\frac{1}{N}\sum_{i=1}^{N}(g_{it} - g_{it}^{f})^2$;

the MSFE of the martingale model is calculated each year as $\frac{1}{N}\sum_{i=1}^{N}(g_{it})^2$;

the MSFE of the sub-martingale model is calculated each year as $\frac{1}{N}\sum_{i=1}^{N}(g_{it} - \overline{g}_{t-1})^2$;

where g_{it} is five year earnings growth from January year t to December year t+4, is forecast of g_{it} reported at April year t and $\overline{g}_{t-1}$ is the average value over all companies of five year earnings growth from January year t−5 to December year t−1. The MSFE for the pooled sample is computed over all firms and years.

Panel B reports the mean forecast error of analysts, calculated as:

$$\text{MFE} = \frac{1}{N}\sum_{i=1}^{N}(g_{it} - g_{it}^{f}),$$

and its standard error. The MFE for the pooled sample is computed over all firms and years.

DeBondt and Thaler (1990) find that while they reject the hypothesis that β is equal to unity for one and two year forecasts, their estimated parameters (0.65 for one year forecasts, 0.46 for two year forecasts) are much larger than those reported here, both statistically and economically. For annual earnings forecasts,

Table 2

Forecast Efficiency

	Panel A: Weak Efficiency			Panel B: The Incremental Information Content of Price-Earnings Based Forecasts				
	$\hat{\beta}$	*SE*	$\overline{R}^2$	$\hat{\beta}$	*SE*	$\hat{\gamma}$	*SE*	$\overline{R}^2$
Pooled sample	0.20	(0.08)	0.00	0.05	(0.09)	0.04	(0.01)	0.02
1982	−0.73	(0.26)	0.04	−0.81	(0.28)	0.03	(0.04)	0.05
1983	0.42	(0.25)	0.01	0.08	(0.27)	0.05	(0.02)	0.04
1984	0.19	(0.27)	0.00	0.03	(0.30)	0.04	(0.02)	0.01
1985	0.05	(0.29)	0.00	0.02	(0.33)	0.01	(0.02)	0.00
1986	0.31	(0.23)	0.01	−0.25	(0.22)	0.10	(0.02)	0.06
1987	0.46	(0.22)	0.01	0.41	(0.22)	0.01	(0.02)	0.01
1988	0.42	(0.21)	0.01	0.43	(0.21)	0.00	(0.01)	0.01
1989	0.08	(0.22)	0.00	−0.03	(0.23)	0.03	(0.02)	0.01
1990	0.28	(0.17)	0.01	0.20	(0.20)	0.02	(0.02)	0.01
1991	0.39	(0.17)	0.01	0.11	(0.50)	0.06	(0.03)	0.03
1992	0.09	(0.27)	0.00	−0.20	(0.31)	0.10	(0.03)	0.05

Notes:
Panel A reports the results of the test of the weak efficiency of analysts' forecasts. The regression for the pooled sample is $g_{it} = \alpha_t + \beta g_{it}^f + u_{it}$ where g_{it} is five year earnings growth from January year t to December year t+4 and g_{it}^f is the median forecast of g_{it} reported in April of year t. The regression for the annual samples is $g_{it} = \alpha_t + \beta_t g_{it}^f + u_{it}$. The Panel reports the estimated slope parameter, its Froot-Newey-West adjusted standard error and the adjusted R-squared statistic.

Panel B reports the results of the test for the incremental information content of price-earnings based forecasts. The regression for the pooled sample is $g_{it} = \alpha_t + \beta g_{it}^f + \gamma g_{it}^p + u_{it}$ where g_{it} is five year earnings growth from January year t to December year t+4, is the median forecast of g_{it} reported in April of year t,

$$g_{it}^p = \frac{p_{it}/pe_{mt} - e_{it}}{e_{it}}, \quad pe_{mt} = \frac{1}{N}\sum_{i=1}^{N}\frac{p_{it}}{e_{it}},$$

e_{it} is the earnings reported in December of year $t-1$, and p_{it} is the price in April of year t. The regression for the annual samples is $g_{it} = \alpha_t + \beta_t g_{it}^f + \gamma_t g_{it}^p + u_{it}$. The Panel reports the estimated slope parameter, its Froot-Newey-West adjusted standard error and the adjusted R-squared statistic.

Givoly (1985) cannot reject the hypothesis that β is unity. Using UK data on the forecasts of individual analysts, Capstaff et al. (1995) find that the estimated coefficient declines with the forecast horizon, with an estimated value of around 0.5 for 20 month forecasts (their longest horizon). The results of this paper therefore strongly support the view (first offered by DeBondt and Thaler, 1990) that forecast earnings growth is too extreme, and that the longer the horizon, the more extreme it becomes. In the annual regressions, β is significantly less than unity in all years, and significantly greater than zero in only three years. In one year, it is actually significantly negative.

(iv) The Incremental Information Content of Price-Earnings Based Forecasts

The results of regression (5), which supplements analysts' forecasts with forecasts that are derived from the assumption that earnings will evolve in such a way that each firm's price-earnings ratio will converge to the current market price-earnings ratio, are reported in Panel B of Table 2. Under the null hypothesis that analysts make optimal use of information about future earnings that is contained in share prices, the coefficient on the price-earnings based forecast, $\hat{\gamma}$, should be zero. In the pooled sample, the estimated coefficient is significantly greater than zero, implying that analysts do not make full use of information that is readily available at the time that their forecasts are made. However, there is much year to year variation in both the statistical and economic significance of the coefficient, with six years in which the coefficient is not statistically different from zero.

The marginal contribution of price-earnings based forecasts can be gauged by comparing the two Panels of Table 2. The inclusion of the price-earnings forecast explains an additional two percent of the variation in actual earnings growth in the pooled sample, while in individual years, this figure varies between zero and five percent. However, the price-earnings based forecast used in the present analysis is derived under the somewhat unrealistic assumption that all firms have a common long run price-earnings ratio. Undoubtedly, more accurate earnings growth forecasts could be imputed by making more sophisticated assumptions about how price-earnings ratios evolve over time. The results presented here therefore almost certainly understate the extent to which analysts neglect information embodied in share prices. The fact that analysts appear to neglect information contained in share prices when forming their long run earnings growth forecasts is consistent with analogous results for their forecasts over shorter horizons (see, for instance, Ou and Penman, 1989; Abarbanell, 1991; Elgers and Murray, 1992; and Capstaff et al., 1995 and 1998).

(v) Forecast Error Decomposition

The preceding results demonstrate that the accuracy of analysts' long run earnings forecasts is extremely low, and that they are very significantly biased and inefficient. In this sub-section, the source of analysts' forecast error is investigated using the two decompositions of mean square forecast error described in Section 3. The first decomposes forecast error into systematic and non-systematic components. The results of this decomposition are given in Panel A of Table 3. It can be seen that by far the largest component of mean square forecast error is random. In the pooled sample, less than twelve percent of the forecast error is the result of the systematic component of analysts' forecast errors. Of the systematic component, about seven percent is due to bias, and about four percent due to inefficiency. A similar pattern holds for the annual samples, although there is considerable year to year variation, with as much as ninety-five percent of mean square forecast error accounted for by the random component in some years. In principle, knowledge of the systematic error in analysts' forecasts permits the use of 'optimal linear correction' techniques in order to improve forecast accuracy. This involves employing the predicted values calculated using the estimated coefficients from regression (4), above, in place of the forecasts themselves. The effect of the ordinary least squares regression is to adjust the forecasts by compensating for their bias and inefficiency. The degree to which accuracy can be enhanced in this way depends upon the proportion of the mean square forecast error that is systematic. The results reported here imply that, assuming that the underlying data generating process for actual earnings growth and the method by which analysts form the expectations of earnings growth remain constant, optimal linear correction of the forecasts will reduce the forecast error only by about twelve percent. This is clearly an important result for the users of analysts' forecasts.

The second decomposition divides the mean square forecast error into the error in forecasting average earnings growth in the economy, the error in forecasting the deviation of average growth in each industry from average growth in the economy, and the error in forecasting the deviation of earnings growth for

Table 3

Forecast Error Decomposition

	Panel A : Decomposition by Error Type			Panel B: Decomposition by Level of Aggregation		
	Bias	*Inefficiency*	*Random*	*Economy*	*Industry*	*Firm*
Pooled sample	7.51	4.07	88.45	9.21	35.53	55.25
1982	17.67	15.41	67.23	17.67	46.06	36.27
1983	4.37	2.12	93.92	4.37	40.21	55.42
1984	2.38	4.64	93.34	2.38	52.27	45.34
1985	6.07	6.68	87.57	6.07	36.45	57.48
1986	8.00	2.96	89.37	8.00	40.59	51.41
1987	16.73	1.86	81.69	16.73	30.15	53.11
1988	14.10	2.04	84.13	14.10	29.77	56.13
1989	20.02	5.32	74.89	20.02	27.45	52.53
1990	9.62	4.49	86.13	9.62	31.68	58.69
1991	3.35	2.63	94.27	3.35	33.05	63.60
1992	0.26	4.78	95.24	0.26	32.13	67.61

Notes:
Panel A reports the results of the decomposition of mean square forecast error for each year t by error type, given by:

$$\text{MSFE} = \frac{1}{N_t}\sum_{i=1}^{N_t}(g_{it} - g_{it}^f)^2 = (\overline{g}_t - \overline{g}_t^f)^2 + (1-\beta_t)^2\sigma^2_{g^f t} + (1-\rho_t^2)\sigma^2_{gt}$$

where N_t is the sample size in year t, g_{it} is five year earnings growth from January year t to December year t+4, g_{it}^f is the median forecast of g_{it} reported in April of year t, $\overline{g}_t$ and $\overline{g}_t^f$ are the average values of g_{it} and g_{it}^f, β_t is the slope coefficient reported in Panel A of Table 2, ρ_t is the correlation coefficient between g_{it} and g_{it}^f, and σ^2_{gt} and $\sigma^2_{g^f t}$ are the variances of g_{it} and g_{it}^f. The decomposition for the pooled sample is computed over all firms and years.

Panel B reports the results of the decomposition of mean square forecast error for each year t by the level of aggregation, given by:

$$\text{MSFE} = \frac{1}{N_t}\sum_{i=1}^{N_t}(g_{it} - g_{it}^f)^2$$

$$= (\overline{g}_t - \overline{g}_t^f)^2 + \frac{1}{N_t}\sum_{j=1}^{J_t}N_{jt}[(\overline{g}_{jt} - \overline{g}_t) - (\overline{g}_{jt}^f - \overline{g}_t^f)]^2 + \frac{1}{N_t}\sum_{i=1}^{N_t}[(g_{it} - \overline{g}_{jt}) - (g_{it}^f - \overline{g}_{jt}^f)]^2$$

where J_t is the number of industries in the sample, N_{jt} is the number of firms in industry j, $\overline{g}_{jt}$ and $\overline{g}_{jt}^f$ are the average values of g_{it} and g_{it}^f in industry j. The decomposition for the pooled sample is the weighted average of the decompositions for the annual samples, with weights proportional to the sample size each year. The table reports each of the components of mean square forecast error as a percentage of total mean square forecast error.

individual firms from average industry growth. The results of this decomposition are reported in Panel B of Table 3. The results demonstrate that analysts' forecast inaccuracy derives mainly from an inability to forecast deviations of individual firm growth from the average growth rate in its industry. The error in forecasting deviations of industry growth from the average growth rate in the economy is also important, but somewhat smaller than the error in forecasting individual firm growth. In contrast, analysts' inability to forecast average earnings growth in the economy contributes relatively little to their inaccuracy. An interesting feature of this decomposition is that the proportion of forecast error generated at the industry level appears to be diminishing over time, while the proportion generated at the individual firm level is increasing. This is potentially related to changes in the methods used by analysts to forecast earnings growth, or changes in accounting standards.

(vi) The Performance of Analysts' Forecasts Conditional on Firm and Forecast Characteristics

The foregoing analysis has considered analysts' long run earnings growth forecasts as a homogenous group. However, it is likely that forecast performance will vary with the characteristics of the firm whose earnings are being forecast. For instance, one would expect that firms with highly variable cash flows, or those for which little information is available about future earnings prospects, would be associated with lower forecast accuracy. Additionally, forecast performance is likely to vary with the size of the forecast itself since the efficiency results indicate that low forecasts are less overly-optimistic than high forecasts.

In order to investigate this issue, the accuracy, bias and efficiency results are reproduced for sub-samples of companies, partitioned on the basis of market capitalisation, price-earnings ratio, market-to-book ratio and the level of the forecast itself. For each variable, the sample is sorted into ascending order of the partitioning variable and split into quintiles, with equal numbers of firms in each quintile.[10] For all the results of this section, results are reported for quintiles pooled across all years only.

Table 4 presents the results for forecast accuracy, with the mean square forecast error for each quintile reported in Panel A.

There is substantial variation in forecast accuracy across market capitalisation, price-earnings ratio and forecast earnings growth, while there is no obvious systematic variation in forecast accuracy across market-to-book. Forecast accuracy increases with market capitalisation, with forecasts for the quintile of largest firms more than twice as accurate as those for the quintile of smallest firms. There is an inverse relationship between forecast accuracy and price-earnings ratio, with forecasts for the lowest quintile almost three times as accurate as those for the highest quintile. The largest variation in forecast accuracy is with the level of the forecast itself, with low forecasts being five times more accurate than high forecasts. In all three cases, variation in forecast accuracy is monotonic (almost monotonic in the case of price-earnings and forecast size), although it does not appear to be linear, with the largest differences occurring in the lowest and highest quintiles.

The results of Panel A show that forecast accuracy varies substantially with market capitalisation, price-earnings ratio and the forecast itself. However, these variables are not independent, and so variation in forecast accuracy with one variable may merely reflect variation with another. In order to identify the marginal effects of firm and forecast characteristics on forecast accuracy, Panel B of Table 4 reports the regression of the squared forecast error on the natural logarithm of market capitalisation, market-to-book, price-earnings and forecast earnings growth. Interestingly, all four variables independently contribute to the explanation of forecast accuracy, with the most influential, in terms of statistical significance, being the price-earnings ratio, followed by the level of the forecast itself. The most accurate forecasts are therefore low forecasts issued for large companies with low price-earnings ratios and high market-to-book ratios. The four variables together explain more than thirteen percent of the variation in forecast accuracy.

The variation of forecast accuracy with market capitalisation is not surprising. Information about future earnings prospects is likely to be more readily available, and of a higher quality, for larger firms. The variation of forecast accuracy with the forecast itself is consistent with the results on forecast efficiency. The inverse relationship between forecast accuracy and price-earnings ratio is harder to explain, but may be driven by the fact that very

Table 4

Forecast Accuracy Conditional on Firm and Forecast Characteristics

Panel A: Forecast Accuracy by Firm and Forecast Characteristics

	Quintile 1 (lowest)	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5 (highest)*
Capitalisation	11.52	8.24	6.35	5.19	4.47
Market-to-Book	7.84	6.51	6.36	7.18	7.88
Price-Earnings	5.30	4.53	5.02	6.13	14.79
Forecast Size	2.77	6.56	5.70	7.46	13.38

Panel B: The Marginal Effect of Firm and Forecast Characteristics on Forecast Accuracy

	Estimated Coefficient	*Standard Error*
Capitalisation	−103.18	(14.39)
Market-to-Book	−17.02	(6.80)
Price-Earnings	24.47	(3.55)
Forecast Growth	42.67	(6.17)
$\overline{R}^2$	0.13	

Notes:

Panel A reports the MSFE in percent for each quintile of firm-year observations sorted in ascending order of market capitalisation, market-to-book ratio, price-earnings ratio and forecast earnings growth.

Panel B reports the estimated slope coefficients from the regression:

$$(g_{it} - g_{it}^f)^2 = \alpha_t + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it}$$

where g_{it} is five year earnings growth from January year t to December year $t+4$, g_{it}^f is the median forecast of g_{it} reported in April of year t, m_{it} is the market capitalisation of firm i in April of year t, mb_{it} is the ratio of market capitalisation of firm i in April of year t to the book value of equity firm i in December of year $t-1$ and pe_{it} is the ratio of the share price of firm i in April of year t to the earnings for the fiscal year ending in December of year $t-1$. Froot-Newey-West adjusted standard errors are reported in parentheses. The regression is estimated for the sample pooled over all years.

high price-earnings ratios arise partly as a result of very low, but transitory earnings, the trajectory of which is likely to be difficult to forecast accurately. The positive relationship between forecast accuracy and market-to-book ratio is potentially explained by the fact that high market-to-book companies, *ceteris paribus*, should on average have high earnings growth. Since forecast earnings growth is generally too optimistic, the size of the forecast error for these companies should on average be lower.

Table 5 presents the results for forecast bias. Again, there is strong variation in forecast bias with market capitalisation, price-earnings ratio and the level of the forecast itself. Consistent with the results for forecast accuracy reported in Table 4, forecast bias decreases (in absolute value) with market capitalisation and increases with forecast size. However, while forecast inaccuracy increases with price-earnings ratio, forecast bias *decreases* with price-earnings ratio, implying that while forecasts become less biased as the price-earnings ratio increases, they nevertheless become less accurate. However, this merely implies that the random component of forecast inaccuracy decreases more rapidly with price-earnings ratio than does the systematic component. The largest variation in forecast bias is again with forecast size, with forecasts in the highest quintile being more than four times as biased as those in the lowest quintile. This is consistent with the results on efficiency reported earlier that demonstrate a significant negative relationship between forecast error and the level of the forecast. There is some variation in forecast bias with market-to-book value of equity, although it is not monotonic across quintiles, and the difference between the lowest and highest quintile is not large. There is no quintile of companies for which it can be concluded that analysts' forecasts are unbiased.

Panel B reports the results of the regression of forecast error on market capitalisation, market-to-book value of equity, price earnings ratio and forecast earnings growth. There is again independent variation in forecast bias with market capitalisation, price-earnings ratio and the level of the forecast itself, with the latter being the strongest factor, statistically speaking. There is no significant variation with market-to-book. The four variables together explain about six percent of the variation in forecast error.

These results are broadly consistent with Frankel and Lee (1996), who investigate the performance of analysts' shorter horizon forecasts in order to operationalise an accounting valuation model based on book value of equity and the market's expectation of earnings growth. They find that analyst over-optimism is associated with low book-to-price ratio (the inverse of the market-to-book ratio used in the present analysis) and high past sales growth. They also find that analyst over-optimism is

Table 5

Forecast Bias Conditional on Firm and Forecast Characteristics

Panel A: Forecast Bias by Firm and Forecast Characteristics

	Quintile 1 (lowest)	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5 (highest)*
Capitalisation	−12.28	−8.15	−5.99	−5.34	−5.00
	(0.87)	(0.75)	(0.67)	(0.60)	(0.56)
Market-to-Book	−5.32	−6.35	−8.61	−8.08	−8.38
	(0.75)	(0.68)	(0.65)	(0.70)	(0.73)
Price-Earnings	−11.66	−6.87	−7.42	−5.48	−5.32
	(0.54)	(0.55)	(0.58)	(0.66)	(1.04)
Forecast Size	−3.98	−3.56	−5.49	−7.59	−16.12
	(0.44)	(0.69)	(0.64)	(0.71)	(0.90)

Panel B: The Marginal Effect of Firm and Forecast Characteristics on Forecast Bias

	Estimated Coefficient	*Standard Error*
Capitalisation	0.76	(0.28)
Market-to-Book	0.05	(0.05)
Price-Earnings	0.23	(0.05)
Forecast Growth	−0.93	(0.09)
$\overline{R}^2$	0.06	

Notes:

Panel A reports the MFE in percent for each quintile of firm-year observations sorted in ascending order of market capitalisation, market-to-book ratio, price-earnings ratio and forecast earnings growth. Standard errors are reported in parentheses.

Panel B reports the estimated slope coefficients from the regression:

$$(g_{it} - g_t^f)^2 = \alpha_t + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g_{it}^f + v_{it}$$

where g_{it} is five year earnings growth from January year t to December year $t+4$, g_{it}^f is the median forecast of g_{it} reported in April of year t, m_{it} is the market capitalisation of firm i in April of year t, mb_{it} is the ratio of market capitalisation of firm i in April of year t to the book value of equity firm i in December of year $t-1$ and pe_{it} is the ratio of the share price of firm i in April of year t to the earnings for the fiscal year ending in December of year $t-1$. Froot-Newey-West adjusted standard errors are reported in parentheses. The regression is estimated for the sample pooled over all years.

associated with forecasts that are high relative to the current level of earnings (i.e. optimistic forecasts). Since forecast earnings growth and actual earnings growth are largely uncorrelated in the present sample, this is consistent with the finding reported above that analyst over-optimism is associated with high forecast earnings growth.

Table 6

Forecast Efficiency Conditional on Firm and Forecast Characteristics

Panel A: Forecast Efficiency by Firm and Forecast Characteristics

	Quintile 1 (lowest)	*Quintile 2*	*Quintile 3*	*Quintile 4*	*Quintile 5 (highest)*
Capitalisation	0.01	0.25	0.12	0.56	1.15
	(0.10)	(0.09)	(0.09)	(0.12)	(0.13)
Market-to-Book	0.05	0.01	0.00	−0.08	0.28
	(0.14)	(0.12)	(0.11)	(0.11)	(0.09)
Price-Earnings	−0.31	0.24	0.08	−0.04	−0.21
	(0.09)	(0.10)	(0.11)	(0.12)	(0.11)
Forecast Size	0.84	0.59	0.57	0.60	−0.11
	(0.26)	(0.86)	(0.98)	(0.84)	(0.13)

Panel B: The Marginal Effect of Firm and Forecast Characteristics on Forecast Efficiency

	Estimated Coefficient	*Standard Error*
Capitalisation	3.87	(2.30)
Market-to-Book	1.99	(1.14)
Price-Earnings	0.12	(0.63)
Forecast Growth	−12.47	(2.31)
$\overline{R}^2$	0.11	

Notes:
Panel A reports the estimate of β in the regression $g_{it} = \alpha_t + \beta g^f_{it} + u_{it}$ for each quintile of firm-year observations sorted in ascending order of market capitalisation, market-to-book ratio, price-earnings ratio and forecast earnings growth. Froot-Newey-West adjusted standard errors are reported in parentheses.

Panel B reports the estimated slope coefficients from the regression:

$$(g^f_{it} - \overline{g}^f_t)[(g_{it} - \overline{g}_t) - (g^f_{it} - \overline{g}^f_t)] = \alpha_t + \beta_1 \ln m_{it} + \beta_2 mb_{it} + \beta_3 pe_{it} + \beta_4 g^f_{it} + v_{it}$$

where g_{it} is five year earnings growth from January year t to December year $t+4$, g^f_{it} is the median forecast of g_{it} reported in April of year t, m_{it} is the market capitalisation of firm i in April of year t, mb_{it} is the ratio of market capitalisation of firm i in April of year t to the book value of equity firm i in December of year $t-1$ and pe_{it} is the ratio of the share price of firm i in April of year t to the earnings for the fiscal year ending in December of year $t-1$. Froot-Newey-West adjusted standard errors are reported in parentheses. The regression is estimated for the sample pooled over all years.

Table 6 presents the results for forecast efficiency. Panel A reveals that there is considerable variation in forecast efficiency across both market capitalisation and the level of the forecast, with some variation across market-to-book. The estimated slope parameter, β, is close to zero for the quintile of smallest firms, and rises monotonically with firm size. For the quintile of largest firms, the efficiency condition that $\beta = 1$ cannot be rejected. The estimated slope parameter decreases with the level of forecast, and for the quintile of firms with the lowest forecasts, the null hypothesis that $\beta = 1$ cannot be rejected either. There is no systematic variation with price-earnings ratio. The most efficient forecasts are therefore low forecasts for large firms with high market-to-book ratios.

Panel B of Table 6 reports the marginal contribution of each of the independent variables to forecast efficiency. Consistent with results of Panel A, there is positive independent variation in forecast efficiency with market capitalisation and market-to-book ratio, although the significance is marginal. Also consistent with the quintile results, the relationship between forecast efficiency and forecast growth is very significantly negative. There is no significant variation in forecast efficiency with price-earnings ratio. The four variables together explain eleven percent of the variation in forecast efficiency.

5. SUMMARY AND CONCLUSIONS

This paper has undertaken a detailed study of the accuracy, bias and efficiency of analysts' forecasts of long run earnings growth for US companies. The results of the paper can be summarised as follows.

(i) The accuracy of analysts' long run earnings growth forecasts is extremely low. Superior forecasts can be achieved simply by assuming that long run earnings growth is zero.
(ii) Analysts' forecasts are excessively optimistic. Forecast earnings growth, on average, exceeds actual earnings growth by about seven percent per annum.
(iii) Analysts' forecasts are weakly inefficient. Forecast errors are not independent of the forecasts themselves. In particular, high forecasts are associated with high forecast errors, while low forecasts are associated with low forecast errors.
(iv) Analysts' forecasts do not incorporate all information contained in current share prices. A superior forecast can be obtained by assuming that each firm's earnings will

evolve in such a way that its price-earnings ratio will converge to the current market-wide price-earnings ratio.

(v) Despite the bias and inefficiency identified in (ii) and (iii) above, the systematic components of analysts' forecast errors contribute relatively little to their inaccuracy. More than eighty-eight percent of the mean square forecast error is random. This is an important result for the users of analysts' long run earnings growth forecasts, since it means that the accuracy of analysts' forecasts cannot be significantly improved using linear correction techniques.

(vi) The largest part of analysts' forecast error is made at the individual firm level. The inability of analysts to forecast average earnings growth in the economy does not contribute substantially to their inaccuracy. However, there is evidence that the level of aggregation at which analysts' errors are being made is changing over time, with increasing accuracy at the industry level, and decreasing accuracy at the individual firm level.

(vii) There is significant heterogeneity in the performance of analysts' forecasts. The most reliable earnings growth forecasts are low forecasts issued for large companies with low price-earnings and high market-to-book ratios. The least biased forecasts are those for low forecasts for companies with low price-earnings ratios, while the most efficient forecasts are low forecasts for large companies with high market-to-book ratios. This is again an important result for the users of analysts' forecasts since it offers some opportunity to discriminate between good and bad forecasts.

(viii) There is very little evidence to suggest that the inaccuracy, bias or inefficiency of analyst' forecasts have diminished over time.

The idea that analysts systematically make over-optimistic forecasts, is not necessarily an indictment of their rationality *per se* since they may have considerable incentives to do so. An earnings growth forecast is not generally the final product delivered by an analyst to the client. In particular, earnings growth forecasts will be typically provided as part of a package of services, including brokerage, advice on mergers and acquisitions, and underwriting, and these related activities may influence the forecasts that an analyst makes (see Schipper, 1991). Sell-side analysts, for instance, have a vested interest in their clients' reaction to earnings forecasts. If earnings forecasts are used to support stock recommendations then high forecasts will tend to generate more business than low forecasts, since there is a larger potential client base for buy recommendations than for sell recommendations. Francis and Philbrick (1993) provide evidence that suggests that analysts may be intentionally over-optimistic in order to cultivate and maintain good management relations.

The decomposition of mean square forecast error by error type revealed that by far the largest component of analysts' forecast errors is random, with the systematic component accounting for less than twelve percent. Inevitably, at such long forecasting horizons, the potential to make accurate forecasts of earnings growth is limited. However, the fact that such a large component of actual earnings growth is random may explain why analysts' forecasts are so biased. The larger the component of the forecast error that is random, the lower the impact of forecast bias on forecast error. Assuming that analysts do have conflicting objectives — one to produce *accurate* earnings growth forecasts, the other to produce *high* earnings growth forecasts — then if analysts know that the first objective is largely unattainable, they will use the forecasting process to satisfy the second. If analysts are also producing short term and interim forecasts for the same company, then the bias in their long term forecasts may be compounded.

A number of papers have now concluded that there is substantial mis-pricing in the stock market as a consequence of irrational long run earnings growth forecasts being incorporated into the market expectation of earnings growth. The results of this paper support the hypothesis that analysts' consensus long run earnings growth forecasts are indeed irrational if they are to be interpreted as optimal forecasts of future earnings growth. However, given the uncertainty over analysts' incentives, it is by no means inevitable that these forecasts will be incorporated without modification into the market expectation of earnings growth. An interesting topic for future research will be to examine to what extent the market recognises the characteristics in forecast long run earnings growth identified in this paper.

NOTES

1 A partial list would include Brown and Rozeff (1978), Brown et al. (1987a and 1987b) and O'Brien (1988) who consider the performance of analysts' quarterly earnings forecasts, and Collins and Hopwood (1980), Fried and Givoly (1982) and Brown et al. (1985), who consider analysts' annual forecasts. International evidence on analysts' forecasts is provided by Capstaff et al. (1995), who analyse the performance of UK analysts, and Capstaff et al. (1998), who consider the forecasts of European analysts. For a comprehensive survey of the literature on analysts' earnings forecasts, see Brown (1993).

2 This was confirmed in conversation with IBES staff.

3 The correlation between the mean and the median forecast in the sample is 0.98. This is accounted for by the fact that most stocks have long term forecasts originating from only one or two analysts.

4 IBES have confirmed that they do receive earnings growth forecasts for companies whose earnings are currently negative. This may be explained by the fact that while analysts use the latest reported earnings as a base for earnings growth when earnings are positive, they use some other unspecified base measure of earnings, such as forecast annual earnings or average historical annual earnings, when earnings are negative.

5 In order to establish the robustness of the results, the analysis was conducted using maximum earnings growth threshold values in the range 50% to 1,000%, and by trimming the sample instead on the basis of initial earnings per share, using a minimum earnings threshold of between 0.10 and 1.00 dollars. The sensitivity of the results to changes in the threshold values was low, and none of the qualitative conclusions were altered. The regressions were additionally estimated using the minimum absolute deviation estimator, which is considerably less sensitive to outliers. This produced results that were almost completely invariant with respect to the choice threshold values. As a further test of the robustness of the results, the analysis was conducted using the change in earnings scaled by price, with the corresponding forecast change in earnings computed using the forecast growth rate. The results of these robustness tests are not reported here, but are available from the author on request.

6 The average growth rate is taken over all firms for which earnings data are available, using the same sample selection criteria as for subsequent earnings growth, namely excluding observations for which earnings are negative at the beginning of the five year period, and those for which the calculated growth rate exceeds 100% in absolute value.

7 This can be seen by subtracting forecast earnings growth, g_{it}^{f}, from each side so that the regression becomes one of forecast error on forecast earnings growth — the constant remains the same while the slope parameter becomes $\beta-1$.

8 Taking the conditional expectation of equations (10) and (11) gives the mean square forecast error and the mean forecast error, respectively, as a function of the independent variables. Regressions (10) and (11) thus measure the marginal contribution of each of the independent variables to forecast accuracy and forecast bias. Taking the conditional expectation of equation (12) gives the covariance between $(g_{it} - g_{it}^{f})$ and g_{it}^{f} as a function of the independent variables. This covariance is the numerator of the estimated slope coefficient in a regression of $g_{it} - g_{it}^{f}$ on g_{it}^{f}. Under the

null hypothesis that forecasts are weakly efficient, this covariance should be equal to zero. If it is less than zero, forecasts are too extreme, while if it is greater than zero, forecasts are too compressed. Regression (12) thus measures the marginal contribution of each of the independent variables to forecast efficiency.

9 See, for example, Brown et al. (1987a) and O'Brien (1988), who consider the accuracy of analysts' quarterly earnings forecasts relative to the forecasts of different time series models, and Fried and Givoly (1982), who consider the relative accuracy of analysts' annual earnings forecasts.

10 Except for the largest quintile, which has an additional observation.

REFERENCES

Abarbanell, J. (1991), 'Do Analysts' Earnings Forecasts Incorporate Information in Prior Stock Price Changes?' *Journal of Accounting and Economics*, Vol. 14, pp. 147–66.

Alford, A., J. Jones and M. Zmijewski (1994), 'Extensions and Violations of the Statutory SEC Form 10-K Filing Requirements', *Journal of Accounting and Economics*, Vol. 17, pp. 229–54.

Brooks, L. and D. Buckmaster (1976), 'Further Evidence on the Time Series Properties of Accounting Income', *Journal of Finance*, Vol. 31, pp. 1359–73.

Brown, L. (1993), 'Earnings Forecasting Research: Its Implications for Capital Markets Research', *International Journal of Forecasting*, Vol. 9, pp. 295–320.

________ (1997), 'Analyst Forecasting Errors: Additional Evidence', Working Paper (University at Buffalo).

________ and M. Rozeff (1978), 'The Superiority of Analysts' Forecasts as Measures of Expectations: Evidence from Earnings', *Journal of Finance*, Vol. 33, pp. 1–16.

________ R. Hagerman, P. Griffin and M. Zmijewski (1987a), 'Security Analyst Superiority Relative to Univariate Time-Series Models in Forecasting Quarterly Earnings', *Journal of Accounting and Economics*, Vol. 9, pp. 61–87.

________ ________ ________ ________ (1987b), 'An Evaluation of Alternative Proxies for the Market's Assessment of Unexpected Earnings', *Journal of Accounting and Economics*, Vol. 9, pp. 159–93.

Brown, P., G. Foster and E. Noreen (1985), 'Security Analyst Multi-Year Earnings Forecasts and the Capital Market', *American Accounting Association Studies in Accounting Research*, Vol. 21.

Bulkley, G. and R. Harris (1997), 'Irrational Analysts' Expectations as a Cause of Excess Volatility in Stock Prices', *Economic Journal*, Vol. 107, pp. 359–71.

Capstaff, J., K. Paudyal and W. Rees (1995), 'The Accuracy and Rationality of Earnings Forecasts by UK Analysts', *Journal of Business, Finance & Accounting*, Vol. 22, pp. 67–85.

________ ________ ________ (1998), 'A Comparative Analysis of Earnings Forecasting in Europe', manuscript (University of Glasgow).

Christie, A. (1987), 'On Cross-Sectional Analysis in Accounting Research', *Journal of Accounting and Economics*, Vol. 9, pp. 231–58.

Collins, W. and W. Hopwood (1980), 'A Multivariate Analysis of Annual Earnings Forecasts Generated from Quarterly Forecasts of Financial Analysts and Univariate Time-Series Models', *Journal of Accounting Research* Vol. 18, pp. 390–406.

DeBondt, W. (1992), *Earnings Forecasts and Share Price Reversals* (Foundation of Chartered Financial Analysts, Charlottsville, Virginia).

________ and R. Thaler (1990), 'Do Security Analysts Overreact?' *American Economic Review*, Vol. 80, pp. 52–57.

Dechow, P. and R. Sloan (1997), 'Returns to Contrarian Investment Strategies: Tests of Naive Expectations Hypotheses', *Journal of Financial Economics*, Vol. 43, pp. 3–27.

Elgers, P., and D. Murray (1992), 'The Relative and Complementary Performance of Analyst and Security-Price Based Measures of Expected Earnings', *Journal of Accounting and Economics*, Vol. 15, No. 303–16.

Elton, E., M. Gruber and M. Gultekin (1984), 'Professional Expectations: Accuracy and Diagnosis of Errors', *Journal of Financial and Quantitative Analysis*, Vol. 19, pp. 351–62.

Foster, G. (1977), 'Quarterly Accounting Data: Time Series Properties and Predictive-Ability Results', *Accounting Review*, Vol. 52, pp. 1–21.

Francis, J. and D. Philbrick (1993), 'Analysts' Decisions as Products of a Multi-Task Environment', *Journal of Accounting Research*, Vol. 31, pp. 216–30.

Frankel, R. and C. Lee (1996), 'Accounting Valuation, Market Expectation, and the Book-to-Market Effect', *Journal of Finance*, Vol. 51, pp. 1040–41.

Fried, D. and D. Givoly (1982), 'Financial Analysts' Forecasts of Earnings', *Journal of Accounting and Economics*, Vol. 4, pp. 85–107.

Froot, K. (1989), 'Consistent Covariance Matrix Estimation with Cross-Sectional Dependence and Heteroskedasticity in Financial Data', *Journal of Financial and Quantitative Analysis*, Vol. 24, pp. 333–55.

Givoly, D. (1985), 'The Formation of Earnings Expectations', *Accounting Review*, Vol. 60, pp. 372–86.

Lakonishok, J., A. Schleifer and R. Vishny, (1994), 'Contrarian Investment, Extrapolation, and Risk', *Journal of Finance*, Vol. 49, pp. 1541–78.

La Porta, R. (1996), 'Expectations and the Cross-Section of Stock Returns', *Journal of Finance*, Vol. 51, pp. 1715–42.

Mincer, J. and V. Zarnovitz (1969), 'The Valuation of Economic Forecasts', in J. Mincer (ed.), *Economic Forecasts and Expectations* (National Bureau of Economic Research).

Newey, W. and K. West (1987), 'A Simple Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix', *Econometrica*, Vol. 55, pp. 703–8.

O'Brien, P. (1988), 'Analysts' Forecasts as Earnings Expectations', *Journal of Accounting and Economics*, Vol. 10, pp. 53–83.

Ou, J. and S. Penman (1987), 'Financial Statement Analysis and the Prediction of Stock Returns', *Journal of Accounting and Economics*, Vol. 11, pp. 295–330.

Ramakrishnan, R. and J. Thomas (1992), 'What Matters from the Past: Market Value, Book Value or Earnings? Earnings Valuation and Sufficient Statistics for Prior Information', *Journal of Accounting, Auditing and Finance*, Vol. 7, pp. 423–64.

Schipper, K. (1991), 'Commentary on Analysts' Forecasts', *Accounting Horizons*, Vol. 5, pp. 105–21.

Theil, H. (1964), *Optimal Decision Rules for Government and Industry* (North Holland).

________ (1966), *Applied Economic Forecasting* (North Holland).

Darden Graduate School of Business
University of Virginia

Working Paper No. 99-08

THE MARKET RISK PREMIUM: EXPECTATIONAL ESTIMATES USING ANALYSTS' FORECASTS

Robert S. Harris
Felicia C. Marston

The Market Risk Premium: Expectational Estimates Using Analysts' Forecasts

Robert S. Harris
Felicia C. Marston

Draft October 12, 1999

Felicia C. Marston is on the faculty at the McIntire School of Commerce, University of VA 22903, (804) 924-1417, fm2v@virginia.edu. Robert S. Harris is on the faculty at the Darden School, University of Virginia, Charlottesville, VA 22906, (804) 924-4823, harrisr@darden.gbus.virginia.edu. We thank Erik Benrud and seminar participants at the University of Virginia and at the SEC for comments. Thanks to Darden Sponsors, TVA, the Walker Family Fund and McIntire Associates for support of this research and to IBES, Inc. for supplying data.

The Market Risk Premium: Expectational Estimates Using Analysts' Forecasts

Abstract

We use expectational data from financial analysts to estimate a market risk premium for U.S. stocks. Using the SP500 as a proxy for the market portfolio, we find an average market risk premium of 7.14% above yields on long-term U.S. government bonds over the period 1982-1998. We also find that this risk premium varies over time and that much of this variation can be explained by either the level of interest rates or readily available forward-looking proxies for risk. The market risk premium appears to move inversely with government interest rates suggesting that required returns on stocks are more stable than interest rates themselves.

The Market Risk Premium: Expectational Estimates Using Analysts' Forecasts

The notion of a market risk premium (the spread between investor required returns on safe and average risk assets) has long played a central role in finance. It is a key factor in asset allocation decisions to determine the portfolio mix of debt and equity instruments. Moreover, the market risk premium plays a critical role in the Capital Asset Pricing Model (CAPM), practitioners most widely used means of estimating equity hurdle rates. In recent years, the practical significance of estimating such a market premium has increased as firms, financial analysts and investors employ financial frameworks to analyze corporate and investment performance. For instance, the increased use of Economic Value Added to assess corporate performance has provided a new impetus for estimating capital costs.

The most prevalent approach to estimating the market risk premium relies on some average of the historical spread between returns on stocks and bonds.[1] This choice has some appealing characteristics but is subject to many arbitrary assumptions such as the relevant period for taking an average. Compounding the difficulty of using historical returns is the well noted fact that standard models of consumer choice would predict much lower spreads between equity and debt returns than have occurred in U.S. markets—the so called equity premium puzzle (see Welch (1998), Siegel and Thaler (1997)). In addition, theory calls for a forward looking risk premium that could well change over time.

[1] Bruner, Eades, Harris and Higgins (1998) provide survey evidence on both textbook advice and practitioner methods for estimating capital costs. Despite substantial empirical assault, the CAPM continues to play a major role in applied finance. As testament to the market for cost of capital estimates Ibbotson Associates (1998) publishes a "Cost of Capital Quarterly."

This paper takes an alternate approach by using expectational data to estimate the market risk premium. The approach has two major advantages for practitioners. First, it provides an independent estimate which can be compared to historical averages. At a minimum, this can help in understanding likely ranges for risk premia. Second, expectational data allow investigation of changes in risk premia over time. Such time variations in risk premia serve as important signals from investors that should affect a host of financial decisions.

The paper updates and extends earlier work (Harris (1986), Harris and Marston (1992)) which incorporates financial analysts' forecasts of corporate earnings growth. Updating through 1998 provides an opportunity to see whether changes in the risk premium are in part responsible for the run up in share prices in the bull market. In addition, we provide new tests of whether changes in risk premia over time are linked to forward-looking measures of risk. Specifically, we look at the relationship between the risk premium and four ex-ante measures of risk: the spread between yields on corporate and government bonds, consumer sentiment about future economic conditions, the average level of dispersion across analysts as they forecast corporate earnings and the implied volatility on the SP500 Index derived from options data.

Section I provides background on the estimation of equity required returns and a brief discussion of current practice in estimating the market risk premium. In Section II, models and data are discussed. Following a comparison of the results to historical returns in Section III, we examine the time-series characteristics of the estimated market premium in Section IV. Finally, conclusions are offered in Section V.

I. Background

The notion of a "market" required rate of return is a convenient and widely used construct. Such a rate (k) is the minimum level of expected return necessary to compensate investors for bearing the average risk of equity investments and receiving dollars in the future rather than in the present. In general, k will depend on returns available on alternative investments (e.g., bonds). To isolate the effects of risk, it is useful to work in terms of a market risk premium (*rp*), defined as

$$rp = k - i, \tag{1}$$

where i = required return for a zero risk investment.

Lacking a superior alternative, investigators often use averages of historical realizations to estimate a market risk premium. Bruner *et al.* (1998) provide recent survey results on best practices by corporations and financial advisors. While almost all respondents used some average of past data in estimating a market risk premium, a wide range of approaches emerged. "While most of our 27 sample companies appear to use a 60+- year historical period to estimate returns, one cited a window of less than ten years, two cited windows of about ten years, one began averaging with 1960, and another with 1952 data" (p. 22). Some used arithmetic averages and some geometric. This historical approach requires the assumptions that past realizations are a good surrogate for future expectations and, as typically applied, that the risk premium is constant over time. Carleton and Lakonishok (1985) demonstrate empirically some of the problems with such historical premia when they are dissaggregated for different time periods or groups of firms. As Bruner *et al* (1998) point out, few respondents cited use of expectational data to supplement or replace historical returns in estimating the market premium.

Survey evidence also shows substantial variation in empirical estimates. When respondents gave a precise estimate of the market premium, they cited figures from 4 to over 7 percent (Bruner *et al* 1998). A quote from a survey respondent highlights the range in practice. "In 1993, we polled various investment banks and academic studies on the issue as to the appropriate rate and got anywhere between 2 and 8%, but most were between 6 and 7.4%." (Bruner *et al* 1998, p. 23). An informal sampling of current practice also reveals large differences in assumptions about an appropriate market premium. For instance, in a 1999 application of EVA analysis, Goldman Sachs Investment Research specifies a market risk premium of "3%

from 1994-1997 and 3.5% from 1998-1999E for the S&P Industrials" (Goldman Sachs (1999, p. 59)). At the same time an April 1999 phone call to Stern Stewart revealed that their own application of EVA typically employed a market risk premium of 6%. In its application of the CAPM, Ibbotson Associates (1998) uses a market risk premium of 7.8%. Not surprisingly, academics don't agree on risk premium either. Welch (1998) surveyed leading financial economists at major universities. For a 30-year horizon, he found a mean risk premium of 6.12% but a range from 2% to 9% with an interquartile range of 2% (based on 104 responses).

To provide additional insight on estimates of the market premium, we use publicly available expectational data. This expectational approach employs the dividend growth model (hereafter referred to as the discounted cash flow or DCF model) in which a consensus measure of financial analysts' forecasts (FAF) of earnings is used as a proxy for investor expectations. Earlier works by Malkiel (1982), Brigham, Vinson, and Shome (1985), Harris (1986) and Harris and Marston (1992) have used FAF in DCF models[2].

II. Models and Data

We employ the simplest and most commonly used version of the DCF model to estimate shareholders' required rate of return, k, as shown in Equation (2):

$$k = \left(\frac{D_1}{P_0}\right) + g, \qquad (2)$$

where D_1 = dividend per share expected to be received at time one, P_0 = current price per share (time 0), and g = expected growth rate in dividends per share[3]. A primary difficulty in using the

[2] Ibbotson Associates (1998) use a variant of the DCF model with forward-looking growth rates as one means to estimate cost of equity; however, they do this as a separate technique and not as part of the CAPM. For their CAPM estimates they use historical averages for the market risk premium. The DCF approach with analysts' forecasts has been used frequently in regulatory settings.

[3] Our methods follow Harris (1986) and Harris and Marston (1992) who provide an overview of earlier research and a detailed discussion of the approach employed here. For instance, theoretically, i is a risk-free rate, though empirically its proxy (e.g., yield to maturity on a government bond) is only a "least risk" alternative that is itself subject to risk. They also discuss single versus multistage growth discounted cash flow models and procedures used in calculating the expected dividend yield. While the model calls for expected growth in dividends, in the long run, dividend growth is sustainable only via growth in earnings. As long as payout ratios are not expected to change, the two growth rates will be the same.

DCF model is obtaining an estimate of *g*, since it should reflect market expectations of future performance. This paper uses published FAF of long-run growth in earnings as a proxy for g. Equation (2) can be applied for an individual stock or any portfolio of companies. We focus primarily on its application to estimate a market premium as proxied by the SP500.

FAF come from IBES Inc. The mean value of individual analysts' forecasts of five-year growth rate in EPS is used as our estimate of *g* in the DCF model. The five-year horizon is the longest horizon over which such forecasts are available from IBES and often is the longest horizon used by analysts. IBES requests "normalized" five-year growth rates from analysts in order to remove short-term distortions that might stem from using an unusually high or low earnings year as a base. Growth rates are available on a monthly basis.

Dividend and other firm-specific information come from COMPUSTAT. D_1 is estimated as the current indicated annual dividend times (1+g). Interest rates (both government and corporate) are gathered from Federal Reserve Bulletins and *Moody's Bond Record.* Table 1 describes key variables used in the study. Data are collected for all stocks in the Standard & Poor's 500 stock (SP500) index followed by IBES. Since five-year growth rates are first available from IBES beginning in 1982, the analysis covers the period from January 1982-December 1998.

We generally adopt the same approach as used in Harris and Marston (1992). For each month, a market required rate of return is calculated using each dividend paying stock in the SP500 index for which data are available. As additional screens for reliability of data, in a given month we eliminate a firm if there are fewer than three analysts' forecasts or if the standard deviation around the mean forecast exceeds 20%. Combined these two screens eliminate fewer than 20 stocks a month. Later we report on the sensitivity of our results to various screens. The DCF model in Equation (2) is applied to each stock and the results weighted by market value of equity to produce the market-required return.[4] The risk premium is constructed by subtracting the interest rate on government bonds.

For short-term horizons (quarterly and annual), past research (Brown, 1993) finds that on average analysts' forecasts are overly optimistic compared to realizations. However, recent research on quarterly horizons (Brown, 1997) suggests that analysts' forecasts for SP500 firms do not have an optimistic bias for the period 1993-1996. There is very little research on the properties of five-year growth forecasts, as opposed to shorter horizon predictions.[5] Any analysts' optimism is not necessarily a problem for our analysis. If investors share analysts' views, our procedures will still yield unbiased estimates of required returns and risk premia. In light of the possible bias, however, we interpret our estimates as "upper bounds" for the market premium.

To broaden our exploration, we tap four very different sources to create ex ante measures of equity risk at the market level. The first proxy comes from the bond market and is calculated as the spread between corporate and government bond yields (BSPREAD). The rationale is that increases in this spread signal investors' perceptions of increased riskiness of corporate activity that would be translated to both debt and equity owners. The second measure, CON, is the consumer confidence index reported by the Conference Board at the end of the month. While the reported index tends to be around 100, we rescale CON as the actual index divided by 100. We also examined use of CON as of the end of the prior month; however, in regression analysis

---

[4] We weighted 1998 results by year-end 1997 market values since our monthly data on market value did not extend through this period. Since we did not have data on firm-specific dividend yields for the last four months of 1998, we estimated the market dividend yield for these months using the dividend yield reported in the *Wall Street Journal* scaled by the average ratio of this figure to the dividend yield for our sample as calculated in the first eight months of 1998. We then made adjustments using growth rates from IBES to calculate the market required return. We also estimated results using an average dividend yield for the month which employed the average of the price at the end of the current and prior months. These average dividend yield measures led to essentially the same regression coefficients as those reported later in the paper but introduced significant serial correlation in some regressions (Durbin-Watson statistics significantly different from 2.0 at the .01 level).

[5] To our knowledge, the only studies of possible bias in analysts' five-year growth rates are Boebel (1991) and Boebel, Harris and Gultekin (1993). They both find evidence of optimism in IBES growth forecasts. In the most thorough study to date, Boebel (1991) reports that this bias seems to be getting smaller over time. His forecast data do not extend into the 1990's.

this lagged measure was generally not statistically significant in explaining the level of the market risk premium[6]. The third measure, DISP, measures the dispersion of analysts' forecasts. Such analyst disagreement should be positively related to perceived risk since higher levels of uncertainty would likely generate a wider distribution of earnings forecasts for a given firm. DISP is calculated as the equally weighted average of firm-specific standard deviations for each stock in the SP500 covered by IBES. The firm-specific standard deviation is calculated based on the dispersion of individual analysts' growth forecasts around the mean of individual forecasts for that company in that month. Our final measure, VOL, is the implied volatility on the SP500 index. As of the beginning of the month, we use a dividend adjusted Black Scholes Formula to estimate the implied volatility in the SP500 index option contract which expires on the third Friday of the month. The call premium, exercise price and the level of the SP500 index are taken from the *Wall Street Journal* and treasury yields come from the Federal Reserve. Dividend yield comes from DRI. We use the option contract that is closest to being at the money.

III. Estimates of the Market Premium

Table 2 reports both required returns and risk premia by year (averages of monthly data). The results are quite consistent with the patterns reported earlier (e.g., Harris and Marston, 1992). The estimated risk premia are positive, consistent with equity owners demanding additional rewards over and above returns on debt securities. The average expectational risk premium (1982 to 1998) over government bonds is 7.14%, slightly higher than the 6.47% average for 1982 to 1991 reported earlier (Harris and Marston, 1992). For comparison purposes, Table 3 contains historical returns and risk premia. The average expectational risk premium

[6] We examined two other proxies for Consumer Confidence. The Conference Board's Consumer Expectations Index yielded essentially the same results as those reported. The University of Michigan's Consumer Sentiment Indices tended to be less significantly linked to the market risk premium though coefficients were still negative.

reported in Table 2 is approximately equal to the arithmetic (7.5%) long-term differential between returns on stocks and long-term government bonds.[7]

Table 2 shows the estimated risk premium changes over time, suggesting changes in the market's perception of the incremental risk of investing in equity rather than debt securities. Scanning the next to last column of Table 2, the risk premium is higher in the 1990's than earlier and especially so in late 1997 and 1998. Our DCF results provide no evidence to support the notion of a declining risk premium in the 1990's as a driver of the strong run up in equity prices.

A striking feature in Table 2 is the relative stability of our estimates of k. After dropping (along with interest rates) in the early and mid-1980's, the average annual value of k has remained within a 75 basis point range around 15 percent for over a decade. Moreover, this stability arises despite some variability in the underlying dividend yield and growth components of k as Table 2 illustrates. The results suggest that k is more stable than government interest rates. Such relative stability of k translates into parallel changes in the market risk premium. In a subsequent section, we examine whether changes in our market risk premium estimates appear linked to interest rate conditions and a number of proxies for risk[8].

We explored the sensitivity of our results to our screening procedures in selecting companies. Our reported results screen out all non-dividend paying stocks on the premise that use of the DCF model is inappropriate in such cases. The dividend screen eliminates an average of 55 companies per month. In a given month, we also screen out firms with fewer than three analysts' forecasts, or if the standard deviation around the mean forecast exceeds 20%. When we repeated our analysis without any of the screens, the average risk premium over the sample

---

[7] Interestingly, for the 1982-1996 period the arithmetic spread between large company stocks and long-term government bonds was only 3.3% per year. The downward trend in interest rates resulted in average annual returns of 14.1% on long-term government bonds over this horizon. Some (e.g., Ibbotson, 1997) argue that only the income (not total) return on bonds should be subtracted in calculating risk premia.

[8] Although our focus is on the market risk premium, in earlier work (Harris and Marston (1992), Marston, Harris and Crawford (1993)), we examined the cross-sectional link between expectational equity risk premia at the firm level and beta and found a significant positive correlation. For comparative purposes, we replicated and updated that

period increased by only 40 basis points, from 7.14% to 7.54%. We also estimated the beta of our sample firms and found the sample average to be one, suggesting that our screens do not systematically remove low or high-risk firms. Specifically, using firms in our screened sample as of December 1997 (the last date for which we had CRSP return data), we used ordinary least squares regressions to estimate beta for each stock using the prior sixty months of data and the CRSP return (SPRTRN) as the market index. The value-weighted average of the individual betas was 1.00.

In the results reported here we use firms in the SP500 as reported by COMPUSTAT in September 1998 which could create a survivorship bias, especially in the earlier months of our sample. We compared our current results to those obtained in our earlier work (Harris and Marston (1992)) for which we had data to update the SP500 composition each month. For the overlapping period, January 1982-May 1991 the two procedures yield the same average market risk premium, 6.47%. This suggests that the firms departing from or entering the SP500 index do so for a number of reasons with no discernable effect on the overall estimated SP500 market risk premium.

IV. Changes in the Market Risk Premium Over Time

With changes in the economy and financial markets, equity investments may be perceived to change in risk. For instance, investor sentiment about future business conditions likely affects attitudes about the riskiness of equity investments compared to investments in the bond markets. Moreover, since bonds are risky investments themselves, equity risk premia (relative to bonds) could change due to changes in perceived riskiness of bonds, even if equities displayed no shifts in risk.

In earlier work covering the 1982-1991 period, Harris and Marston (1992) reported regression results indicating that the market premium decreased with the level of government

---

analysis through 1998 and reached very similar conclusions. At the firm level our expectational estimates of risk interest rates and increased with the spread between corporate and government bond yields (BSPREAD). This bond yield spread was interpreted as a time series proxy for equity risk. We introduce three additional ex ante measures of risk shown in Table 1: CON, DISP and VOL. The three measures come from three independent sets of data and are supplied by different agents in the economy (consumers, equity analysts and investors (via option and share price data)). Table 4 provides summary data on all four of our risk measures.

Table 5 replicates and updates earlier analysis.[9] The results confirm the earlier patterns. For the entire sample period, Panel A shows that risk premia are negatively related to interest rates. This negative relationship is also true for both the 1980's and 1990's as displayed in Panels B and C. For the entire 1982 to 1998 period, the addition of the yield spread risk proxy to the regressions lowers the magnitude of the coefficient on government bond yields, as can be seen by comparing Equations 1 and 2 of Panel A. Furthermore, the coefficient of the yield spread (0.487) is itself significantly positive. This pattern suggests that a reduction in the risk differential between investment in government bonds and in corporate activity is translated into a lower equity market risk premium.

In major respects, the results in Table 5 parallel earlier findings. The market risk premium changes over time and appears inversely related to government interest rates but positively related to the bond yield spread, which proxies for the incremental risk of investing in equities as opposed to government bonds. One striking feature is the large negative coefficients on government bond yields. The coefficients indicate the equity risk premium declines by over 70 basis points for a 100 basis point increase in government interest rates.[10] This inverse

---

premia are significantly positively correlated to beta.

[9] OLS regressions with levels of variables generally showed severe autocorrelation. As a result, we used the Prais-Winsten method (on levels of variables) and also OLS regressions on first differences of variables. Since both methods yielded similar results and the latter had more stable coefficients across specifications, we report only the results using first differences. Tests using Durbin-Watson statistics from regressions in Tables 5 and 6 do not accept the hypothesis of autocorrelated errors (tests at .01 significance level, see Johnston 1984, pp. 321-325).

[10] The Table 5 coefficients on i are significantly different from -1.0 suggesting that equity required returns do respond to interest rate changes. However, the large negative coefficients imply only minor adjustments of required

relationship suggests much greater stability in equity required returns than is often assumed. For instance, standard application of the CAPM suggests a one-to-one change in equity returns and government bond yields.

Table 6 introduces three additional proxies for risk and explores whether these variables, either individually or collectively, are correlated with the market premium. Since our estimates of implied volatility start in May 1986, the table shows results for both the entire sample period and for the period during which we can introduce all variables. Entered individually each of the three variables is significantly linked to the risk premium with the coefficient having the expected sign. For instance, in regression (1) the coefficient on CON is -.014 which is significantly different from zero (t = -3.50). The negative coefficient signals that higher consumer confidence is linked to a lower market premium. The positive coefficients on VOL and DISP indicate the equity risk premium increases with both market volatility and disagreement among analysts. The effects of the three variables appear largely unaffected by adding other variables. For instance, in regression (4) the coefficients on CON and DISP both remain significant and are similar in magnitude to the coefficients in single variable regressions.

Even in the presence of the new risk variables, Table 6 shows that the market risk premium is affected by interest rate conditions. The large negative coefficient on government bond rates implies large reductions in the equity premium as interest rates rise. One feature of our data may contribute to the observed negative relationship between the market risk premium and the level of interest rates. Specifically, if analysts are slow to report updates in their growth forecasts, changes in our estimated *k* would not adjust fully with changes in the interest rate even if the true risk premium were constant. To address the impact of "stickiness" in the measurement of *k*, we formed "quarterly" measures of the risk premium which treat *k* as an average over the

---

returns to interest rate changes since the risk premium declines. In earlier work (Harris and Marston (1991)) the coefficient was significantly negative but not as large in absolute value. In that earlier work we reported results quarter. Specifically, we take the value of k at the end of a quarter and subtract from it the average value of i for the months ending when k is measured. For instance, to form the risk premium for March 1998 we take the March value of k and subtract the average value of i for January, February and March. This approach assumes that in March k still reflects values of g that have not been updated from the prior two months. We then pair our quarterly measure of risk premium with the average values of the other variables for the quarter. For instance, the March 1998 "quarterly" risk premium would be paired with averaged values of BSPREAD over the January through March period. To avoid overlapping observations for the independent variables, we use only every third month (March, June, September, December) in the sample.

As reported in Table 7, sensitivity analysis using "quarterly" observations suggests that delays in updating may be responsible for a portion, but not all, of the observed negative relationship between the market premium and interest rates. For example, when we use quarterly observations the coefficient on i in regression (2) of Table 7 is -.527, well below the earlier estimates but still significantly negative[11].

As an additional test, we look at movements in the bond risk premium (BSPREAD). Since BSPREAD is constructed directly from bond yield data it does not have the potential for reporting lags that may affect analysts' growth forecasts. Regression 3 in Table 7 shows BSPREAD is negatively linked to government rates and significantly so[12]. While the equity premium need not move in the same pattern as the corporate bond premium, the negative coefficient on BSPREAD suggests that our earlier results are not due solely to "stickiness" in measurements of market required returns.

---

using the Prais-Winsten estimators. When we use that estimation technique and recreate the second regression in Table 5, the coefficient for i is -.584 (t = 12.23) for the entire sample period 1982-1998.

[11] Sensitivity analysis for the 1982-1989 and 1990-1998 subperiods yields results similar to those reported.

[12] We thank Bob Conroy for suggesting use of BSPREAD. Regression 3 in Table 7 appears to have autocorrelated errors: the Durbin-Watson (DW) statistic rejects the hypothesis of no autocorrelation. However, in subperiod analysis, the DW statistic for the 1990-98 period is consistent with no autocorrelation and the coefficient on i is essentially the same (-.24, t = -8.05) as reported in Table 7.

The results in Table 7 suggest that the inverse relationship between interest rates and the market risk premium may not be as pronounced as suggested in earlier tables. Still, there appears to be a significant negative link between the equity risk premium and government interest rates. The quarterly results in Table 7 would suggest about a 50 basis point change in risk premium for each 100 basis point movement in interest rates.

Overall, our ex ante estimates of the market risk premium are significantly linked to ex ante proxies for risk. Such a link suggests that investors modify their required returns in response to perceived changes in the environment. The findings provide some comfort that our risk premium estimates are capturing, at least in part, underlying economic changes in the economic environment. Moreover, each of the risk measures appears to contain relevant information for investors. The market risk premium is negatively related to the level of consumer confidence and positively linked to interest rate spreads between corporate and government debt, disagreement among analysts in their forecasts of earnings growth and the implied volatility of equity returns as revealed in options data.

II. Conclusions

Shareholder required rates of return and risk premia are based on theories about investors' expectations for the future. In practice, however, risk premia are typically estimated using averages of historical returns. This paper applies an alternate approach to estimating risk premia that employs publicly available expectational data. The resultant average market equity risk premium over government bonds is comparable in magnitude to long-term differences (1926 to 1998) in historical returns between stocks and bonds. As a result, our evidence does not resolve the equity premium puzzle; rather, our results suggest investors still expect to receive large spreads to invest in equity versus debt instruments.

There is strong evidence, however, that the market risk premium changes over time. Moreover, these changes appear linked to the level of interest rates as well as ex ante proxies for risk drawn from interest rate spreads in the bond market, consumer confidence in future economic conditions, disagreement among financial analysts in their forecasts and the volatility of equity returns implied by options data. The significant economic links between the market premium and a wide array of risk variables suggests that the notion of a constant risk premium over time is not an adequate explanation of pricing in equity versus debt markets.

Our results have implications for practice. First, at least on average, our estimates suggest a market premium roughly comparable to long-term historical spreads in returns between stocks and bonds. Our conjecture is that, if anything, our estimates are on the high side and thus establish an upper bound on the market premium. Second, our results suggest that use of a constant risk premium will not fully capture changes in investor return requirements. As a specific example, our findings indicate that common application of models such as the CAPM will overstate changes in shareholder return requirements when government interest rates change. Rather than a one-for-one change with interest rates implied by use of constant risk premium, our results indicate that equity required returns for average risk stocks likely change by half (or less) of the change in interest rates. However, the picture is considerably more complicated as shown by the linkages between the risk premium and other attributes of risk.

Ultimately, our research does not resolve the answer to the question "What is the right market risk premium?" Perhaps more importantly, our work suggests that the answer is conditional on a number of features in the economy—not an absolute. We hope that future research will harness ex ante data to provide additional guidance to best practice in using a market premium to improve financial decisions.

Table 1. Variable Definitions

k	=	Equity required rate return.
P_0	=	Price per share.
D_1	=	Expected dividend per share measured as current indicated annual dividend from COMPUSTAT multiplied by $(1 + g)$.
g	=	Average financial analysts' forecast of five-year growth rate in earnings per share (from IBES).
i	=	Yield to maturity on long-term U.S. government obligations (source: Federal Reserve, 30-year constant maturity series).
rp	=	Equity risk premium calculated as $rp = k - i$.
BSPREAD	=	spread between yields on corporate and government bonds, BSPREAD = yield to maturity on long-term corporate bonds (Moody's average across bond rating categories) minus i.
CON	=	Monthly consumer confidence index reported by the Conference Board (divided by 100).
DISP	=	Dispersion of analysts' forecasts at the market level.
VOL	=	Volatility for the SP500 index as implied by options data.

Table 2. Bond Market Yields, Equity Required Return, and Equity Risk Premium, 1982-1998

Values are averages of monthly figures in percent. *i* is the yield to maturity on long-term government bonds, *k* is the required return on the SP500 estimated as a value weighted average using a discounted cash flow model with analysts' growth forecasts. The risk premium $rp = k - i$. The average of analysts' growth forecasts is *g*. *Div yield* is expected dividend per share divided by price per share.

Year	*Div yield*	*g*	*K*	*i*	$rp = k - i$
1982	6.89	12.73	19.62	12.76	6.86
1983	5.24	12.60	17.86	11.18	6.67
1984	5.55	12.02	17.57	12.39	5.18
1985	4.97	11.45	16.42	10.79	5.63
1986	4.08	11.05	15.13	7.80	7.34
1987	3.64	11.01	14.65	8.58	6.07
1988	4.27	11.00	15.27	8.96	6.31
1989	3.95	11.08	15.03	8.45	6.58
1990	4.03	11.69	15.72	8.61	7.11
1991	3.64	11.99	15.63	8.14	7.50
1992	3.35	12.13	15.47	7.67	7.81
1993	3.15	11.63	14.78	6.60	8.18
1994	3.19	11.47	14.66	7.37	7.29
1995	3.04	11.51	14.55	6.88	7.67
1996	2.60	11.89	14.49	6.70	7.79
1997	2.18	12.60	14.78	6.60	8.17
1998	1.80	12.95	14.75	5.58	9.17
Average	3.86	11.81	15.67	8.53	7.14

Table 3. Average Historical Returns on Bonds, Stocks, Bills, and Inflation in the U.S., 1926-1998

Historical Return Realizations	Geometric Mean	Arithmetic Mean
Common Stock (large company)	11.2%	13.2%
Long-term government bonds	5.3%	5.7%
Treasury bills	3.8%	3.8%
Inflation rate	3.1%	3.2%

Source: Ibbotson Associates, Inc., *1999 Stocks, Bonds, Bills and Inflation*, 1999 Yearbook.

Table 4. Descriptive Statistics on Ex Ante Risk Measures

Entries are based on monthly data. BSPREAD is the spread between yields on long-term corporate and government bonds. CON is the consumer confidence index. DISP measures the dispersion of analysts' forecasts of earnings growth. VOL is the volatility on the SP500 index implied by options data. Variables are expressed in decimal form, e.g., 12% = .12.

A. Variable
Monthly Levels

	Mean	Standard Deviation	Minimum	Maximum
BSPREAD	.0123	.0040	.0070	.0254
CON	.9500	.2240	.473	1.382
DISP	.0349	.0070	.0285	.0687
VOL	.1599	.0696	.0765	.6085

B. Variable
Monthly Changes

	Mean	Standard Deviation	Minimum	Maximum
BSPREAD	-.00001	.0011	-.0034	.0036
CON	.0030	.0549	-.2300	.2170
DISP	-.00002	.0024	-.0160	.0154
VOL	-.0008	.0592	-.2156	.4081

C. Correlation Coefficients for Monthly Changes
*significantly different from zero at the .05 level
**significantly different from zero at the .01 level

	BSPREAD	CON	DISP	VOL
BSPREAD	1.00	-.16*	.05	.22**
CON	-.16*	1.00	.07	-.09
DISP	.05	.07	1.00	.03
VOL	.22**	-.09	.03	1.00

Table 5. Changes in the Market Equity Risk Premium Over Time

The table reports regression coefficients (*t*-values). Regression estimates use all variables expressed as monthly changes to correct for autocorrelation. The dependent variable is the market equity risk premium for the SP500 index. BSPREAD is the spread between yields on long-term corporate and government bonds. The yield to maturity on long-term government bonds is denoted as *i*. For purposes of the regression, variables are expressed in decimal form, e.g., 12% = .12.

Time period	Intercept	*i*	BSPREAD	R^2
A. 1982-1998	-.0002	-.8696		.57
	(-1.49)	(-16.54)		
	-.0002	-.749	.487	.59
	(-1.11)	(-11.37)	(2.94)	
B. 1980's	-.0005	-.887		.56
	(-1.62)	(-10.97)		
	-.0004	-.759	.508	.57
	(-1.24)	(-7.42)	(1.99)	
C. 1990's	-.0000	-.840		.64
	(-0.09)	(-13.78)		
	-.0000	-.757	.347	.65
	(0.01)	(-9.85)	(1.76)	

Table 6. Changes in the Market Equity Risk Premium Over Time and Selected Measures of Risk

The table reports regression coefficients (*t*-values). Regression estimates use all variables expressed as monthly changes to correct for autocorrelation. The dependent variable is the market equity risk premium for the SP500 index. BSPREAD is the spread between yields on long-term corporate and government bonds. The yield to maturity on long-term government bonds is denoted as *i*. CON is the change in consumer confidence index. DISP measures the dispersion of analysts' forecasts of earnings growth. VOL is the volatility on the SP500 index implied by options data. For purposes of the regression, variables are expressed in decimal form, e.g., 12% = .12.

Time period		Intercept	*i*	BSPREAD	CON	DISP	VOL	*Adj.* R^2
A. 1982-1998								
	(1)	0.0002 (.97)			-0.014 (-3.50)			0.05
	(2)	-0.0001 (-.96)	-0.737 (-11.31)	0.453 (2.76)	-0.007 (-2.48)			0.60
	(3)	0.0002 (.78)				0.244 (2.38)		0.02
	(4)	-0.0001 (-.93)	-0.733 (-11.49)	0.433 (2.69)	-0.007 (-2.77)	0.185 (3.13)		0.62
B. May 1986-1998								
	(5)	0.0000 (.03)	-0.821 (-11.16)	0.413 (2.47)	-0.005 (-2.22)	0.376 (3.74)		0.68
	(6)	0.0001 (.53)					0.011 (2.89)	0.05
	(7)	0.0000 (.02)	-0.831 (-11.52)	0.326 (1.95)	-0.005 (-2.12)	0.372 (3.77)	0.006 (2.66)	0.69

Table 7. Regressions Using Alternate Measures of Risk Premia to Analyze Potential Effects of Reporting Lags in Analysts' Forecasts

The table reports regression coefficients (*t*-values). Regression estimates use all variables expressed as changes (monthly or quarterly) to correct for autocorreclation. BSPREAD is the spread between yields on long-term corporate and government bonds. *rp* is the risk premium on the SP500 index. The yield to maturity on long-term government bonds is denoted as *i*. For purposes of the regression, variables are expressed in decimal form, e.g., 12% = .12.

Dependent Variable	Intercept	*i*	BSPREAD	*Adj.* R^2
(1) Equity Risk Premium (*rp*) Monthly Observations (same as Table 5)	-.0002 (-1.11)	-.749 (-11.37)	.487 (2.94)	.59
(2) Equity Risk Premium (*rp*) "Quarterly" nonoverlapping observations to account for lags in analyst reporting	-.0002 (-.49)	-.527 (-6.18)	.550 (2.20)	.60
(3) Corporate Bond Spread (BSPREAD) Monthly Observations	-.0001 (-1.90)	-.247 (-11.29)		.38

REFERENCES

R.B. Boebel, "The Information Content of Financial Analysts' Forecasts: Short-term vs. Long-term," University of North Carolina at Chapel Hill, Dissertation, 1991. University Microfilms International order number 9207936.

R.B. Boebel, M.N. Gultekin and R.S. Harris, "Financial Analysts' Forecast of Corporate Earnings Growth: Top Down Versus Bottom Up," in *Handbook of Security Analyst Forecasting and Asset Allocation*, John B. Guerard, Jr. and Mustafa N. Gultekin (eds.), London, JAI Press, 1993, pp. 185-192.

D. Brigham, D. Shome, and S. Vinson, "The Risk Premium Approach to Measuring A Utility's Cost of Equity," *Financial Management* (Spring 1985), pp. 33-45.

L.D. Brown, " Earnings Forecasting Research: Its Implications for Capital Market Research," *International Journal of Forecasting* (November 1993), pp. 295-320.

L.D. Brown, "Analyst Forecasting Errors: Additional Evidence," *Financial Analysts Journal* (November/December 1997), pp. 81-90.

R.F. Bruner, K. Eades, R. Harris and R. Higgins, Best Practices in Estimating the Cost of Capital: Survey and Synthesis," *Financial Practice and Education* (Spring/Summer 1998), pp. 13-28.

W.T. Carleton and J. Lakonishok, "Risk and Return on Equity: The Use and Misuse of Historical Estimates," *Financial Analysts Journal* (January/February 1985), pp. 38-47.

L. Chan, Y. Hamao, and J. Lakonishok, "Fundamental and Stock Returns in Japan," *Journal of Finance*, 46, (5), (December 1991) pp. 1739-64.

J. Cragg and B.G. Malkiel, *Expectations and the Structure of Share Prices*, National Bureau of Economic Research, Chicago, University of Chicago Press, 1982.

Goldman Sachs, *EVA Company Analysis*, January 1999.

R.S. Harris, "Using Analysts' Growth Forecasts to Estimate Shareholder Required Rates of Return," *Financial Management* (Spring 1988), pp. 58-67.

Ibbotson Associates, Inc., *1998 Cost of Capital Quarterly*, 1998 Yearbook

Ibbotson Associates, Inc., *1997 Stocks, Bonds, Bills, and Inflation*, 1997 Yearbook.

J. Johnston, *Econometric Methods*, New York, McGraw-Hill, 3rd edition, 1984.

P. Kennedy, *A Guide to Econometrics*, Cambridge, MIT Press, 2nd edition, 1985.
G.S. Madalla, *Econometrics*, New York, McGraw-Hill, 1977.

B. Malkiel, "Risk and Return: A New Look," in *The Changing Role of Debt and Equity in Financing U.S. Capital Formation*, B.B. Friedman (ed.), National Bureau of Economic Research, Chicago, University of Chicago Press, 1982.

B. Malkiel, "The Capital Formation Problem in the United States," *Journal of Finance* (May 1979), pp. 291-306.

F. Marston, R. Harris, and P. Crawford, "Risk and Return in Equity Markets: Evidence Using Financial Analysts' Forecasts," in *Handbook of Security Analysts' Forecasting and Asset Allocation*, J. Guerard and M. Gultekin (eds.), Greenwich, CT, JAI Press, 1992, pp. 101-122.

F. Marston and R.S. Harris, "Risk, Return, and Equilibrium: A Revisit Using Expected Returns," *Financial Review*, (February 1993).

J. Siegel and R. Thaler, "Anomalies: The Equity Premium Puzzle," *Journal of Economic Perspectives*, 11 (1), (1997), pp. 191-200.

J. VanderWeide and W.T. Carleton, "Investor Growth Expectations: Analysts vs. History," *Journal of Portfolio Management* (Spring 1988), pp. 78-82.

I. Welch, "Views of Financial Economists on the Equity Premium and Other Issues," UCLA/Anderson Finance Working Paper 10-98, May 18, 1998.

How the Risk Premium Factor Model and Loss Aversion Solve the Equity Premium Puzzle

Stephen D. Hassett

Hassett Advisors
www.HassettAdvisors.com
SHassett@HassettAdvisors.com

This Draft: September 22, 2010
First Draft: September 20, 2010

Abstract

The term "equity premium puzzle" was coined in 1985 by economists Rajnish Mehra and Edward C. Prescott. The equity premium puzzle in considered one of the most significant questions in finance. A number of papers have explored the fundamental questions of why the premium exists and has not been arbitraged away over time. This paper expands upon the findings implicit in the Risk Premium Valuation Model (Hassett 2010) that the equity risk premium is a function of risk free rates. Since 1960 the equity risk premium has been 1.9 – 2.48 times the risk free rate. The long term consistency of this relationship with loss aversion coefficients associated with Prospect Theory (Kahneman and Tversky, 1979) suggest it as a solution to the equity premium puzzle and support the experimental findings of Myopic Loss Aversion (Thaler, Tverseky, Kahneman and Schwartz, 1997).

Introduction

The equity premium puzzle in considered one of the most significant questions in finance. The term "equity premium puzzle" was coined by Mehera and Prescott in their 1985 paper, "The Equity Premium, A Puzzle,"[1] referring to the inability to reconcile the observed equity risk premium with financial models.

In the analysis, they use short-term treasuries as the risk free rate to calculate the real return on equities over numerous historical periods. They conclude that on average short-term treasuries have produced a real return of about 1% over the long-term, while equities have yielded 7%, implying a premium of about 6% or seven times the risk free return. Unable to reconcile a 7 x premium with financial models, they term it a puzzle.

Since then numerous papers have also attempted to explain the difference, including Shlomo Benartzi; Richard H. Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle"[2] which attempts to explain it in relation of loss aversion as first described in a paper by Daniel Kahneman and Amos Tversky in 1979.[3] They state:

> "The second behavioral concept we employ is mental accounting [Kahneman and Tversky 1984; Thaler 1985]. Mental accounting refers to the implicit methods individuals use to code and evaluate financial outcomes: transactions, investments, gambles, etc. The aspect of mental accounting that plays a particularly important role in this research is the dynamic aggregation rules people follow. Because of the presence of loss aversion, these aggregation rules are not neutral."

Our mental accounting for gains and losses determines how we perceive them.

Loss Aversion

Loss aversion refers to the fact that people are more sensitive to decreases in wealth than increases. Empirical estimates find that losses are weighted about twice as strongly as gains (e.g., Tversky and Kahneman (1992)[4]; Kahneman, Knetsch, and Thaler (1991)[5], Thaler, Tversky, Kahneman, Schwartz (1997)[6]). The pain of losing $100 is roughly twice the perceived benefit of gaining $100, so on average their subjects required equal odds of winning $200 to compensate for the potential loss of $100. In other words, the average subject required a gain of twice the potential loss to take a gamble that had equal chance of loss or gain. This was in stark contrast to the belief that people, as rational beings, evaluated the expected value and would be indifferent to a chance of gaining $100 to losing $100 if the odds were 50/50; if the gain were tilted to be slightly favorable they should take the bet. In reality, losing hurts more; people on average do not find the prospect of gaining $101 along with an equal chance of losing $99 to be an attractive wager. In their experiments, they found that subjects required about $200 to be willing to accept the 50/50 proposition of losing $100. Kahneman won the Nobel Prize in Economics in 2002 after Tversky passed away in 1996. Of course all people do not behave this way all the time, otherwise Las Vegas would not exist!

Loss Aversion and Corporate Decision Making

Incorporating loss aversion into financial thinking is in many ways a significant departure from how finance is often taught and practiced. In business school, I was taught to rely on net present value and expected value. A project with positive net present values should be pursued and that when faced with a range of outcomes, the expected value can be calculated by assigning probabilities to each outcome. The mantra: Pursue all NPV positive projects.

My experience has been that the business world rarely works this way. Due to corporate as much as individual loss aversion, decision makers are often much more risk averse, viewing the consequence of failure much greater than the rewards for success. Investments that have only slightly positive NPV or expected value are usually not pursued. Even the more risk tolerant individuals would tend to avoid risk if the organization takes a very dim view of loss.

This is why it is so important for organizations to employ incentive structures that reward sustainable growth in value and prudent risk taking. My own experience is that organizations without such incentives tend to be very risk averse. When decisions come down the internal calculus that investing successfully results in no reward, while failure results in unemployment or at least limited advancement, investment and growth are sure to slow. I would also argue that this also explains risk taking for traders on Wall Street where outsized rewards are given for success compared to the stigmas and punishments for failure. It's not that traders have high tolerance for risk, it's that in using OPM (Other People's Money) the penalty for failure is small.

Attempts to Solve The Equity Premium Puzzle

As discussed above, Mehra and Prescott(1985) coined the phrase "Equity Premium Puzzle" because they estimated that investors would require a very high coefficient of relative risk aversion (of the order of 40 or 50) to justify the observed equity risk premium of 7%. Mehra and Prescott revisited the topic two decades later with their 2003 paper, "The Equity Premium in Retrospect" where they continued to try and solve the puzzle by comparing real returns and ask whether the equity premium is due to a premium for bearing non-diversifiable risk. They conclude the answer is no unless you assume the individual has an extreme aversion to risk; many times higher than the 2x return seen in the lab.

They approach the problem using a general equilibrium model and compared short-term real risk free rates to observed equity premium. While I am not in a position to opine on the use of these models in evaluating equity premium, for several reasons I will discuss shortly, I believe that the use of short-term real rates is mistaken. I am not surprised they could not explain the rational for investors to such a dramatic disparity, since in my opinion they are not making the right comparison. Rather than using short-term real rates, they should be using long-term nominal rates.

What they did was a bit like measuring the speed of one moving vehicle from another moving vehicle. If Car A is moving at 60 mph and Car B is behind it at 66 mph and car C is next traveling at 61 mph, car C will see itself gaining on car A at just 1 mph. From the perspective of car C, car B is gaining on car A at a rate of 6 mph or 6 x faster than itself. This is all fine unless we care about their speed relative to a neutral observer who is not moving. Relative to the neutral observer, Car B is only going 10% faster than Car A.

Mehra and Prescott did not pick the right relative observation point. By using real returns they are measuring the difference from a moving vehicle. If we look at this from the perspective of real returns then the relative premium looks huge. But if we look at from the perspective of nominal returns, the neutral observer, then the premium it is not unreasonable. This is consistent with both the way individuals have been shown to evaluate gains and losses and with financial theory.

The mental accounting of investors focuses on the nominal returns. It's what investors track and how money managers are compensated. So it makes sense that that proper basis for evaluating the risk premium relative to the risk free rate is long-term nominal returns. For example, let's assume inflation is 2%. If an investor is considering a $1,000 investment with Treasuries at 4%, the yield is guaranteed to be $40 per year with a full return of principal. While the investor is exposed to interim fluctuations in value, the coupon and return of principal are guaranteed. Alternatively, the same investor considering an investment in the S&P 500 Index, would be evaluating the expected return relative to the nominal long-term rate rather than the real short term rate. In this case, expected equity returns of 10% would look good, yielding on average $100 per year rather than $40. If we calculate real returns by subtracting the 2% inflation, the $80 return for equities dwarfs the $20 for treasuries.

Now let's assume that expected inflation rises to 6% and the risk free rate jumps to 8%, so a new $1,000 bond would yield $80. If you applied the same 6% premium for equities, you get an expected yield of $140. Sure the real returns are the same, but doesn't the risky $140 look less attractive compared to a guaranteed $80?

Is it the right thing to track? Maybe not, but it is the reality. If investors compare their returns on equities to the nominal return of other investments, any attempt to explain the premium must compare the relative return as perceived by investors. Nominal not real returns should be used.

Long-term Treasury rates are used in determining cost of capital since they embody the market's best guess on long-term inflation. Even though this means they are not truly risk free, it is the best market estimate of expected interest rate and inflation risk; it is the right reference point. While it's true that using real equity returns accounts for the actual inflation component, it does not account for interest rate risk. In order account for expected inflation, most practitioners use long-term treasuries as the risk free rate. In doing so, they also incorporate a risk factor for interest rates.

Required return can be thought of as follows:

Nominal Equity Return = Real Equity Return + Inflation (1)
= Short-term Risk Free Rate + Inflation + Interest Rate Risk Premium + Equity Risk Premium (2)

If you subtract inflation from both sides to derive the real required return, you are still left with interest rate risk, which includes risk of unexpected inflation. So by using real equity returns and short-term risk free rate, you still have to account for the interest rate risk premium.

Real Equity Return = Short-term Risk Free Rate + Interest Rate Risk Premium + Equity Risk Premium (3)

Essentially, what Mehra and Prescott were calling the equity risk premium, was really the equity risk premium plus the interest rate risk premium.

Some believe that interest rates do not have a material impact on equity returns since inflation will result in earnings growth and since equities are priced as a multiple of earnings, as earnings grow equity prices increase with inflation. As I will discuss later, inflation has a huge impact on equity prices.

In "Myopic Loss Aversion and The Equity Premium Puzzle," Benzarti and Thaler (1995) they posit that the high degree of loss aversion is due to "myopic loss aversion" in that investors are sensitive to interim losses as equity markets fluctuate. They suggest that investors look at nominal returns since that is what is reported, therefore that's what investors look at. They find that a loss aversion factor of 2.25 to 2.78 is consistent with observed risk premiums if investors evaluate their portfolios about once a year and overall results are very sensitive to frequency of evaluation. In "The Effect of Myopia and Loss Aversion on Risk," Thaler, Tversky, Kahneman, Schwarts (1995), looked at this question through lab experiments found that subjects were more loss averse when they evaluated their returns more frequently and that they viewed guaranteed outcomes as a reference point with an evaluation period of about one year (13 months). In other words, investors evaluate their portfolios annually and expect a premium proportionate to the nominal risk free rate. As we will see below the RPF Valuation Model provides real world support for these findings.

Determining the Equity Risk Premium

In introducing the Risk Premium Valuation Model[7] (Hassett 2010), I posited that rather than being a fixed premium, the Equity Risk Premium fluctuates with the risk free rate, maintaining a constant proportionate relationship. The Equity Risk Premium equaled the Risk Free Rate times a constant factor. That factor (Risk Premium Factor) ranged from 0.9 – 1.48 between 1960 and today. So substituting into the formula where Cost of Equity = Rf + ERP,

Cost of Equity = Risk Free Rate + Risk Free Rate x Risk Premium Factor (RPF) (4)

Simplifying to:

Cost of Equity = Risk Free Rate x (1 + RPF) (5)

The RPF does not change frequently. In fact it has shifted only twice since 1960:

Period	RPF
1960 – 1980	1.24
1981 – Q2 2002	0.90
Q3 2002 – Present	1.48

Table 1: Estimated Risk Premium Factors

A Risk Premium Factor of 0.9 – 1.48, means Cost of Equity equals the Risk Free Rate times 1.9 – 2.48, very close to the findings on loss aversion factors.

The factor was determined by applying a set of simplifying assumptions to the constant growth formula:

$$P = E / (C - G) \text{ or } P/E = 1 / (C - G) \quad (6)$$

Variables and assumptions used are as follows:

P = Price (Value of S&P 500)

E = Actual Earnings (Annualize operating earnings for the prior four quarters as reported by S&P). Earnings, while not ideal, are used as a proxy for cash flow and seem to work very well

G = Expected long term projected growth rate, which is broken down into Real Growth and Inflation, so $G = G_R + I_{LT}$

G_R = Expected long-term real growth rate. Long-term expected real growth rate (G_R) is based on long-term GDP growth expectations on the basis that real earnings for a broad index of large-cap equities will grow with GDP over the long-term. A rate of 2.6% is used with the same rate applied historically.[8]

I_{LT} = Expected long-term inflation, as determined by subtracting long-term expected real interest rates (Int_R) from the 10 Year Treasury, where Int_R is 2%; based on the average 10 Year TIPs Yields from March 2003 – present.[9]

C = Cost of Capital is derived using Capital Asset Pricing Model, where for the broad market, $C = R_f + ERP$

R_f = Risk Free Rate as measured using 10 Year Treasury yields

ERP = Risk Premium Factor (RPF) x R_f

RPF = 1.24 for 1960 – 1980; 0.90 for 1981 – 2001; and 1.48 for 2002 – present. The RPF for each period was arrived at using a linear regression to fit the assumptions above to actual PE. All data used in the analysis is available for download at: http://sites.google.com/a/hassett-mail.com/marketriskandvaluation/Home

Including all assumptions, the formula reduces to:

$$P = E / (R_f \times (1+RPF) - (R_f - Int_R) - 2.6\%) \quad (7)$$

$$\text{Or } P/E = 1/ (R_f \times (1+RPF) - (R_f - Int_R) - 2.6\%) \quad (8)$$

The model explains stock prices from 1960 - 2009 with R Squared around 90%[10] to actual index levels from 1960 – 2009 as shown in graph below.

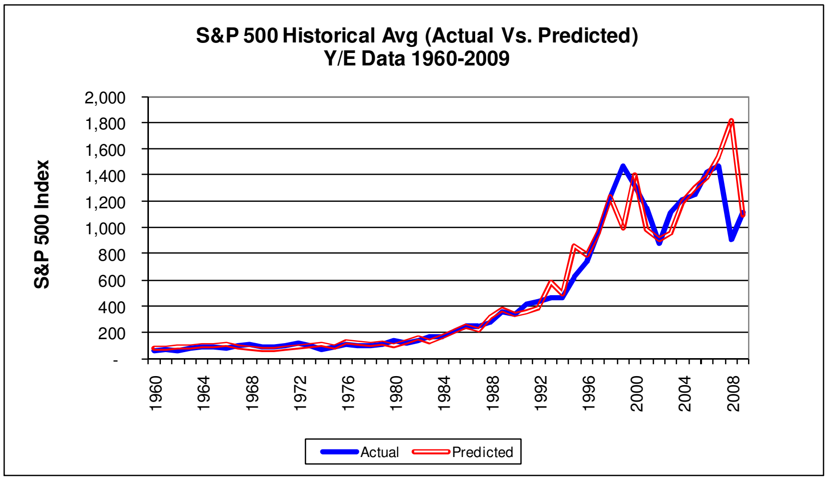


Figure 1: S&P 500 Actual vs. Predicted - 1960- 2009

The model only works if we assume that the Equity Risk Premium is conditioned on the Risk Free Rate, meaning that it gets bigger when the Treasury yields increase and smaller when they shrink. In fact one reason that I suspect many studies compared real returns, rather than nominal returns, may be the belief that inflation does not impact valuation. One common belief is that since profits will grow with inflation, inflation does not matter when discounted back. Another look at the constant growth equation can help understand this thinking:

$$\text{P / E} = 1 / (\text{C} - \text{G}), \text{ where} \quad (9)$$

$$\text{C} = \text{Rf} + \text{ERP} \quad (10)$$

$$\text{G} = \text{Real Growth} + \text{Expected Inflation} \quad (11)$$

$$\text{Rf} = \text{Real Interest Rate} + \text{Expect Inflation} \quad (12)$$

We can restate the equation for P/E as:

$$\text{P/E} = 1/ (\ (\text{Real Interest Rate} + \text{Expect Inflation}) - (\text{Real Growth} + \text{Expected Inflation}), \quad (13)$$

Expected Inflation is canceled out and:

$$\text{P/E} = 1/ (\text{Real Interest Rate} + \text{Real Growth}) \quad (14)$$

Since we assume the Real Interest Rate and Real Growth are a constant over the long term, P/E is also a constant. And, this would be true if the Equity Risk Premium were a constant. But if we assume that the Equity Risk Premium moves with the Risk Free Rate, then we get the relationship charted above, which is a very good fit with historical data.

Impact of Inflation on Value

Some argue that inflation should not have an impact on equity values, since higher costs can be passed on in the form of higher prices, so on average, earnings growth should keep up with inflation. If you assume P/E ratios should be a constant, say, 19 then with earnings of $2.00 share a company would trade at $38.00. With 5% inflation, earnings would grow to $2.10 and the share price to $39.90 – a gain of 5% which just matches inflation.

We get the same result using a constant growth model and a fixed Equity Risk Premium. Let's assume the Equity Risk Premium is 6%, the Risk Free Rate is 7%, which embodies 5% inflation, and real long term growth rate of 2.6%. Using the formula P/E = 1 / (C-G) we get, P/E = 1 / ((7%+6%) – (5%+2.6%) for a P/E of 18.5. If we lower the inflation rate to 2% the risk free rate drops to 4% and we calculate P/E = ((4%+6%)-(2%+2.6%) = 18.5. As shown earlier, any change inflation cancels itself out.

However, if we derive the Equity Risk Premium using the RFP Model, then the Equity Risk Premium varies with inflation. More inflation results in a higher risk premium. Using a 2% real interest rate, Table 2 below demonstrates the impact of inflation on P/E:

Inflation	R_f	ERP	Cost of Equity	G	Predicted P/E
2.0%	4.0%	5.9%	9.9%	4.6%	18.8
3.0%	5.0%	7.4%	12.4%	5.6%	14.7
4.0%	6.0%	8.9%	14.9%	6.6%	12.1
5.0%	7.0%	10.4%	17.4%	7.6%	10.2
6.0%	8.0%	11.8%	19.8%	8.6%	8.9

Table 2: Inflation Drives Valuation

Since investors expect a proportionately higher return over risk free, as inflation rises they apply a greater discount to future earnings, resulting in a lower present value, resulting in a lower multiple.

Back to Loss Aversion

We know that individuals have different tolerances for risk. If the RPF is 1.48, that implies the market as a whole has a loss aversion coefficient of 2.48. That is the average of all investors, not every individual. We would expect some to have lower coefficients and others higher. Gambling addicts destroy their own lives, knowing the odds are not better than even, implying a loss aversion coefficient of less than 1.0**.** Likewise, some people are more risk averse than average. This is one of the factors that act to set price.

The prices for individual stocks are set at the margin. For example, Google closed today at $476 and traded about 2.5 million shares. But with 320 million shares outstanding, that is less than 1%. The price is set by the investors trading that 1%. The implication is that the owners of the remaining 99% think Google is worth more than the current $476 and some number of investors would be will to buy Google at a lower price. Mechanically the way this works is that sellers offer to sell a number of shares at a certain price, called the Ask, and potential buyers offer to buy at a specified price, called the Bid. The Bid for Google might be 200 shares at $476.07 and the Ask 700 shares at $476.18. The difference, $0.11 in this case, is called the Bid-Ask spread. These are the current best offers to buy and sell. For high volume stocks like Google, the Bid-Ask spread is small, just 0.02% in this case. For lower volume equities the spread will generally be higher.

If an investor places a marker order to, say, buy 500 shares, the first 200 shares will be filled at the current Bid price for 200 shares at $476.17. The remaining 300 shares will be filled by the next best ask price, which will be $476.17 or higher. It is not the consensus or average estimate of value that determines the price, but the price at which investors at the margin are willing to buy or sell at any moment. So if I don't own shares of Google and I think it's worth just $400 or even $100, I am not a factor in setting the price. But if in the moment described above, I enter a bid for 200 shares at $476.18, the order is immediately filled and, for that moment, I am the price setter.

Similarly, investors with loss-aversion coefficients at the extremes should not be expected to have much market impact. An investor with a loss aversion coefficient well above 2.5 will be risk averse and have portfolio skewed towards government bonds, while and investor with a loss aversion coefficient near 1.0 will always have a portfolio that is mostly equities. Therefore neither will have much impact on price setting. On the other hand, investors with loss aversion coefficients around 2.5 will be more likely to be shifting their portfolios between bonds and equities and have a larger impact on pricing.

Conclusion

Loss aversion is hard wired into us and drives a number of decision processes that seems to include how investors set prices in the stock market. Thaler, Tversky, Kahneman, Schwarts (1995) found evidence of what they called Myopic Loss Aversion and demonstrated the expectations of risk premiums were consistent experimental findings for loss aversion if portfolios were evaluated annually. The Risk Premium Factor Valuation Model (Hassett 2010) provides real world evidence that the market actually behaves this way. Combing evidence that the risk premium varied with the risk free rate in a proportion consistent the findings in behavioral studies, suggests that Loss Aversion is the answer to the equity premium puzzle.

Endnotes

[1] Mehra, Rajnish, and Edward C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economies*, vol. 15, no. 2 (March):145–161.

[2] 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.

[3] Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica*, vol. 46, no. 2 (March):171–185.

[4] Kahneman, D., A. Tversky. 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty*, 5:297-323.

[5] Kahneman, Daniel., Jack L. Knetsch and Richard.H. Thaler, 1991." The endowment effect, loss aversion, and status quo bias." *Journal of Economic Perspectives,* vol. 5, no. 1 (Winter): 193-206

[6] 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

[7] Hassett, S. D. (2010), The RPF Model for Calculating the Equity Market Risk Premium and Explaining the Value of the S&P with Two Variables. *Journal of Applied Corporate Finance*, 22: 118–130. http://bit.ly/d6b1Py

[8] "Economic Projections and The Budget Outlook." Whitehouse.gov, Access Date March 15, 2009, http://www.whitehouse.gov/administration/eop/cea/Economic-Projections-and-the-Budge-Outlook/

[9] "H.15 Selected Interest Rates", The Federal Reserve Website, Accessed March-July 2009, http://www.federalreserve.gov/datadownload/Choose.aspx?rel=H.15

[10] See Hassett (2010)

The RPF Model for Calculating the Equity Market Risk Premium and Explaining the Value of the S&P with Two Variables

by Stephen D. Hassett, Hassett Advisors

While driving increases in shareholder value is one of the most important responsibilities of any business leader, many executives are handicapped by their limited understanding of what drives value. And they are not alone. Even prominent economists say that stock market valuation is not fully understood. For example, in a 1984 speech to the American Finance Association, Lawrence Summers said,

It would surely come as a surprise to a layman to learn that virtually no mainstream research in the field of finance in the past decade has attempted to account for the stock-market boom of the 1960s or the spectacular decline in real stock prices during the mid-1970s.[1]

Some people see the stock market as arbitrary and random in setting values. But despite occasional bouts of extreme volatility (including, of course, the recent crash), most academics (and many practitioners) would likely agree with the proposition that the market does a reasonably good job of incorporating available information in share prices. At the same time, however, certain factors can clearly cause the market to misprice assets. These include problems with liquidity, imperfect information, and unrealistic expectations that can knock valuations out of line for a period of time. But such limitations notwithstanding, over a longer horizon the market appears to be reasonably efficient in correcting these aberrations.

The RFP Valuation Model introduced in this article is intended to explain levels and changes in market values and, by so doing, to help identify periods of likely mispricing. As such, the model offers a general quantitative explanation for the booms, bubbles, and busts—that is, the series of multiple expansions and contractions—that we have experienced over the past 50 years. The model explains stock prices from 1960 through the present (March 2010), including the 2008/09 "market meltdown." And it does so using a surprisingly simple approach—one that combines generally accepted approaches to valuation with a simple way of estimating the Market or Equity Risk Premium (ERP) that produces remarkably good explanations of market P/E ratios and overall market levels.

To show you what I mean, Figure 1 shows how the P/E ratio predicted by model, when applied to S&P Operating Earnings, explains levels of the S&P 500 over the past 50 years, the earliest date for which I had reliable earnings data.

My approach to estimating the Equity Risk Premium is the most original part of this overall hypothesis. Many if not most finance theorists have assumed that the Equity Risk Premium is a constant that reflects the historical difference between the average return on stocks and the average return on the risk-free rate (generally the return on the 10-year U.S. government bonds). But if we also assume that long-term real interest rates do not change and that real growth can be approximated by real long-term GDP growth (also generally assumed to be stable), then the market-wide P/E would also be absolutely constant over time.

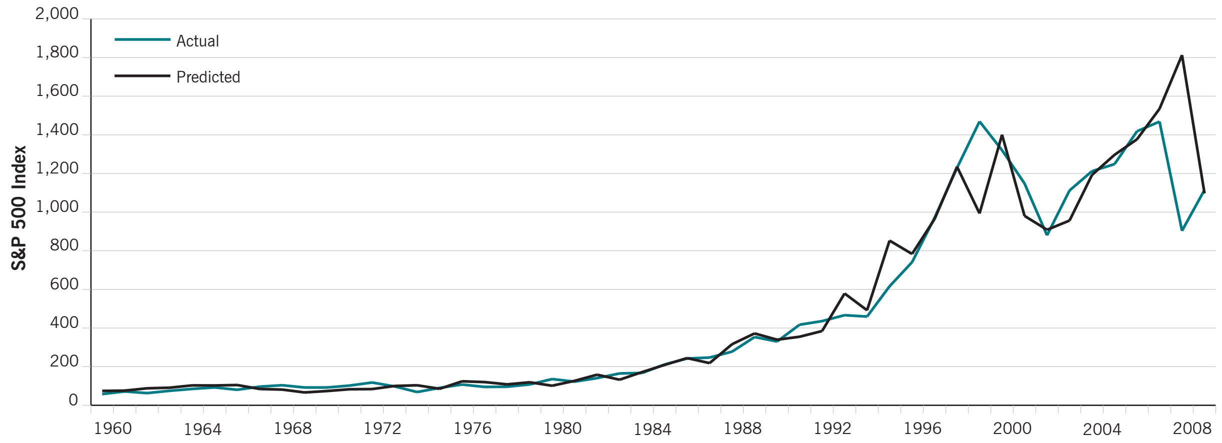
But, of course, the P/E multiple on the earnings of the S&P 500 is volatile, with year-end values ranging from 7.3 in 1974 to 29.5 in 2001. One possible objection to the idea of a constant risk premium is its implication that, when the risk-free rate increases, investors are satisfied with a premium that is smaller as a proportion of the risk-free rate. In this article, I suggest that the Equity Risk Premium is not a fixed number but a variable that fluctuates in direct proportion to the long-term risk-free rate as a fixed percentage, not a fixed premium. When used with the constant growth model, the cost of capital can be determined by the following formula:

$$\text{Equity Risk Premium} = \text{Risk-Free Long-Term Rate} \times \text{Risk Premium Factor} \qquad (1)$$

This relationship can be used to explain why and how the risk premium varies over time; as interest rates vary, so does the risk premium. This Risk Premium Factor (RPF) appears to have held steady for long periods of time, changing just twice during the 50-year period from 1960 to the present (July 2009). Based on my calculations, the RPF was 1.24 from 1960-1980, 0.90 from 1981-June 2002, and 1.48 from July 2002 to the present. As we saw earlier in Figure 1, the model does a very good job of predicting market levels, even through the present financial crisis.

1. Quoted by Justin Fox, *The Myth of the Rational Market: History of Risk, Reward and Delusion on Wall Street*, p. 199. (Harper Collins, New York, 2009).

Figure 1 **S&P 500 Actual vs. Predicted—1960–2009**



This result is also consistent with investor "loss aversion," the well-documented (by Kahneman and Tversky) willingness of investors to sacrifice significant gains to avoid considerably smaller losses. One of their studies produced a loss aversion coefficient of 2.25,[2] which implies that participants, on average, would be indifferent to the outcome of a coin flip promising either an expected but uncertain $325 or a guaranteed $100. The analogous calculation for the RPF model suggests that if the risk-free rate were 4% and the RPF 1.48, investors contemplating a $1,000 investment would assign roughly equal value to a guaranteed (bond-like) $40 and equities with an expected return of $99.

Valuing Constant Growth

The place to start is with the simplest valuation model, the Constant Growth Equation. This model derives from, and represents a specific case of, the Discounted Cash Flow (DCF) model that is used to determine the net present value of a projected stream of future cash flows. In the case in question, it is a perpetual stream of cash flows with a constant rate of growth. Instead of assuming different levels of earnings in each period, it assumes a constant growth rate off the base year and a constant cost of capital.

The DCF model can be expressed as follows:

$$P = \sum E_1 / (1+C)^1 + E_2 / (1+C)^2 + \ldots + E_n / (1+C)^n \qquad (2)$$

where E is cash flow and C is cost of capital. If you assume that E grows at a constant rate (G),

$$P = \sum (E_0 \text{ x } (1 + G)^1) / (1+C)^1 + (E_0 \text{ x } (1 + G)^2) / (1+C)^2 + \ldots + (E_0 \text{ x } (1 + G)^n) / (1+C)^n \qquad (3)$$

the result simplifies to:

$$P = E / (C - G) \qquad (4)$$

This equation, which is not so much a theory as an indisputable mathematical concept, is the expanded form of the core insight that the value of a perpetual stream is the amount of the payments divided by the required rate of return. In other words, the value of a guaranteed $100 perpetual annuity in a market where the long-run risk-free return is 10% is $1,000 ($100/.10).

The next step is to take the constant growth version of this model (equation 4) and apply it to market valuation by substituting S&P operating earnings for the variable E above.

P = Price (Value of S&P 500 Index)

E = Earnings (Reported operating earnings for the prior four quarters as reported by S&P) as a proxy for cash flow

G = Expected long term growth rate

C = Cost of equity capital

This formula can also be restated to predict the Price-Earning (P/E) ratio of the S&P 500 as follows:

$$P/E = 1 / (C - G) \qquad (5)$$

2. Daniel Kahneman and Amos Tversky, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5 (1992):297-323.

Table 1 **Growth Drives P/E**

Long-term Growth	Predicted P/E
0%	12.6
2%	16.7
4%	25
6%	50

These two equations, when used with the right assumptions (as discussed below) can be helpful in understanding the valuations of both individual companies and the overall market.

Some academics and practitioners argue that equity should be valued as the present value of not earnings or cash flows, but of the dividend payments actually made to shareholders—an argument that is embodied in the Gordon (or Dividend) Growth Model. Some proponents of this model advocate a modified approach that values all corporate distributions, share repurchases as well as dividends. One well-known advocate of this model is Nobel Laureate Paul Krugman, who wrote:

Now earnings are not the same as dividends, by a long shot; and what a stock is worth is the present discounted value of the dividends on that stock—period, end of story.[3]

I disagree, and for several reasons. For starters, Modigliani and Miller demonstrated in their famous 1961 article on the "irrelevance" of dividend policy, that it is the underlying expected earnings power of companies, not their dividend payouts, that determine corporate market values.[4] Dividend policy is as much a reflection of a company's capital structure and investment opportunity set as of its expected future profits—and decisions to pay out capital may often reflect a maturing of the business and a scarcity of profitable investment opportunities. What's more, most promising growth companies pay no or minimal dividends—and certainly for those companies, the current levels and changes in earnings are likely to be more reliable indicators than dividends of future profitability.

Why Growth Rate and Cost of Capital Matter—Lessons from the Constant Growth Equation

Assume you have an asset with a cost of capital of 12%, a growth rate of 2% and cash flow of $100. Using the Constant Growth model, the value can be calculated as follows: $100 / (12% - 2%) = $1,000. This might be called the "intrinsic value" of the asset and, as such, it offers the best guide to what it should trade for.

We can also apply this model to a share of stock to determine its intrinsic value. In place of cash flow, we use earnings per share (EPS) of $2.00 with the same cost of capital and growth rate, and the result is $2.00/(12% - 2%) = $20.00. Since EPS is $2.00 and price is $20.00, the Price to Earnings Ratio (P/E) is $20/$2 or a P/E of 10. While the market may value it differently, if these assumptions are true, this formula tell us its intrinsic value.

P/E ratios are often used to assess whether share prices are expensive or cheap. A P/E of 8 is considered very low, but when Google had a P/E of 60 or more, some thought it was very high. Is a company with a P/E of 10 a bargain compared to a company with a P/E of 20? We can explore this question using the constant growth equation.

Take the same company and now assume that its cost of capital drops to 8%, its growth rate increases to 3%, and its earnings stay the same. These might seem like small changes, but their impact is dramatic: $2.00/(8% - 3%) = $40.00, a doubling of value with the P/E rising to 20. If growth increases to 5% (in line with nominal long-term GDP growth), the share price rises to $66, and the P/E is 33. (For additional examples of how P/E varies based on growth for a company with an 8% cost of capital, see Table 1.)

The formula P = E / (C – G) shows that earnings relate directly to price. What many managers fail to realize is that investors don't look at earnings in a vacuum; they parse the information in earnings in order to estimate growth. And that's why the reporting of earnings often causes the P/E to change.

So, for all its simplicity, the Constant Growth model has some important lessons:

1. Small changes in growth make a big difference in value
2. Cost of capital is important, so we better get it right
3. Earnings drive value (stock price) but also contain information

While it may not be difficult to project current earnings, the big challenges are forecasting growth and getting the right cost of capital.

A Short Overview of Risk Premiums

The Capital Asset Pricing Model (CAPM) can be used to determine the cost of equity for an individual firm or the market overall. The model takes the form of the following equation: **Cost of Equity = R_f + β x (ERP),** where R_f = Risk-Free Rate (and we will use the yields on 10-year Treasuries as a proxy); β = Beta, which measures the sensitivity of the stock to market risk (which, by definition, is 1.0 for the entire

3. Krugman, Paul, "Dow 36,000: How Silly Is It?", *The Official Web Page of Paul Krugman*, accessed August 2009, http://web.mit.edu/krugman/www/dow36K.html.

4. Franco Modigliani, Merton H. Miller, "Dividend Policy, Growth, and the Valuation of Shares," *Journal of Business*. 1961, vol. 34, no. 4.

Table 2 **ERP Drives Valuation**

R_f	ERP	Cost of Equity	GDP + Inflation	Predicted P/E
5%	3%	8%	5%	33
5%	4%	9%	5%	25
5%	5%	10%	5%	20
5%	6%	11%	5%	17
5%	7%	12%	5%	14

market); and ERP = Equity Risk Premium (the calculation of which will be the main subject of this discussion). Given that the Beta of the broad market is 1.0, the Cost of Equity for the market as a whole can be expressed as $C = R_f + ERP$.

While the risk-free rate is easily determined, the risk premium is not. In fact, there is no clear consensus on how this should be done. The Equity Risk Premium (ERP) is the expected return an investor requires above the risk-free rate for investing in a portfolio of equities. It makes sense that if 10-year Treasury yields represent the safest (risk-free) long-term investment, then investors will require higher expected rates of return to buy riskier securities like corporate bonds or equities. My own considerable experience in valuing businesses has made it clear to me how sensitive valuations can be to one's estimate of the ERP (a topic I return to later).

The most common way of estimating the ERP is to measure the historical premiums that investors have received relative to Treasury yields and assume that investors will expect that rate of return in the future. Depending on method and time-period, this can range from 3% to 7% or more. Other methods include surveys and forward-looking estimates based on current stock market levels. There is a huge body of research on measuring equity risk premiums. Indeed, entire books have been written on the subject.

Many researchers have argued that the Equity Risk Premium changes over time—and that such fluctuations are a major source of stock price changes—and also that the ERP has experienced a "secular" decline during the past few decades. In their book *Dow 36,000*, for example, Kevin Hassett (no relation) and James Glassman pushed this argument to its reduction ad absurdum when suggesting that the risk premium could vanish entirely since, given a sufficient amount of time, stocks appeared virtually certain to outperform bonds.[5] In *The Myth of the Rational Market*, Justin Fox quotes Eugene Fama, one of the pioneers of the efficient market hypothesis, as saying, "My own view is that the risk premium has gone down over time basically because we've convinced people that it's there."[6] Roger Ibbotson, a well-known compiler of ERP statistics, has suggested that the recent decline in the risk premium should be viewed as a permanent, but non-repeating event, "We think of it as a windfall that you shouldn't get again," he said.[7]

The Effects of Risk Premium on Valuation

Table 2 shows the expected effects of differences in ERP (ranging from 3% to 7%) on valuations and P/E ratios. Using the constant growth model, P/E = 1 / (C – G), if we assume that the market will grow with long-term estimates of real GDP at 3% plus long-term inflation at 2%, our estimate of stock market P/E would have P/E = 1 / (C – 5%). (Note: Real GDP + Inflation is Nominal GDP). With Treasury yields at 5%, and ERPs ranging from 3%-7%, our range of cost of capital (R_f + ERP) is from 8% to 12%. Table 2 also shows the P/E implied for the overall market given this range of estimates of ERP and cost of capital. To provide some perspective on these numbers, if the S&P 500 were at 1,200 with its current P/E of 19, it would increase more than 25% to 1,593 with a P/E of 25 and the same level of earnings!

A New ERP Theory: The Risk Premium Factor (RPF) Model

Conventional theory says that if the Equity Risk Premium were 6.0% and 10-year Treasury yield was 4.0% then investors would expect equities to yield 10%. The theory also implies that if the 10-year Treasury was 10%, then investors would require a 16% return, which represents a proportionally smaller premium.

For reasons discussed below, I will argue that investors expect to earn a premium that is not fixed, as in the conventional CAPM, but varies directly with the level of the risk-free rate in accordance with a "Risk Premium Factor" (RPF). While this proportional RPF is fairly stable, it can and does change over longer periods of time.

To illustrate the concept, with an RPF of 1.48, equities are expected to yield 9.9% when Treasury yields are at 4.0%. But if Treasury yields suddenly rose to 10%, equities would have to return 24.8% (10 + 1.48 x 10 = 24.8) to provide investors with the same *proportional* compensation for risk. In this example, an increase in interest rates (and inflation) causes the risk premium to jump from about 6% to 15%, suggesting that interest rates have a greater impact on valuation and market price than is generally recognized.

To test this approach, we must determine not only the

5. James K. Glassman and Kevin A. Hassett, *Dow 36,000: The New Strategy for Profiting From the Coming Rise in the Stock Market*, (Times Business, New York, January 1, 1999).

6. Justin Fox, *The Myth of the Rational Market: History of Risk, Reward and Delusion on Wall Street*, p. 263. (Harper Collins, New York, 2009).

7. Ibid.

Figure 2 **10-Year Treasury Yields—1960–2009**[12]

Source: U.S. Treasury

Risk Premium Factor, but estimates for the other variables in the following equation:

$$P/E = 1 / (C - G) \quad (11)$$

In the analysis that follows, I use the following variables and assumptions:

- P = Price (Value of S&P 500)
- E = Actual Earnings (Annualized operating earnings for the prior four quarters as reported by S&P). Earnings, while not ideal, are used as a proxy for cash flow and seem to work very well
- G = Expected long-term projected growth rate, which is broken down into Real Growth and Inflation, so $G = G_R + I_{LT}$

G_R = Expected long-term real growth rate. Long-term expected real growth rate (G_R) is based on long-term GDP growth expectations on the basis that real earnings for a broad index of large-cap equities will grow with GDP over the long-term. A rate of 2.6% is used with the same rate applied historically.[8]

- I_{LT} = Expected long-term inflation, as determined by subtracting long-term expected real interest rates (Int_R) from the 10-year Treasury, where Int_R is 2%; based on the average 10-year TIPs Yields from March 2003 to the present.[9]
- C = Cost of Capital is derived using Capital Asset Pricing Model, where for the broad market, $C = R_f + ERP$
- R_f = Risk-Free Rate as measured using 10-year Treasury yields
- ERP = Risk Premium Factor (RPF) x R_f
- RPF = 1.24 for 1960 – 1980; 0.90 for 1981 – 2001; and 1.48 for 2002 – present. The RPF for each period was arrived at using a linear regression to fit the assumptions above to actual PE.[10]

When using these assumptions for the present period—that is, with an RPF of 1.48—the formula reduces to:

$$P/E = 1/ (R_f \times (1+RPF) - (R_f - 2\%) - 2.6\%) \quad (12)$$

Explanatory Value of the RPF Valuation Model

As can be seen in Figures 2-6, the actual values deviated significantly from the predicted values at the end of 2008 and the first quarter of 2009, but had returned to something like parity by June 2009. I believe that these deviations from the model were attributable mainly to the abnormally low yields for 10-year Treasuries that had been in effect since late 2008, when the "flight to quality," along with the Federal Reserve's purchase of notes beginning in March 2009, caused the 10-year Treasuries to be overpriced.[11] As shown in Figure 2, yields then fell to as low as 2.2%, as compared to a more "normal" range of 4.1% to 5.1% in 2006 and 2007 (and rarely

8. "Economic Projections and The Budget Outlook," Whitehouse.gov, Access Date March 15, 2009, http://www.whitehouse.gov/administration/eop/cea/Economic-Projections-and-the-Budge-Outlook/.
9. "H.15 Selected Interest Rates", The Federal Reserve Website, Accessed March-July 2009, http://www.federalreserve.gov/datadownload/Choose.aspx?rel=H.15.
10. All data used in the analysis is available for download at: http://sites.google.com/a/hassett-mail.com/marketriskandvaluation/Home.
11. "Fed in Bond-Buying Binge to Spur Growth," *The Wall Street Journal Online*, March 19, 2009, http://online.wsj.com/article/SB123739788518173569.html.
12. H.15 Selected Interest Rates", The Federal Reserve Website, accessed March-January 2010, http://www.federalreserve.gov/datadownload/Choose.aspx?rel=H.15.

Figure 3 **S&P 500 P/E Actual vs. Predicted—1960–2009 (Annual)**

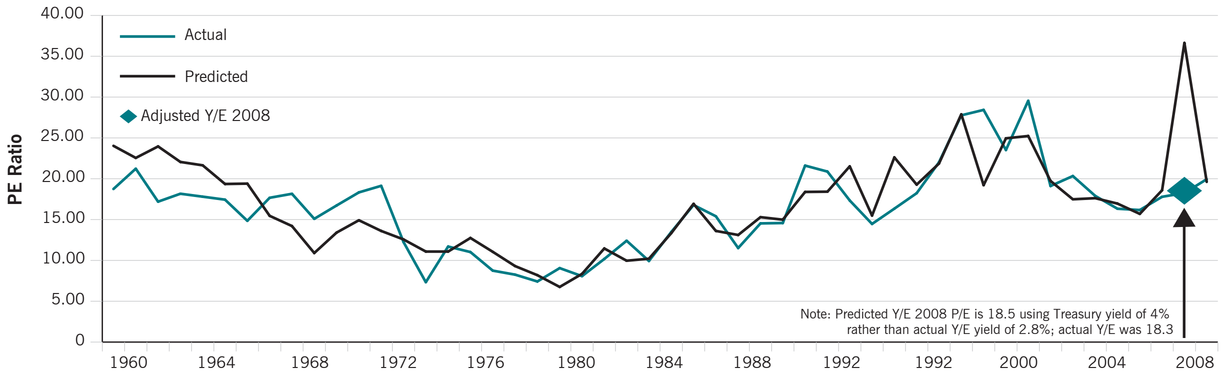
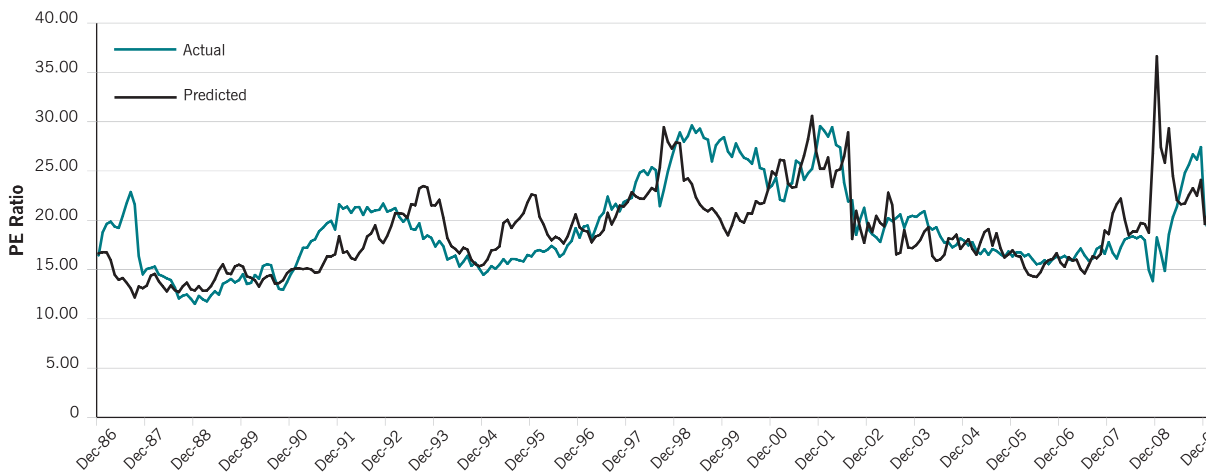


Figure 4 **S&P 500 P/E Actual vs. Predicted—1988–March 2010 (Monthly)**



less than 4% since 1960).

To compensate for these abnormally low Treasury yields Figure 3 shows the P/E ratios that would likely have prevailed if Treasury yields had remained at a still low, but more normal yield of 4%.[13] And as shown in each of Figures 3-5, when we normalize the 2008 R_f variable in this way, the actual year-end valuations correspond closely with the predicted values. One use of the model is to spot anomalies—and I believe that Treasury yields during the 2008/09 financial crisis were an anomaly.

Also plainly visible in Figure 3 is the decline in P/E ratios in the 1970s, reflecting the increase in interest rates during that period. It also shows the jump in P/Es during the 1980s, reflecting the drop in inflation and interest rates.

Figure 4 shows the application of the same model using monthly data from the end of 1986 through March 2010.[14] Like Figure 3, Figure 4 shows the return of values to parity by middle of 2009. And as can be seen in Figure 5, the RPF model explains overall market valuation levels when actual S&P operating earnings are applied to the P/E ratio during the period 1960–2009.[15] Using both year-end annual data for the past 50 years and monthly data for the past 20 years, then, the RPF model appears to do a very good job explaining valuations. And that in turn would suggest that, at any

13. While earnings are released quarterly, the model was extended to monthly and daily price data by using actual closing prices for S&P 500 and 10-Year Treasury yields along with S&P 500 operating earnings as a constant for each month in the quarter. The quarterly earnings were applied for the month preceding quarter end (i.e., Dec – Feb = Q1) under the assumption that market expectations would have incorporated earning expectations. Again, it assumed that as the end of quarter approaches earnings estimates should be within a reasonably close to those actual earnings ultimately reported and embodied in share prices. Earnings and S&P Averages 1960-1988 from Damodaran Online: Home Page for Aswath Damodaran (New York University) http://pages.stern.nyu.edu/~adamodar/; S&P Earnings and levels from 1988 – Present from Standard and Poors Website, http://www2.standardandpoors.com/portal/site/sp/en/us/page.topic/indices_500/2,3,2,2,0,0,0,0,0,1,5,0,0,0,0,0.html; Calculations and methodology by the Author.

14. See Note 13.

15. See Note 13.

Figure 5 **S&P 500 Actual vs. Predicted—1988–March 2010**

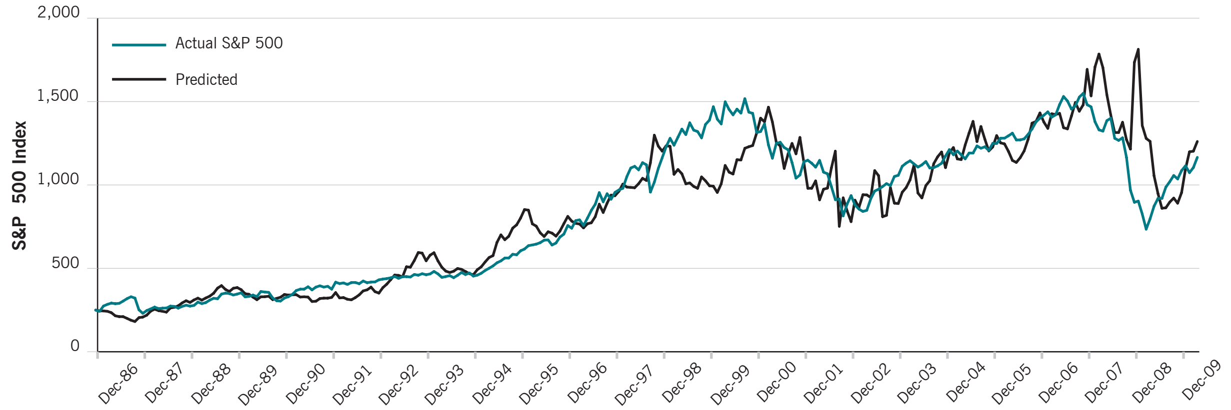


Table 3 **Estimated Risk Premium Factors**

Period	RPF
1960 – 1980	1.24
1981 – Q2 2002	0.90
Q3 2002 – Present	1.48
6%	50

Table 4 **RPF Valuation Model R Squared Results**

	R Squared	
Dataset	Full Dataset	Excluding Meltdown
1960 – 2008 (Annual)	89.5%	96.3%
1986 – September 2009 (Quarterly)	80.6%	88.0%
January 1986 – September 2009 (Monthly)	86.3%	90.8%
January 1986 – September 2009 (Daily)	86.5%	90.9%

point in time, the general level of market pricing and P/E ratios are driven mainly by just two factors: interest rates and expected earnings.

Estimating the Risk Premium Factor (RPF)

The RPF was estimated by fitting the model to actual levels of the S&P 500 over the period 1960 to the present. This analysis revealed two distinct shifts in the RPF since 1960. Table 3 shows the RFP factors that provide the best fit for each period.

The overall fit was assessed by calculating the R^2s of the regressions using the appropriate RPF for each time period. As previously discussed, the meltdown after September 2008 drove down the risk-free rate to an unsustainable level and left a trail of historical earnings that clearly did not reflect expectations. As also discussed previously, these factors are now back in line. To adjust for this recent anomaly, the R^2 was calculated excluding meltdown time period beginning September 2008.

As reported in Table 4, after excluding the meltdown period, the RPF Valuation Model explains a remarkably high 96% variation of stock prices over the past 50 years, as well as 91% of the daily variation.[16]

Consistency with Prospect Theory/Loss Aversion

As mentioned earlier, Daniel Kahneman and Amos Tversky first developed "prospect theory" in 1979, proposing that individuals have a sufficiently strong preference for avoiding losses that they are willing to pass up considerably larger gains. (Kahneman won the Nobel Prize in Economics in 2002 after Tversky passed away in 1996.) Such "loss aversion" in turn causes individuals to seek compensation for risk that is greater than what would be indicated by expected value of the outcomes. For example, if you were offered a certain $100 or $201 for correctly guessing a coin flip, you *should* prefer the coin flip. Not surprisingly, most people require higher levels of compensation to take the bet.

Numerous studies have been conducted to determine how much additional compensation is required; this is called the loss aversion coefficient. In a 1992 study, Kahneman and

16. For daily calculation, actual closing prices for S&P 500 and 10-Year Treasury are used; daily earnings were derived using same approach as monthly earnings as explained in Note 13.

Tversky reported finding a coefficient equal to 2.25.[17] In other words, people on average were indifferent to a coin flip for $325 versus a guaranteed $100. Other studies found coefficients of loss aversion in the range of 1.43 to 4.8.[18]

Such coefficients are consistent with my RPF findings, in which equities require premiums ranging from 90% to 148% over 10-year Treasury yields (roughly equivalent to loss aversion coefficients between 1.90 and 2.48). And the two concepts appear to have another important similarity. Stock market investors, like the subjects in these studies, appear to expect an incremental return for bearing risk that increases proportionally with the level of the risk-free interest rate. For example, if you were indifferent between $10 guaranteed and $30 on a coin flip, you probably would not accept that same fixed $20 premium over the expected value if the stakes were raised and you were offered a choice between a certain $100 and a contingent $220. Likewise, if the risk-free rate is 4% and the RPF is 1.48, a $1,000 investment in bonds would offer a guaranteed $40 and equities an expected return of $99, or a $59 premium. But if bonds instead yielded 10% and the guaranteed return rises to $100, a $59 premium would probably look much less attractive.

Potential Causes for Shifts in The Risk Premium Factor (RPF)

The RPF has shifted twice in the past 50 years, once in 1981 and again in July 2002. The period from 1960-1981 was characterized by increasing inflation expectations, rising from 1.8% in 1960 to 11.7% in 1981.[19] In 1981, the trend reversed and inflation expectations began to decline. The 1981 shift in RPF from 1.24 to 0.90 could have resulted from this change in inflation expectations driven by world events, with the decline in inflation resulting in higher real after-tax equity returns. Events during 1981 that could have contributed this change include:

• Resolution of the Iran hostage crisis. The reduction of tensions could have increased expectations of stability and a secure oil supply bringing with it lower inflation and less risk of an economic shock.[20]

• Inauguration of the Reagan era, with tax reduction leading to higher real after-tax returns.

At the same time, my analysis shows that the RPF increased from 0.90 to 1.48 in mid-2002. The decline of the rate of long-term inflation ended in 2002, with long-term inflation expectations having declined from a peak of 11.7% in 1981 to 2.0% in 2002. From 2002–2008, the rate of inflation has remained fairly stable, fluctuating in the 2% - 3% range. Other events that could have caused or contributed to the shift in 2002 include:

• Department of Justice investigation into Enron. Enron, Tyco and WorldCom's destruction of confidence in reported earnings may have led to increase risk premium factor.

• The enactment of Sarbanes Oxley in response to accounting scandals. The act faced severe criticism for imposing significant costs on public companies. Some suggested high compliance costs would cause capital to flee to less regulated markets, increasing the premium required for U.S. equities.

• Congressional authorization of war in Iraq. Expectations of a protracted war with Iraq could have increased expectations that increased borrowing to fund the war would lead to increased inflation and tax rates in the future.

Potential Weaknesses in RPF Theory and Methodology

Proper application of the model requires an understanding of its potential weaknesses:

• *All data points are current actual or historical.* While the market is forward looking, all data in the analysis are based on actual results. Even 10-year Treasury yields, which embody expectations about future real interest and inflation, were sampled at a single point in time, along with earnings that are not released until well after the quarter ends. Analysts' estimates are widely accepted as being embodied in current share price and would be expected to be reasonably close to actual before the end of each quarter.

• *Reasons for changes in Risk Premium Factor (RPF) are not fully explained.* The RPF has changed twice over the past 50 years and has historically held for long periods of time. While I have suggested a few possible reasons for the two changes in the RPF over the past 50 years, it is clear that further explanation and understanding is necessary.

• *The RPF may seem to be set arbitrarily to fit actual.* Given the good linear regression fit across a relatively large number of data points, the RPF seems to make sense and provide good result. Nevertheless, this remains a valid concern.

• *RPF cannot be projected.* Thus far it only seems possible to discern the RPF with hindsight. Still this would seem superior to other methods for determining risk premiums that produce less definitive results. For example, if the RPF changed just two times over 50 years, one might argue that in any given year there is a 96% chance (48 out of 50) that the RPF will remain constant over the next year.

17. Kahneman and Tversky. (1992), cited earlier.

18. Abdellaoui, Mohammed, Bleichrodt, Han and Paraschv, Corina, Loss Aversion Under Prospect Theory: a Parameter-Free Measurement (October 2007). Management Science, 10:1659-1674.

19. Calculation of inflation expectations based on difference between 10-Year Treasury yield and assumed 2% long-term real interest rate

20. "1981: Tehran frees US hostages after 444 days" BBC Website, Accessed March 15, 2009, http://news.bbc.co.uk/onthisday/hi/dates/stories/january/21/newsid_2506000/2506807.stm.

Declining Interest Rates Explain More than Half of S&P 500 Index Growth Since 1981

Interest rates are much more important than is generally recognized. Some contend that the effects of interest rates on corporate values are limited to the direct impact on corporate borrowing and consumer spending. Such observers tend to argue that although the cost of capital rises with inflation, for the market as a whole, the negative effect of this increase is directly offset by the positive effects of inflation on earnings. In other words, in the equation $V = E / (C - G)$, since C and G increase by the same amount (inflation), the expected impact of inflation is zero.

By contrast, the RFP Model suggests that since the ERP increases proportionally with the risk-free rate, it rises faster than the growth in earnings, causing a decline in valuations. So, in addition to the direct negative impact of interest rates on earnings, higher rates also have a large impact on P/E multiples.

The highest monthly finish of the S&P 500 was October 2007, when it closed at 1549. The highest annual finish of the risk-free rate was 1981, when the 10-year Treasury yield ended the year at 13.7%. Between these two mileposts, the S&P 500 Index increased 1264%, from 122 to 1549. During the same period, S&P Operating Earnings increased only 588%, rising from 15.2 to 89.3. Thus, earnings accounted for only 47% (588%/1264%) of the growth of the S&P 500 during this period.

And since the increase in S&P earnings account for less than half of the increase in its value, much of the remaining increase can be attributed to decreases in the risk-free rate—and with the 10-year Treasury yields falling to 4.47% in October 2007, the cost of capital dropped from over 26% at the end of 1981 to about 11% in 2007. And according to the RPF model, over 50% of the appreciation over the past 29 years is explained by reductions in both the RPF and risk-free rate. More specifically, the model provides a way of explaining the remarkable increases in corporate P/E multiples since the 1960s—one that relies largely on changes in interest rates (which embody expected inflation) during that period.

The RPF Model and Market Efficiency: Exploring Major Market Events From 1986–2009

The RPF Model can help demystify valuation and also help explain major market vents over the past 20 or so years. The exploration of these events may also serve to shed some light on the efficient market hypothesis.

The Efficient Market Hypothesis (EMH) was first fully proposed by Eugene Fama in his doctoral thesis at the University of Chicago in the 1960s. In short, it states that the markets are "informationally efficient" in the sense that all available information is incorporated in the current stock price. The implication is that since all information is embodied in the current price, it should be difficult for investors to beat the market year in and year out.

Over time it has been much debated and variations have emerged that allow exceptions for holders of private information (say, management) small stocks that are not heavily traded. The EMH has been much criticized, particularly by professional money managers who would be out of work if the market were perfectly efficient. After all, if the pros can't outperform the market, why not just buy index funds?

Many people take the EMH to mean that the markets are *always* right. Today even Fama admits the market makes mistakes: "In a period of high uncertainty, it's very difficult to figure out what the right prices are for stocks."[21] And Ken French, a frequent collaborator with Fama and Professor at the Tuck School of Business at Dartmouth, said in an interview jointly conducted with Fama that:

> *The efficient market hypothesis is just a model and, like all interesting models, it is not literally true. There are mistakes in prices even if one considers just publicly available information and, since people use financial prices to help decide how to allocate resources, those mistakes must affect the underlying reality. Of course, the existence of mistakes does not imply they are easy to find.*[22]

How the RPF Valuation Model Explains October 19, 1987 (Black Monday)

U.S. and global markets plunged on October 19, 1987, with the S&P 500 declining more than 20%. The cause of the decline has been much discussed, with program trading often cited as the main culprit along with portfolio insurance (derivatives).[23]

The application of the RPF Model to this period is revealing. As shown in Figure 6, which shows actual versus predicted S&P levels,[24] the market appears to have gotten "ahead of itself"—thereby creating a bubble of sorts—in anticipating an increase in earnings and values. As can be seen in Figure 7, interest rates began to climb in March 1987, rising from 7.25% in March to 9.25% in October, driving down the predicted P/E and the predicted level of the S&P 500.[25] Yet despite flat earnings, the market grew by 12% from February to September (and a total of 25%

21. "CBS Money Watch, http://moneywatch.bnet.com/investing/article/eugene-fama-why-you-cant-time-the-market/277142/.

22. "Fama/French Forum" http://www.dimensional.com/famafrench/2009/04/qa-bias-in-the-emh.html.

23. "Black Monday 10 Years Later: 1987 Timeline," *The Motley Fool Website*, accessed March 2009, http://www.fool.com/features/1997/sp971017crashanniversary-1987timeline.htm.

24. See Note 13.

25. See Note 14.

Figure 6 **Actual vs. Predicted During October 1987 Crash**[32]

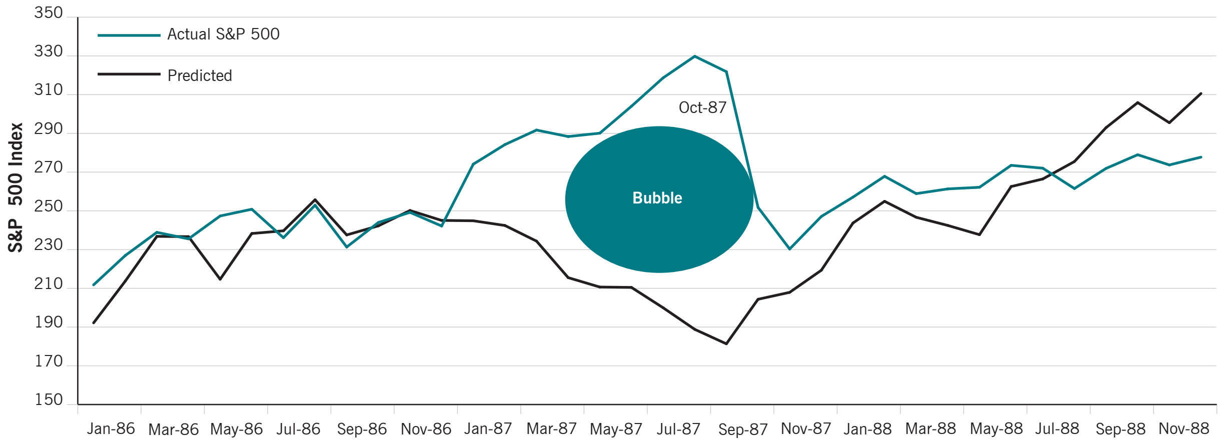
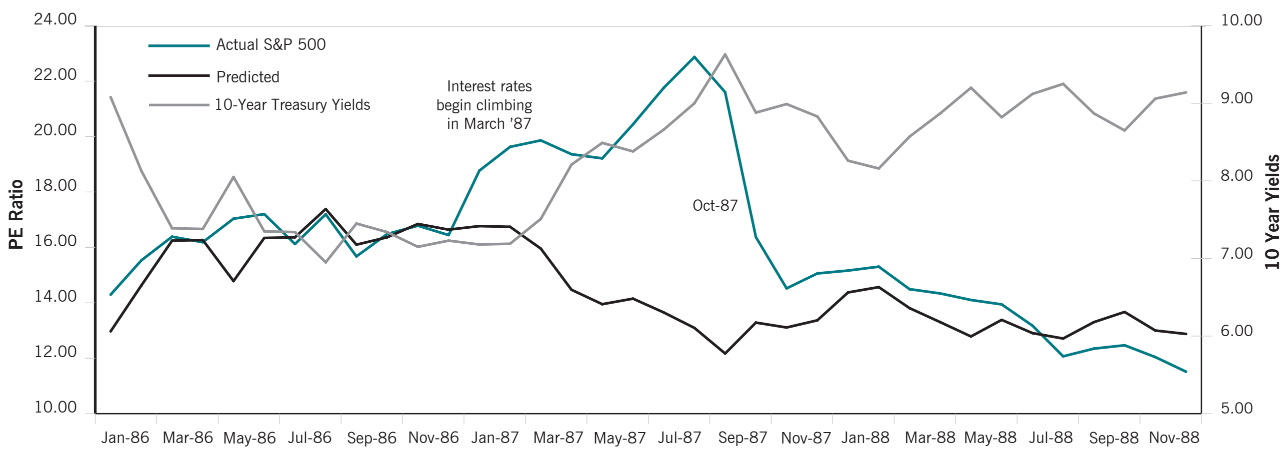


Figure 7 **Interest Rate Impact on October 1987 Crash, Actual S&P 500 Month-end data–10-Year Treasury Yields**



from December). With the market crash in October, the predicted and actual fell back into parity, with both figures suggesting the creation and bursting of a bubble.[26]

The suggestion offered by the RPF model in this case is that the underlying cause of the crash was excessive valuation relative to the sharp rise in interest rates. While actual and predicted levels often deviate, without a shift in the RPF, they tend to fall back in line.

But why did the market fall on October 19 and not November 19? The market began its decline in August. During the days before October 19, Iran had attacked a U.S flagged tanker, exacerbating fears that oil prices would continue to rise.[27] Perhaps this solidified the belief that earnings would not rise and inflation would stay high, keeping interest rates high. And this point of view was rapidly assimilated into the market. My own belief is that these developments were nothing more than the pinpricks that popped the balloon—actions that, while not particularly momentous in and of themselves, were enough to cause an unbalanced state to return to a more sustainable equilibrium. While derivatives and program trading may have aggravated the market decline once the decent began, they were not the fundamental cause, but rather part of the mechanism that helped to restore equilibrium.

26. See Note 14.

27. "Iranian Attacks on Kuwaiti Port Called Cause for U.S. to Retaliate," *The New York Times*, October 18, 1987, http://www.nytimes.com/1987/10/18/world/iranian-attacks-on-kuwaiti-port-called-cause-for-us-to-retaliate.html.

Figure 8 **Actual vs. Predicted during the 2000 dot.com Bubble, S&P 500 Month-end data–10-Year Treasury Yields**

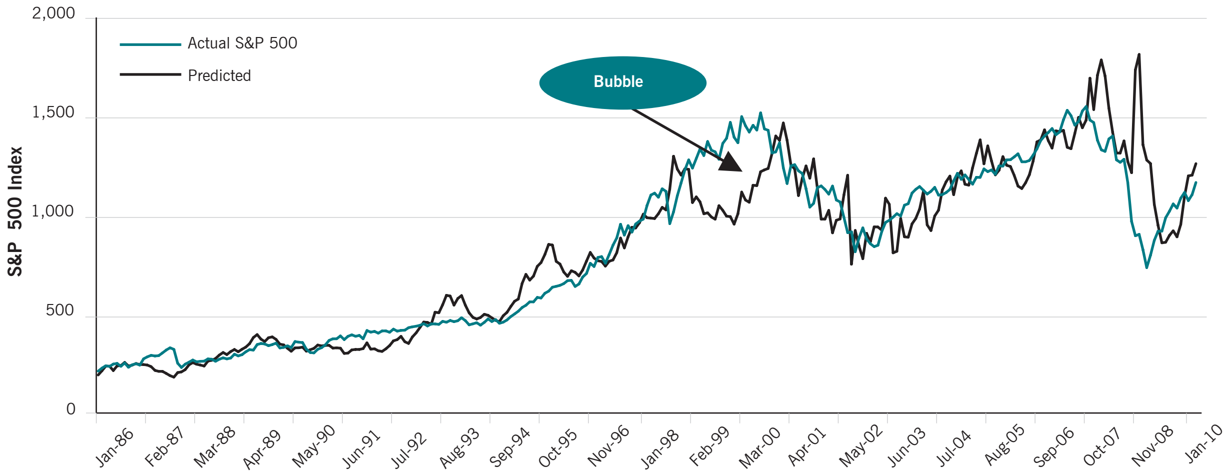
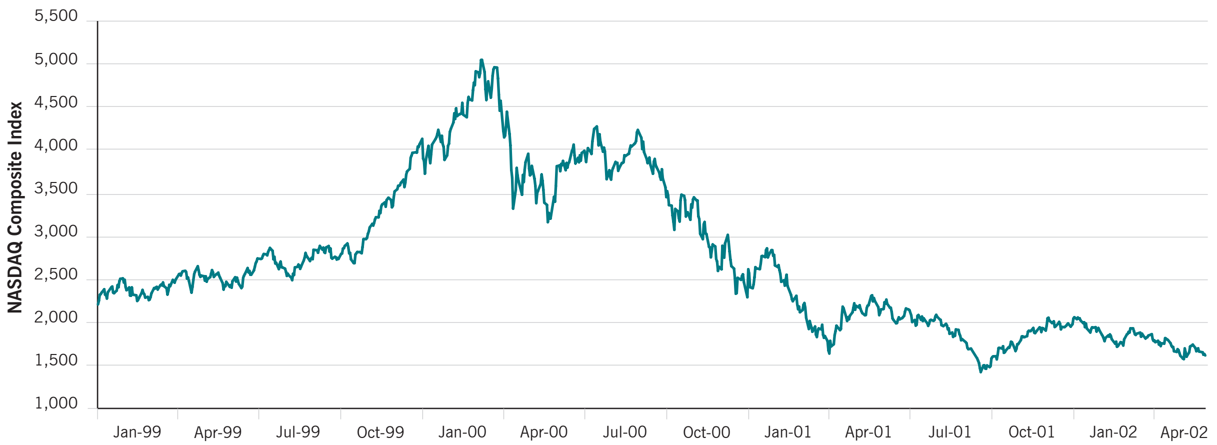


Figure 9 **NASDAQ January 1999–May 2002**



Source: Yahoo! Finance

2000 "Dot Com" Bubble: RPF Model Suggests Significant Bubble for the S&P 500

The NASDAQ peaked on March 10, 2000, at 5,132 in what is widely considered to be a bubble driven by excessive valuations of the Internet and other technology companies. Many economists such as Robert Schiller, author of *Irrational Exuberance*, argued that the entire market was embroiled in a speculative bubble throughout this period.[28]

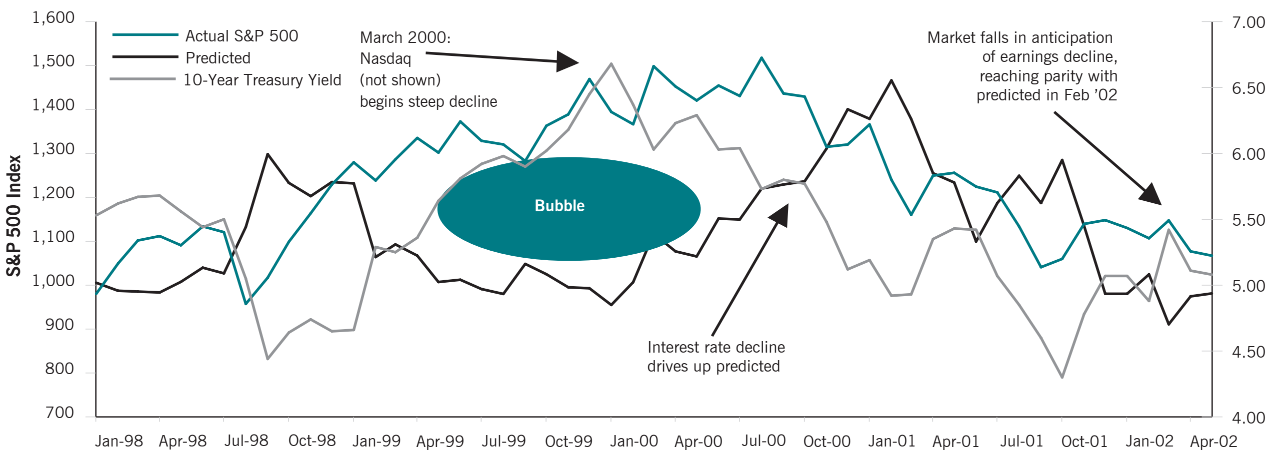
Application of the RPF Model to the S&P 500, strongly suggests that a significant bubble did exist. Indeed, Figure 8 suggests that the dot.com bubble of the late 90s was by far the largest during the period 1986 through 2009.

The model was not applied to the NASDAQ because it would be inappropriate to assume that the long-term growth of the smaller cap and technology heavy NASDAQ would equal long-term GDP growth and that volatility (Beta) would be the same as the S&P 500. As shown in Figure 9, the NASDAQ had declined by 32% in mid-April 2000 from its March 10 high, and by 51% by the end of 2000.

What explains this plunge in prices? From November 1998 until March 2000, 10-year Treasury yields increased from 4.6% to 6.2%. While the NASDAQ began to run up in late 1999, as can be seen in Figure 10, the S&P 500 Index began to diverge from RPF Model predictions in January

28. Robert J. Schiller, *Irrational Exuberance*, (Princeton University Press).

Figure 10 **Dot.com Bubble Close Up, Actual S&P 500 Month-end data–10-Year Treasury Yields** [32]



1999. As also shown in the figure, the S&P 500 Index did not begin its decline until August 2000. (Remember the model is applied using actual reported operating earnings, so predicted levels at any point are backward looking and do not reflect expectations.) However, the market began to anticipate that the NASDAQ meltdown would have a negative impact on earnings and the index followed.[29] And since S&P earnings fell by 27% from March 2000 to December 2001, the RFP Model appears to have "signaled" that earnings would fall well in advance of the actual reported drop.

The implication, then, is that the bubble was created by the combination of inflated earnings levels with rising 10-year Treasury yields that the market was somehow slow to recognize. To the extent the increases in interest rates were orchestrated by the Fed to cool an overheating economy, investors may have misread the signal and expected the increase in interest rates to be temporary. But, as the rate increases began to affect earnings, the market began a sharp repricing as the new point of view was assimilated.

How the RPF Valuation Model Explains 2008–2009 Meltdown and Recovery

The bursting housing bubble and mortgage crisis ultimately led to the meltdown that began September 2008. By August 2008, the S&P 500 had already fallen by 16% from its May 2007 peak. During this period, 10-year Treasury yields declined from around 5% to less than 4%. As illustrated in Figure 11, this led to an increase in predicted levels of the S&P 500 index.

According to the Case-Schiller Home Price Index, home prices fell more than 10% from second quarter of 2006 to the fourth quarter of 2007 and a total of 18% by the second quarter of 2008.[30] This historically large decline led to (well-founded) concerns about financial instability and the elimination of an important source of disposable income. Once again, in anticipation of a decline in earnings, the S&P 500 index fell while the RPF Model (using reported operating earnings) showed an increase in predicted levels as interest rates declined. The lines for expected and actual S&P values in Figure 11 begin to converge in August 2008, just before the worst of meltdown began in September and October. Investors were unable to absorb the seriousness of the pending crisis, so while the market fell in anticipation of an earnings decline, the expectations did not come close to reflecting the magnitude of the situation.

As can be seen in Figure 11, the flight to quality and resulting drop in Treasury rates clearly drove up the predicted levels to abnormal highs. But, as interest rates returned to a more normal level by June 2009, the predicted and actual levels returned to parity.

RPF Model implications for efficient markets?

• Over a longer period of time, the market is efficient if one allows for oscillations around true value, but is also subject to making mistakes. These mistakes can create bubbles.

• Over time the bubbles are deflated and the market returns to predicted levels as new long-term views are assimilated.

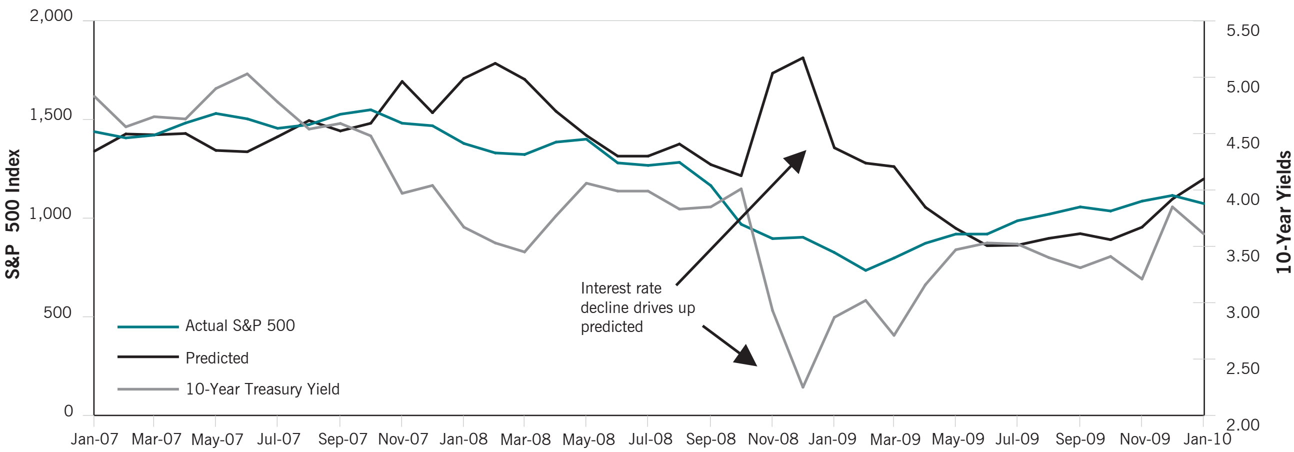
• The RPF Valuation model has shown to be useful in identifying bubbles before they pop.

This pattern supports the contention that the valuation model would have worked well during this period with a

29. See Note 13.

30. "S&P/Case-Schiller Home Price Indices," *Standard and Poors Website*, accessed March to April 2009, http://www2.standardandpoors.com/spf/pdf/index/csnational_values_022445.xls.

Figure 11 **Actual vs. Predicted During 2008–2009 Meltdown, S&P 500 Month-end data–10-Year Treasury Yields**



normalized interest rate. It also shows how the market led predicted levels as it incorporated expected rather than actual historical operating earnings.

In sum, analysis of these major market events with the RPF Model supports the contention that markets make mistakes in processing information. It also suggests that market prices oscillate around a true fair value price. But, as highlighted throughout this discussion of three major market events, these deviations can be very large.

2010 Outlook

As of this writing, on April 14, 2010, the S&P 500 Index closed at 1,211, as compared to a predicted level of 1,260—still 4% below the predicted level. In addition to looking at the market today, the model can help inform an opinion about the future. S&P estimates 2010 operating earnings of $75.27. If we also assume the 10-year Treasury remains unchanged at 3.83%, the S&P 500 Index would be predicted to end the year at 1,485—a gain of another 23%. But if the bond rate rises to 5%, even with the growth in earnings, the S&P's predicted value at year end is 1,107—a drop of 9% from the current level.

Conclusions

Many people view the market valuation process as a black-box driven by emotion, leaving many managers unsure what strategies they can pursue to increase shareholder value. Using two main variables, the RPF Valuation model highlights a number of important principles that can be used to inform the valuation of all companies in most (though not all) circumstances:

1. The Equity Risk Premium is not a constant, but a relatively stable Risk Premium Factor (RPF) that is applied to the risk-free rate (10-year Treasury yields).

2. The Risk Premium Factor is consistent with the loss aversion coefficient associated with the prospect theory (of Kahneman and Tversky).

3. The Risk Premium Factor Valuation Model [$P = E / (R_f \times (1+RPF) – (R_f – Int_R + G_R))$] effectively explains both P/E and S&P 500 Index levels using readily available information and simplifying assumptions.

4. Growth is a critical component of valuation, and the impact of growth on value is easily quantified using the RPF model.

5. Interest rates drive market value—and the fair value of the market (P/E Ratio) cannot be estimated without considering interest rates.

6. Interest rates have a greater impact on market price and valuation than is generally recognized, with low rates more beneficial and high rates more punishing.

7. Declining interest rates were a major factor in the long bull market from 1980 through 2007.

8. The RPF model suggests that if Treasury yields remain in the low 4%–5% range and earnings recover to 2006/07 levels, the market could stage a rally and recover to record levels, with the S&P 500 Index rising to the range of 1,300–1,700.

9. Though efficient and rational over longer time periods, the market is prone to occasional, generally short-lived oscillations and pricing errors.

STEVE HASSETT is president of Hassett Advisors based in Atlanta, Georgia, which specializes in corporate development and growth strategies. Previously, he was VP-international and emerging businesses at the Weather Channel, founder of a Web and mobile software company, and a corporate finance consultant with Stern Stewart & Co.

Journal of Applied Corporate Finance (ISSN 1078-1196 [print], ISSN 1745-6622 [online]) is published quarterly, on behalf of Morgan Stanley by Wiley Subscription Services, Inc., a Wiley Company, 111 River St., Hoboken, NJ 07030-5774. Postmaster: Send all address changes to JOURNAL OF APPLIED CORPORATE FINANCE Journal Customer Services, John Wiley & Sons Inc., 350 Main St., Malden, MA 02148-5020.

Information for Subscribers *Journal of Applied Corporate Finance* is published in four issues per year. Institutional subscription prices for 2010 are:

Print & Online: US$416 (US), US$499 (Rest of World), €323 (Europe), £255 (UK). Commercial subscription prices for 2010 are: Print & Online: US$556 (US), US$663 (Rest of World), €429 (Europe), £338 (UK). Individual subscription prices for 2010 are: Print & Online: US$105 (US), £59 (Rest of World), €88 (Europe), £59 (UK). Student subscription prices for 2010 are: Print & Online: US$37 (US), £21 (Rest of World), €32 (Europe), £21 (UK).

Prices are exclusive of tax. Asia-Pacific GST, Canadian GST and European VAT will be applied at the appropriate rates. For more information on current tax rates, please go to www3.interscience.wiley.com/about us/journal_ordering_and_payment.html#Tax. The price includes online access to the current and all online back files to January 1997, where available. For other pricing options, including access information and terms and conditions, please visit www.interscience.wiley.com/journal-info.

Journal Customer Services:For ordering information, claims and any enquiry concerning your journal subscription please go to interscience.wiley.com/support or contact your nearest office.
Americas: Email: cs-journals@wiley.com; Tel: +1 781 388 8598 or +1 800 835 6770 (toll free in the USA & Canada).
Europe, Middle East and Africa: Email: cs-journals@wiley.com; Tel: +44 (0) 1865 778315.
Asia Pacific: Email: cs-journals@wiley.com; Tel: +65 6511 8000.
Japan: For Japanese speaking support, Email: cs-japan@wiley.com; Tel: +65 6511 8010 or Tel (toll-free): 005 316 50 480. Further Japanese customer support is also available at www.interscience.wiley.com/support
Visit our Online Customer Self-Help available in 6 languages at www.interscience.wiley.com/support

Production Editor: Joshua Gannon (email:jacf@wiley.com).
Delivery Terms and Legal Title Prices include delivery of print journals to the recipient's address. Delivery terms are Delivered Duty Unpaid (DDU); the recipient is responsible for paying any import duty or taxes. Legal title passes to the customer on despatch by our distributors.

Back Issues Single issues from current and recent volumes are available at the current single issue price from cs-journals@wiley.com. Earlier issues may be obtained from Periodicals Service Company, 11 Main Street, Germantown, NY 12526, USA. Tel: +1 518 537 4700, Fax: +1 518 537 5899, Email: psc@periodicals.com

This journal is available online at Wiley InterScience. Visit www.interscience.wiley.com to search the articles and register for table of contents e-mail alerts.

Access to this journal is available free online within institutions in the developing world through the AGORA initiative with the FAO, the HINARI initiative with the WHO and the OARE initiative with UNEP. For information, visit www.aginternetwork.org, www.healthinternetwork.org, www.healthinternetwork.org, www.oarescience.org, www.oarescience.org

Wiley's Corporate Citizenship initiative seeks to address the environmental, social, economic, and ethical challenges faced in our business and which are important to our diverse stakeholder groups. We have made a long-term commitment to standardize and improve our efforts around the world to reduce our carbon footprint. Follow our progress at www.wiley.com/go/citizenship

Abstracting and Indexing Services
The Journal is indexed by Accounting and Tax Index, Emerald Management Reviews (Online Edition), Environmental Science and Pollution Management, Risk Abstracts (Online Edition), and Banking Information Index.

This journal is printed on acid-free paper.

Financial Analysts Journal
Volume 66 • Number 4
©2010 CFA Institute

Conflicts of Interest and Analyst Behavior: Evidence from Recent Changes in Regulation

Armen Hovakimian and Ekkachai Saenyasiri

Regulation FD made analysts less dependent on insider information and diminished analysts' motives to inflate their forecasts. The Global Research Analyst Settlement had an even bigger impact on analyst behavior: The mean forecast bias declined significantly, whereas the median forecast bias essentially disappeared. These results are similar for all analysts.

Our investigation of the impact of recent changes in regulation on analysts' forecasting behavior follows a number of studies that argued that analysts were motivated to produce research reports that did not reflect their true opinions. Analysts tended to make excessive "buy" recommendations and inflated earnings forecasts for several reasons, two of which gained considerable attention from regulators in the United States. First, analysts may have felt compelled to favor managers in covered companies in order to gain privileged access to information flow (Lim 2001). Second, although analysts are supposed to provide investors with accurate and truthful research reports, conflicts of interest could occur because analysts' compensation was tied to profits generated from investment banking business and brokerage commissions (Lin and McNichols 1998; Carleton, Chen, and Steiner 1998).

In the early part of the first decade of this century, in an effort to restore public confidence in U.S. capital markets, U.S. regulators enacted several rules and regulations, prosecuted analysts whose research reports were tainted by conflicts of interest, and fined banks that failed to prevent research analysts' conflicts of interest. Two of the main regulatory developments during this period were (1) Regulation Fair Disclosure (Reg FD), which became effective on 23 October 2000, and (2) the Global Research Analyst Settlement (Global Settlement), which was announced on 20 December 2002.[1]

Although the primary goals of these two regulatory actions are different, they both have the potential to improve the quality of analyst forecasts. One of the stated goals of Reg FD is to prohibit private communication between companies and analysts, thereby helping to level the playing field so that market participants can have equal access to information and making analysts less dependent on such communication. In prohibiting companies from selectively disclosing private information to analysts, Reg FD may reduce analyst forecast bias by eliminating the incentive for analysts to inflate their earnings forecasts in order to gain access to insider information.

The Global Settlement is an important enforcement agreement between U.S. regulators and 12 large investment banks (the Big-12 banks) designed to eliminate research analysts' conflicts of interest. If successful, the Global Settlement should reduce optimistic bias in analyst forecasts.

Our study considered whether these two actions by U.S. regulators reduced the bias in analysts' earnings forecasts documented in previous studies. We focused on annual earnings forecast bias for several reasons. First, investors may use analyst forecasts to form expectations of earnings and cash flows, both of which are important inputs for stock valuation models. Inflated earnings forecasts can drive stock prices above their fair values if investors fail to adjust for the bias.[2]

Second, given the flurry of new regulations, regulators clearly consider analyst behavior an important factor in maintaining investor confidence in financial markets. Regulation is costly because of the significant expenses associated with analyzing problematic situations and developing remedies. Moreover, restrictions and reporting requirements imposed on various market participants result in ongoing compliance costs. These costs can be justified only if the new regulations help reduce analysts' conflicts of interest and thereby generate an important benefit for financial markets.

Armen Hovakimian is professor of finance at Baruch College, New York City. Ekkachai Saenyasiri is assistant professor of finance at Providence College, Providence, Rhode Island.

Third, most studies that have examined the impact of Reg FD and the Global Settlement on analyst behavior focused on forecast accuracy and forecast dispersion (Bailey, Li, Mao, and Zhong 2003; Agrawal, Chadha, and Chen 2006).[3] These aspects of analyst behavior, however, are little affected by conflicts of interest, the focus of our study.

Other studies have examined forecast bias. Clarke, Khorana, Patel, and Rau (2006) found that the Global Settlement had no impact on relative bias in analyst forecasts. Focusing on the impact of Reg FD on bias in quarterly earnings forecasts between October 1999 and December 2001, Mohanram and Sunder (2006) found that these forecasts became more optimistic after Reg FD but attributed the increase to unexpectedly low realized earnings during the 2001 recession. Our longer study period (1996–2006) allowed us to control for macroeconomic conditions in our regression analysis. Furthermore, we examined longer-term (up to 24 months) earnings forecasts in which the forecast bias is more apparent (Richardson, Teoh, and Wysocki 2004). Although Herrmann, Hope, and Thomas (2008) found some evidence of decline in forecast bias following Reg FD, they focused on internationally diversified companies only; we examined all U.S. companies, and our primary focus was on changes in forecast bias after the Global Settlement.

Lastly, the ability of analysts to forecast earnings accurately can be easily and straightforwardly verified because actual earnings are observed at the end of the forecast period. Barber, Lehavy, McNichols, and Trueman (2006) studied the change in distribution of stock recommendations made from 1996 to 2003. They found that the percentage of buys decreased starting in mid-2000.[4] How unbiased the new distribution of stock recommendations is, however, remains uncertain. But we know that the bias should be zero at the aggregate level when analysts make their forecasts on the basis of their true opinions.

Institutional Background

Historically—and especially before recent regulations—analysts have tended to make unduly optimistic earnings forecasts. In this section, we discuss the possible reasons for this optimistic bias and the potential impacts of the recent regulations on such bias.

Why Do Analysts Make Overoptimistic Earnings Forecasts? A number of studies have documented that analysts regularly make overoptimistic earnings forecasts (Brown 1997; Chopra 1998; Beckers, Steliaros, and Thomson 2004). Optimistic bias tends to be larger for longer-term forecasts and smaller for forecasts made closer to the earnings announcement date. This phenomenon is usually referred to as the walk-down trend (Richardson, Teoh, and Wysocki 2004). Several explanations have been offered for analyst optimism.

First, analysts may be influenced by conflicts of interest if their compensation is tied to investment banking fees and brokerage commissions. Lin and McNichols (1998) found that analysts affiliated with underwriters make more favorable stock recommendations and long-term earnings growth forecasts than analysts not so affiliated. Agrawal and Chen (2005) discovered that optimism in long-term earnings growth forecasts is high when analysts work for financial institutions whose revenues come mainly from brokerage business. Carleton, Chen, and Steiner (1998) found that stock recommendations made by brokerage firms are more optimistic than those of nonbrokerage firms. Using Australian data, Jackson (2005) noted that optimistic analysts generate more trades for their brokerage firms than do less optimistic analysts. Chan, Karceski, and Lakonishok (2007) showed that analysts' earnings forecasts are influenced by their desire to win investment banking clients. Doukas, Kim, and Pantzalis (2005) reported that stocks with excess analyst coverage yield lower future returns, consistent with the conflict-of-interest hypothesis. Hong and Kubik (2003) found that brokerage houses reward optimistic analysts; optimistic analysts at low-status brokerage houses are more likely to move up to higher-status brokerage houses than are less optimistic analysts.

Second, analysts may feel compelled to maintain good relations with company management in order to gain access to insider information that can help improve the accuracy of their forecasts (Lim 2001). Third, analysts may tend to cover stocks for which they have positive views and drop or avoid stocks for which they have negative views, which can induce a self-selection bias (McNichols and O'Brien 1997). Fourth, analysts may have a cognitive bias that leads them to overreact to good earnings information and underreact to bad earnings information (Easterwood and Nutt 1999; Nutt, Easterwood, and Easterwood 1999). Finally, the walk-down trend may be driven by the "earnings guidance game," in which analysts issue optimistic forecasts at the start of the fiscal year and then revise their estimates until the company can beat the forecast at the earnings announcement date (Richardson, Teoh, and Wysocki 2004).

Recent Regulations. Before Reg FD, analysts and institutional investors often had an informational advantage over small investors through private communications with management and conference calls in which company managers discussed past performance and provided guidance on future prospects. Such timely information gave these investment professionals an unfair advantage that allowed them to trade stocks profitably at the expense of uninformed investors.

To gain access to this information flow, analysts may have had to maintain good relations with insiders by making optimistic forecasts and buy recommendations in their research reports. Analysts' excessively optimistic views of the stocks were misleading and contributed to the deterioration of investor confidence in capital market integrity. Through Reg FD, which was introduced in October 2000, the U.S. SEC intended to improve fairness and restore public confidence in the markets by requiring U.S. public companies to disclose material information simultaneously to all market participants.

Other sources of conflicts of interest, however, remained unaddressed by Reg FD. For instance, analysts could be pressured to make optimistic forecasts and buy recommendations in order to favor investment banking clients and generate trading volume. The SEC and such self-regulatory organizations (SROs) as the National Association of Securities Dealers (NASD; now the Financial Industry Regulatory Authority [FINRA]) and the NYSE paid significant attention to this issue and introduced a number of new rules and regulations to curb the negative consequences of these conflicts of interest.

The Sarbanes–Oxley Act of 2002 (SOA), also known as the Public Company Accounting Reform and Investor Protection Act of 2002, became law on 30 July 2002. The SOA is a broad piece of legislation that covers various business practices, including auditor independence, corporate responsibility, enhanced financial disclosure, analysts' conflicts of interest, and corporate and criminal fraud accountability. The SOA amended the Securities Exchange Act of 1934 by creating Section 15D, which requires FINRA and the NYSE to adopt rules reasonably designed to address research analysts' conflicts of interest.

To comply with the SOA, the NASD released Rule 2711 (Research Analysts and Research Reports) and the NYSE amended its Rule 351 (Reporting Requirements) and Rule 472 (Communications with the Public). Most provisions of these rules went into effect on 9 July 2002. These rules mitigate analysts' conflicts of interest by separating research analysts from the influence of the investment banking and brokerage businesses. Research analysts' compensation can no longer be tied to the performance of these businesses. In addition, analysts are restricted from personal trading in the stocks they cover.

On 6 February 2003, the SEC adopted Regulation Analyst Certification (Reg AC).[5] Reg AC provides guidelines for proper disclosure of potential conflicts of interest of sell-side analysts, including their association with investment banking clients and the structure of their compensation.

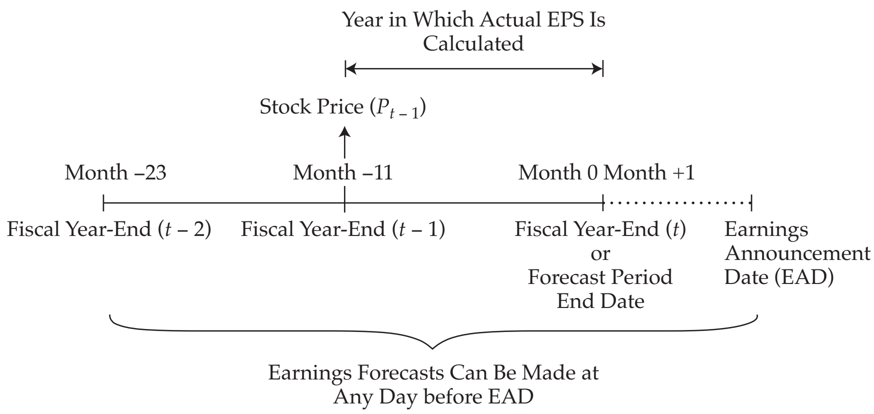
Regulatory objectives have also received support from rigorous enforcement actions. Following a joint investigation by the SEC, NASD, NYSE, and New York State Attorney General, 10 large U.S. and multinational investment banks agreed to pay a fine of $1.435 billion in the Global Research Analyst Settlement for their failure to adequately address research analysts' conflicts of interest. Announced on 20 December 2002, the terms of the Global Settlement initially covered 10 banks.[6] The final agreement was announced on 28 April 2003. Two more banks reached settlements on 26 August 2004.[7] The Global Settlement and the SRO rules share the same spirit in that their mutual objective is to eliminate analysts' conflicts of interest.

The introduction of these rules and regulations allows us to differentiate among the alternative explanations for analyst forecast bias proposed in the literature. First, a reduction in forecast bias after Reg FD would support the argument that analysts were overoptimistic owing to their need for insider information, especially if such a reduction were stronger for informationally more opaque companies. Second, a reduction in bias after the Global Settlement and Rule 2711 would be consistent with the hypothesis that analyst behavior was unduly influenced by conflicts of interest.[8] In contrast, self-selection and cognitive biases may exist even in a world without conflicts of interest. Therefore, if these biases are the main reasons for analysts' overoptimistic forecasts, then these regulatory changes should have no effect on forecast bias.[9]

Sample and Variables

We downloaded sell-side analysts' earnings forecasts for fiscal year-end dates between 1996 and 2006 from the Detail file of the I/B/E/S database. We used forecasts for current- and subsequent-year earnings per share (EPS), which are made for the upcoming and following years' earnings announcement dates.[10] **Figure 1** illustrates the timeline of analyst forecasts. The earliest analyst forecasts for a specific fiscal year-end EPS are made 24 months before the forecast fiscal year-end (in forecast month –23). For each EPS, analysts can

Figure 1. Timeline of Analyst Forecasts



make multiple forecasts over the course of the next 24 months. Some analysts may continue to make forecasts after the forecast fiscal year ends because companies announce their annual earnings after a delay of several months. Because the length of the EPS announcement delay could be affected by how high or low the realized EPS is relative to the consensus, we retained only those forecasts made no more than one month after the forecast fiscal year-end (in forecast month +1), which left us with a total of 2,297,792 forecasts.

For each forecast, I/B/E/S provides actual earnings, forecast date, forecast period (fiscal year) end, earnings announcement date, analyst code identity, broker code identity, and number of analysts used for consensus calculation.[11] We used the I/B/E/S Broker Translation file to convert broker codes into brokers' names, which we used to identify analysts who worked for the Big-12 banks. Stock prices are from the I/B/E/S Summary file.[12] We downloaded real GDP growth rates from the website of the U.S. Bureau of Economic Analysis. We downloaded SIC codes from the CRSP monthly file.

We defined analyst forecast bias, the focus of our analysis, as the average analyst forecast error and calculated it as follows:

$$Bias_{j,t,m} = 100\left(F_{j,t,m} - A_{j,t}\right)/P_{j,t-1}, \tag{1}$$

$$F_{j,t,m} = \frac{1}{I_{j,t,m}} \sum_{i=1}^{I} F_{j,t,m,i}, \tag{2}$$

and

$$F_{j,t,m,i} = \frac{1}{K_{j,t,m,i}} \sum_{k=1}^{K} F_{j,t,m,i,k}, \tag{3}$$

where

$A_{j,t}$ = the actual earnings per share for company j in fiscal year t

$F_{j,t,m,i}$ = the average of annual earnings forecasts for fiscal year-end t of company j, made in month m by analyst i

$K_{j,t,m,i}$ = the number of forecasts made in month m by the same analyst i for the same company j and fiscal year t

$I_{j,t,m}$ = the number of analysts making forecasts in month m for company j and fiscal year t

$P_{j,t-1}$ = the stock price of company j one year before the fiscal year-end t[13]

Note that all EPS forecasts made for the same company and the same fiscal year are normalized by the same stock price. Using the same stock price as the denominator guarantees that any changes in forecast bias across forecast months (m) are the result of changes in analyst forecasts, not of changes in the stock price. In our calculations according to Equations 1–3, we used only new forecasts made in month m. Stale forecasts from earlier months (m – 1, etc.) were not carried over into month m. In other words, each forecast participated in the calculation of the forecast bias only once, in the month in which the forecast was made. In our sample, an average analyst made 4.5 forecasts for each annual EPS. Because for each annual EPS we tracked 25-month forecasts (from month –23 to month +1), the implication is that an average analyst in our sample made a forecast for each covered company about once every six months.

To minimize the influence of outliers and misreported data in our analysis, we replaced with missing values any extreme observations of forecast bias, company size, market-to-book ratio, the number of stocks, and the number of industry analysts following.[14] We dropped from the sample all forecasts made in October 2000 and December 2002 (1.5 percent of our sample) and observations with missing values of any relevant variable. We were left with 1,586,000 individual analyst forecasts, which we used to calculate 434,268 average forecast errors. For each fiscal year and for each of our 7,315 sample companies, our sample contained up to 25 monthly observations of forecast bias ($Bias_{j,t,m}$).

Table 1 reports the summary statistics for the overall sample of 434,268 observations and for each of the three subperiods. The period before Reg FD represents 53 percent of our sample observations, with the period between Reg FD and the Global Settlement and the period after the Global Settlement representing 18 percent and 29 percent of the sample observations, respectively. The mean forecast bias across all sample observations is 1.39 percent of stock price. This result is consistent with prior evidence that analysts' forecasts are optimistically biased (Brown 1997; Chopra 1998). No significant difference exists between the mean forecast bias before Reg FD (1.72) and the mean forecast bias between Reg FD and the Global Settlement (1.97). The mean forecast bias is more than four times smaller after the Global Settlement (0.41), with the difference statistically significant at the 1 percent level.

The average market capitalization of companies in our sample was $4.5 billion, and the average market-to-book ratio was 3.57. On average, 8.41 analysts covered a company in any particular month. The analysts in our sample worked for brokers that, on average, each employed 65.7 analysts. A typical analyst followed 16.30 stocks from 4.78 industries and, at the time of the forecast, had been in the I/B/E/S database for 6.24 years and making forecasts for the covered stock for 2.5 years. Around 17 percent of forecasts were made for companies with negative earnings, and 36 percent of forecasts were made for companies whose earnings were declining relative to earnings in the prior fiscal year.

Test Results

In this section, we present the results of the univariate tests and of the regression analysis of the effects of Reg FD and the Global Settlement on bias in analyst forecasts.

Univariate Results by Forecast Month. **Table 2** presents the median forecasts by the month in which the forecasts were made and by the fiscal year for which they were made. The numbers in the leftmost column represent the month (relative to the fiscal year-end) of the forecast. The numbers in the top row represent the fiscal years for which the

Table 1. Summary Statistics

				Number of Observations			Mean		
Description	Variable	Number of Observations	Mean	Before Reg FD	Between Reg FD and GS	After GS	Before Reg FD	Between Reg FD and GS	After GS
Forecast bias	Bias	434,268	1.39	231,096	77,305	125,867	1.72	1.97	0.41
Reg FD indicator	RegFD	434,268	0.18	231,096	77,305	125,867	0.00	1.00	0.00
Global Settlement indicator	Glob	434,268	0.29	231,096	77,305	125,867	0.00	0.00	1.00
Company characteristics									
Analyst coverage	NumA	434,268	8.41	231,096	77,305	125,867	8.21	8.23	8.88
Market cap ($ millions)	CompanySize	434,268	4,470.00	231,096	77,305	125,867	3,480.00	5,250.00	5,800.00
Market-to-book ratio	MB	434,268	3.57	231,096	77,305	125,867	3.78	3.47	3.23
Negative EPS	EPSLoss	434,268	0.17	231,096	77,305	125,867	0.16	0.26	0.14
Declining EPS	EPSDecline	434,268	0.36	231,096	77,305	125,867	0.37	0.45	0.27
Litigation	Litigation	434,268	0.27	231,096	77,305	125,867	0.25	0.30	0.27
Labor intensive	Labor	434,268	0.61	231,096	77,305	125,867	0.60	0.63	0.63
Analyst characteristics									
Company-specific experience	YearStk	434,268	2.50	231,096	77,305	125,867	2.55	2.43	2.46
General experience	YearIBES	434,268	6.24	231,096	77,305	125,867	6.45	6.19	5.87
No. of stocks covered	NumStk	434,268	16.30	231,096	77,305	125,867	18.18	14.31	14.06
No. of industries covered	NumInd	434,268	4.78	231,096	77,305	125,867	5.46	4.15	3.93
Broker size	BrokerSize	434,268	65.70	231,096	77,305	125,867	54.98	89.03	71.06

Note: This table presents the summary statistics for the overall sample and for the three subperiods.

Table 2. Forecast Bias by Fiscal Year and Forecast Month

	Forecast Period End Year										
Month	96	97	98	99	00	01	02	03	04	05	06
–23	0.1	0.4	1.4	1.6	–0.3	1.9	2.3	1.2	–0.2	–0.3	–0.3
–22	0.3	0.5	0.9	1.3	0.5	2.2	2.7	1.3	0.0	–0.1	0.0
–21	0.3	0.5	1.1	1.6	0.5	2.1	2.6	1.3	0.0	0.0	0.2
–20	0.4	0.5	1.1	1.3	0.6	2.2	2.2	1.4	–0.1	0.0	0.0
–19	0.5	0.7	1.1	1.6	0.5	2.1	2.1	1.3	–0.1	0.0	0.1
–18	0.5	0.4	1.2	1.4	0.6	2.1	1.8	1.1	–0.2	0.0	0.1
–17	0.4	0.4	1.2	1.1	0.5	2.1	1.4	1.0	–0.2	0.0	0.1
–16	0.4	0.5	1.3	1.3	0.6	2.0	1.5	1.1	–0.1	0.0	0.2
–15	0.4	0.4	1.1	0.8	0.4	1.7	0.9	0.8	–0.3	0.0	0.2
–14	0.4	0.3	0.9	0.6	0.4	**FD**	0.6	0.4	–0.2	0.0	0.1
–13	0.4	0.3	1.0	0.6	0.4	1.5	0.5	0.3	–0.2	0.1	0.2
–12	0.3	0.2	0.8	0.4	0.3	1.6	0.4	**GS**	–0.2	–0.1	0.1
–11	0.3	0.3	0.8	0.3	0.1	1.3	0.3	0.1	–0.1	0.0	0.1
–10	0.2	0.2	0.5	0.1	0.2	1.1	0.2	0.0	–0.1	–0.1	–0.1
–9	0.2	0.2	0.6	0.1	0.1	1.1	0.2	0.0	–0.1	0.0	–0.1
–8	0.1	0.1	0.5	0.1	0.1	0.7	0.2	–0.1	–0.1	–0.1	–0.1
–7	0.1	0.0	0.5	0.1	0.1	0.6	0.2	–0.1	0.0	0.0	0.0
–6	0.1	0.1	0.4	0.0	0.1	0.5	0.2	–0.1	–0.1	–0.1	0.0
–5	0.0	0.0	0.2	0.0	0.0	0.2	0.1	–0.1	0.0	–0.1	0.0
–4	0.0	0.0	0.1	0.0	0.0	0.2	0.1	–0.1	0.0	0.0	0.0
–3	0.0	0.0	0.1	0.0	0.0	0.1	0.0	–0.1	–0.1	0.0	0.0
–2	0.0	0.0	0.0	0.0	**FD**	0.0	0.0	–0.1	–0.1	0.0	0.0
–1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	–0.1	0.0	0.0	–0.1
0	0.0	0.0	0.0	–0.1	0.0	0.0	**GS**	–0.1	0.0	0.0	–0.1
1	0.0	–0.1	0.0	–0.1	–0.1	0.0	0.0	–0.1	–0.1	–0.1	–0.3
Median bias	0.2	0.2	0.8	0.3	0.1	1.2	0.4	0.0	–0.1	0.0	0.0
Mean bias	1.2	1.1	1.8	2.2	1.4	3.0	2.1	1.6	0.1	0.5	0.3
Mean forecast	6.2	5.3	4.6	5.1	5.3	3.7	3.0	4.0	4.4	4.2	5.0
Mean actual earnings	5.0	4.1	2.8	2.9	3.9	0.7	0.9	2.4	4.2	3.7	4.7
Mean stock return (%)	0.2	0.2	0.0	0.2	0.0	0.0	–0.2	0.6	0.2	0.1	0.2
GDP (%)	3.7	4.5	4.2	4.5	3.7	0.8	1.6	2.5	3.9	3.2	3.3

Notes: *Forecast bias* is the difference between the mean of all forecasts made in a particular month for a particular company and a particular fiscal year and the realized EPS, scaled by the stock price and multiplied by 100. *Forecast period end year* is the fiscal year for which the forecast was made. *Month* is the month of the forecast relative to the fiscal year-end. *FD* is the month in which Reg FD became effective (October 2000). *GS* is the month in which the Global Settlement was announced (December 2002). Stock returns were calculated from our samples.

forecasts were made. For example, forecasts made in September 2000 for the fiscal year ended December 2000 (i.e., three months before the fiscal year-end) are in row –3 and column 00. The two solid lines separate the forecasts made before and after Reg FD and the forecasts made before and after the Global Settlement. The six bottom rows present forecast bias for each fiscal year averaged across all forecast months, along with the realized earnings per share, average forecasts, annual stock returns, and real GDP growth rates.[15] To align fiscal year-end dates with annual variables, such as real GDP growth rates, we used only forecasts for companies with December fiscal year-ends.

For each year before the Global Settlement, the median forecast errors are significantly positive. Furthermore, for each year before the Global Settlement, we observe the walk-down trend with forecast bias steadily declining as forecasts are made closer to the fiscal year-end. After the Global Settlement, we observe a significant drop in the forecast bias. The results show a total absence of bias in the median forecast errors for 2004–2006 (–0.1 percent, 0.0 percent, and 0.0 percent, respectively). The walk-down trend in median forecast errors is also practically nonexistent for fiscal years 2004–2006.

These results suggest that analysts' conflicts of interest indeed led to excess optimism in earnings forecasts before the Global Settlement and that the Global Settlement has been effective in neutralizing analysts' conflicts of interest. Alternative interpretations of the forecast bias, such as self-selection, cognitive bias, and need for insider information, cannot explain these findings because the Global Settlement should have no effect on these factors.

Unusually high stock valuations and/or realized earnings, rather than less optimistic forecasts, could be responsible for the decline in the average forecast errors after the Global Settlement. A quick look at the actual and forecasted EPS, stock returns, and real GDP growth rates before and after the Global Settlement, however, does not seem to support this idea. Neither aggregate economic performance nor stock valuations seem to be out of the ordinary in the post-settlement years. The actual earnings, stock returns, and GDP growth rates seem to be unusually low in the period between Reg FD and the Global Settlement. We controlled for the effects of these and other potentially relevant factors by examining the effects of Reg FD and the Global Settlement in a regression framework.

Regression Analysis. To examine how Reg FD and the Global Settlement affect bias in analyst forecasts while controlling for the confounding effects of company and analyst characteristics, as well as economic conditions, we estimated the following regression model:

$$\begin{aligned} Bias_{j,t,m} = {} & \alpha_0 + \alpha_1 RegFD_{t,m} + \alpha_2 Glob_{t,m} + \alpha_3 NumA_{j,t,m} \\ & + \alpha_4 CompanySize_{j,t,m-1} + \alpha_5 MB_{j,t,m-1} \\ & + \alpha_6 YearStk_{j,t,m} + \alpha_7 YearIBES_{j,t,m} \\ & + \alpha_8 NumStk_{j,t,m} + \alpha_9 NumInd_{j,t,m} \\ & + \alpha_{10} BrokerSize_{j,t,m} + \alpha_{11} EPSLoss_{j,t} \\ & + \alpha_{12} EPSDecline_{j,t} + \alpha_{13} Litigation_j \\ & + \alpha_{14} Labor_{j,t,m-1} + \alpha_{15} ActualGDP_t \\ & + \alpha_{16} UnexpectedGDP_{t,m} + \beta Month_m + \gamma Year_t \\ & + \delta_j \sum DCompany_j + \varepsilon_{j,t,m}. \end{aligned} \tag{4}$$

In Equation 4, $Bias_{j,t,m}$ is the mean forecast error for all forecasts for company j made in month m relative to the end of fiscal year t, calculated according to Equations 1–3. $RegFD_{t,m}$ equals 1 for forecasts made between 23 October 2000 and 20 December 2002. $Glob_{t,m}$ equals 1 for forecasts made after 20 December 2002. A negative sign for the coefficient of $RegFD_{t,m}$ or $Glob_{t,m}$ would indicate a decline in the bias following, respectively, Reg FD and the Global Settlement.

Lim (2001) argued that the forecast bias is higher when a company's information environment is less transparent—for example, when the company is small and has less analyst coverage. Beckers, Steliaros, and Thomson (2004) showed that the number of analysts following a stock affects the accuracy of the consensus earnings forecast. Hence, we used analyst coverage and company size as proxies for the degree of information transparency. Analyst coverage, $NumA_{j,t,m}$, is defined as the number of outstanding forecasts used in I/B/E/S's monthly consensus calculation. Analyst coverage represents the number of analysts following company j in month m for fiscal year t. $CompanySize_{j,t,m-1}$ is defined as the natural log of the company's market capitalization at the end of the previous month.

Analysts tend to forecast more accurately when they have more experience and resources (Clement 1999; Lim 2001). We measured company-specific experience as the number of years analyst i has been following company j ($YearStk_{j,t,m}$). We measured general experience as the number of years since analyst i first appeared in the I/B/E/S database ($YearIBES_{j,t,m}$). $BrokerSize_{j,t,m}$ is the number of analysts who work for the same employer during the same forecast year as the analyst who makes the forecast. Analysts who work for larger firms tend to have more resources at their disposal.

Clement (1999) found that analysts' forecasts are less accurate the more stocks and the more industries they follow. $NumStk_{j,t,m}$ is the number of stocks for which analyst i supplies at least one forecast within the calendar year. $NumInd_{j,t,m}$ is the number of two-digit SIC industries for which analyst i supplies at least one forecast within the calendar year.

Previous studies have found that forecasting is more difficult when companies report a loss or a decline in earnings (Brown 2001). The $EPSLoss_{j,t}$ indicator equals 1 when the corresponding actual earnings of company j are negative. The $EPSDecline_{j,t}$ indicator equals 1 when actual earnings in fiscal year t are lower than actual earnings in the previous year.

Matsumoto (2002) argued that companies in industries with a higher risk of shareholder lawsuits and/or greater reliance on implicit claims with stakeholders are more likely to avoid missing analyst forecasts. The $Litigation_j$ indicator equals 1 for companies in high-litigation-risk industries: SIC codes 2833–2836 (biotechnology), 3570–3577 and 7370–7374 (computers), 3600–3674 (electronics), and 5200–5961 (retailing).

Matsumoto (2002) also argued that labor-intensive companies try to avoid missing analyst forecasts because their stakeholders are concerned about company credit risk. Labor intensity, $Labor_{j,t,m-1}$, is defined as 1 minus the ratio of gross plant, property, and equipment (PPE) to total gross assets, where gross PPE is the quarterly Compustat item 118 and total gross assets is item 44 plus item 41. $Labor_{j,t,m-1}$ is measured at the end of the last quarter preceding forecast month m.

Richardson, Teoh, and Wysocki (2004) found lower forecast bias for companies with high growth opportunities. We used the market-to-book ratio ($MB_{j,t,m-1}$) at the end of the last quarter preceding the forecast month as a proxy for growth opportunities. The ratio is calculated as the market value of equity divided by the book value of common equity (Compustat quarterly data item 14 multiplied by item 61 and divided by item 59).

We used both the real GDP growth rate and the unexpected change in the real GDP growth rate to capture analysts' inability to forecast earnings accurately if the state of the economy changes substantially. $ActualGDP_t$ is the actual real GDP growth rate in fiscal year t. $UnexpectedGDP_{t,m}$ is defined as the difference between the expected real GDP growth rate and the actual real GDP growth rate in fiscal year t. For earnings forecasts made more than nine months before the fiscal year-end date, the expected real GDP growth rate in fiscal year t is defined as the real GDP growth rate in the quarter for which analysts made earnings forecasts. For forecasts made in Q2 (seven to nine months before the fiscal year-end date), we calculated the expected real GDP growth rate as (Growth in Q1 + 3 × Growth in Q2)/4. For forecasts made in Q3 (four to six months before the fiscal year-end date), we calculated the expected real GDP growth rate as (Growth in Q1 + Growth in Q2 + 2 × Growth in Q3)/4. For forecasts made within the three months before the fiscal year-end date, $UnexpectedGDP_{t,m}$ is set to zero.

Prior research and our results in Table 2 show that forecasts made earlier in the fiscal year are less accurate (Richardson, Teoh, and Wysocki 2004). To control for forecast horizon, we used $Month_m$, defined as the number of months until the fiscal year-end date. For example, for an analyst forecast made in October 1999 for the fiscal year ended December 1999, $Month_m$ equals 2. Richardson, Teoh, and Wysocki (2004) found that forecast bias has been declining gradually since the early 1990s. To address the concern that our results may be driven by this trend, we included a calendar year variable, $Year_t$, in the regression model (Equation 4). To control for unobserved company effects, we estimated the regressions with fixed company effects ($DCompany_j$).

The first set of estimation results in **Table 3** is for the regression model (Equation 4). The results imply that forecast bias declined by 0.24 percent of the stock price after the introduction of Reg FD. This finding confirms our earlier conjecture that the increase in forecast bias following Reg FD (observed in our univariate results) was driven by unexpectedly poor macroeconomic conditions. The decline in forecast bias following Reg FD is consistent with Lim's prediction (2001) that analysts become less optimistic when they rely less on insider information.

After the Global Settlement, the forecast bias is lower by 0.96 percent of the stock price compared with the forecast bias before Reg FD. This result is consistent with our univariate findings and implies that the Global Settlement and related regulations successfully neutralized analysts' conflicts of interest. The positive coefficient on *Month* suggests the presence of the walk-down trend. Forecast bias is high for earlier forecasts and becomes lower over time. On average, forecast bias increases by 0.14 percent of the stock price per month with the length of the forecast horizon.

Because the Global Settlement is an enforcement agreement between U.S. regulators and the Big-12 banks, we next examined whether the impact of the Global Settlement is limited to the Big-12 banks or whether there are spillover effects on other analysts.[16] In a recent study, Barber, Lehavy, McNichols, and Trueman (2006) reported that the proportion of buy recommendations declined significantly among all analysts after the implementation of NASD Rule 2711. They also documented that the decline was stronger for the sanctioned banks. Whether the Global Settlement has had a differential impact on analyst forecast bias, however, remains an open question.

To identify the differential impacts of Reg FD and the Global Settlement on Big-12 analysts, we compared the bias in the forecasts of Big-12 analysts with the bias in the forecasts of other analysts. In a univariate comparison, we found that, on average, the forecasts of analysts working for the Big-12 banks are statistically significantly less biased than the forecasts of their counterparts in each of the three periods. The differences, however, are economically trivial. For example, the difference between the mean forecast bias of Big-12 analysts and that of other analysts is –0.04 percent of the share price in the pre–Reg FD period, –0.09 percent after Reg FD, and –0.05 percent after the Global Settlement.

Table 3. The Impact of Reg FD and the Global Settlement on Forecast Bias

	(1)		(2)	
	Coefficient	*t*-Statistic	Coefficient	*t*-Statistic
RegFD	–0.24**	–3.29	–0.16*	–2.05
Glob	–0.96**	–10.68	–0.86**	–9.51
CompanySize	0.65**	16.89	0.67**	17.52
NumA	0.02**	3.39	0.01**	2.68
MB	–0.03**	–5.97	–0.03**	–5.59
YearStk	0.01	1.58	0.01**	2.59
YearIBES	0.00	1.54	0.00	0.78
NumStk	0.00*	–2.38	0.00*	–2.05
NumInd	–0.01	–1.18	–0.01	–1.40
BrokerSize	0.00	–1.64	0.00	–0.41
EPSLoss	5.40**	43.20	5.23**	40.53
EPSDecline	2.40**	62.82	2.38**	60.63
Litigation	–0.03	–0.24	–0.08	–0.66
Labor	0.52	2.12	0.47	1.89
ActualGDP	–0.04*	–2.05	–0.03	–1.23
UnexpectedGDP	–0.03**	–6.26	–0.04**	–6.61
Big12			–0.06**	–3.05
Big12 × RegFD			–0.07*	–2.04
Big12 × Glob			0.03	1.34
Month	0.14**	51.70	0.13**	47.76
Year	0.03*	2.16	0.02	1.09
Adjusted R^2	0.46		0.45	
No. of observations	434,268		434,268	
No. of companies	7,315		7,315	

Notes: This table presents the coefficients obtained from Equation 4. The dependent variable is earnings forecast bias, defined as the difference between the mean of all forecasts made in a particular month for a particular company and a particular fiscal year and the realized EPS, scaled by the stock price and multiplied by 100. The RegFD indicator equals 1 for forecasts made between 23 October 2000 and 20 December 2002. The Glob indicator equals 1 for forecasts made after 20 December 2002. Analyst coverage, NumA, is the number of outstanding forecasts used by I/B/E/S to calculate monthly consensus. CompanySize is the natural log of a company's market capitalization. Market-to-book ratio, MB, is the market value of equity divided by the book value of common equity. Company-specific experience, YearStk, is the number of years since the analyst made her first forecast for a particular stock. General experience, YearIBES, is the number of years since the first day the analyst appeared in I/B/E/S. NumStk and NumInd are the number of stocks and the number of industries covered by the analyst, respectively. The EPSLoss indicator equals 1 when the corresponding actual earnings of company *j* are negative. The EPSDecline indicator equals 1 when the realized earnings in fiscal year *t* are lower than the realized earnings in the previous year. BrokerSize is the number of analysts working for the employer of the analyst who makes the forecast. The litigation risk indicator, Litigation, equals 1 for companies in high-litigation-risk industries. Labor intensity, Labor, is (1 – Gross PPE/Total gross assets). The regressions are estimated with fixed company effects. The reported *t*-statistics reflect robust standard errors adjusted for heteroscedasticity and clustering by company.

*Significant at the 5 percent level.
**Significant at the 1 percent level.

To see whether the differential impacts of Reg FD and the Global Settlement on Big-12 and other analysts change when we control for company and analyst characteristics, as well as economic conditions, we re-estimated the regression model (Equation 4) with the Big-12 indicator and its interactions with the Reg FD and Global Settlement indicators included as additional independent variables.[17] The second set of results in Table 3 is for this regression. Consistent with our univariate results, the Big-12 indicator and its interaction with Reg FD are significant in statistical but not in economic terms. More importantly, the interaction of the Big-12 indicator with the Glob indicator is insignificant, both statistically and economically.

These results imply that both Big-12 and other analyst forecasts were biased before Reg FD, which is consistent with Lin and McNichols (1998), who found no difference between the earnings forecasts of analysts affiliated with banks involved in underwriting deals with the covered companies and the forecasts of unaffiliated analysts. These results also imply that the impact of the Global Settlement and related regulations is the same among Big-12 and other analysts. This finding may reflect the fear of non-Big-12 firms that they may become targets of similar investigations. In addition, because Big-12 banks no longer reward optimism, the incentive for lower-tier analysts to make optimistic forecasts as a means of moving up to the bigger banks has also been reduced. Finally, the rules and regulations introduced by the SEC, NYSE, and NASD around the time of the Global Settlement covered all analysts.

We checked the robustness of our main conclusion—that forecast bias declined after both Reg FD and the Global Settlement—in a number of ways. First, we used an alternative definition of the forecast bias by normalizing it by the book value of equity per share.[18] Second, we changed the cutoff dates for each period by using the effective date of Rule 2711 instead of the announcement date of the Global Settlement. Third, to ensure that our conclusions were unaffected by changes in the sample composition across the three subperiods, we required at least one forecast by the same analyst for the same company in all three periods. Fourth, we dropped observations with stock prices under $5 to avoid any potential biases induced when the scaling factor is a small number. Fifth, we extended our sample period to include an earlier period (January 1984–December 1995). In all these cases, the results (not reported here) remain qualitatively the same as those reported in Table 3, confirming that forecast bias declined after Reg FD and especially after the Global Settlement.

We also examined the breadth of these effects by estimating forecast bias regressions (Equation 4) separately for 12 business sectors and for subsamples formed on the basis of annual quintile sorts by company size and analyst coverage.[19] The results (not reported here) show that the effects of the Global Settlement are negative for 11 of 12 sectors and are statistically significant for 9 sectors. The effects of Reg FD are negative for 8 of 12 sectors, but significantly so for only 6 sectors. Our results also show that the effect of Reg FD is concentrated among smaller companies and companies with low analyst coverage, whereas the effect of the Global Settlement is more widespread, with no clear cross-sectional pattern.

Conclusion

Analysts' conflicts of interest were evident before the Global Research Analyst Settlement and were not limited to the 12 banks covered by it. Reg FD made analysts less dependent on insider information and thus diminished analysts' motives to favor company managers by inflating their earnings forecasts. The impact of Reg FD is more significant for companies with a less transparent information environment in which insider information has the most value.

Introduced in 2002, the Global Settlement and related regulations had an even bigger impact than Reg FD on analyst behavior. After the Global Settlement, the mean forecast bias declined significantly, whereas the median forecast bias essentially disappeared. Although disentangling the impact of the Global Settlement from that of related rules and regulations aimed at mitigating analysts' conflicts of interest is impossible, forecast bias clearly declined around the time the Global Settlement was announced. These results suggest that the recent efforts of regulators have helped neutralize analysts' conflicts of interest.

We thank Donal Byard, Terrence Martell, and seminar participants at Baruch College for helpful comments. Armen Hovakimian gratefully acknowledges the financial support of the PSC-CUNY Research Foundation of the City University of New York.

This article qualifies for 1 CE credit, inclusive of 1 SER credit.

Notes

1. Several rules and regulations were enacted around the Global Research Analyst Settlement—for example, NASD Rule 2711, NYSE Rule 472, and Regulation Analyst Certification. Because they were introduced over a relatively short period, determining the separate impact of each one of these regulatory actions is impossible. Nevertheless, all these rules and regulations share the same goal of reducing analysts' conflicts of interest. Therefore, we use the term Global Settlement to represent all the rules and regulations enacted around the Global Research Analyst Settlement to address analysts' conflicts of interest.
2. Scherbina (2004) found a negative relationship between the estimated bias that arises from self-selection in coverage and subsequent stock returns. Her results suggest that retail

investors fail to adjust for the bias. Malmendier and Shanthikumar (2007) found that retail investors react to stock recommendations literally. Institutional investors buy stocks that have "strong buy" ratings and sell stocks that have "buy" ratings, whereas retail investors buy in both cases. Kwag and Shrieves (2006) found that persistence in forecast errors can lead to potentially profitable trading strategies.

3. Overall, these studies found either no change (Bailey, Li, Mao, and Zhong 2003) or a decrease in forecast accuracy (Agrawal, Chadha, and Chen 2006; Mohanram and Sunder 2006) and forecast dispersion (Agrawal, Chadha, and Chen 2006) following Reg FD.
4. Kadan, Madureira, Wang, and Zach (2009) documented that stock recommendations have become less optimistic since the Global Settlement. Furthermore, they found that the likelihood of an optimistic recommendation is no longer associated with analyst affiliation. Ferreira and Smith (2006) found that investors have not changed the way they respond to analysts' changes in recommendations since Reg FD. Examining bid–ask spreads and trading activity following Reg FD, Lee, Rosenthal, and Gleason (2004) found no significant increase in volatility or in the adverse-selection component of bid–ask spreads.
5. Reg AC took effect on 14 April 2003. See the joint report by the NASD and NYSE (2005) for the effectiveness of the new rules.
6. The 10 investment banks are Bear Stearns, Citigroup, Credit Suisse First Boston, Goldman Sachs, J.P. Morgan, Lehman Brothers, Morgan Stanley, Merrill Lynch, UBS, and U.S. Bancorp Piper Jaffray. In 2008, Bear Stearns and Merrill Lynch were taken over because of their deteriorating financial positions, whereas Lehman Brothers ended up in bankruptcy. Because our sample period ends in 2006, these events did not affect our results.
7. These two investment banks are Deutsche Bank and Thomas Weisel Partners.
8. Because prior studies (e.g., Lin and McNichols 1998) found no cross-sectional differences in forecast bias between affiliated and unaffiliated analysts, one would not reasonably expect cross-sectional differences in the impact of the Global Settlement on these two analyst types.
9. Therefore, one would not reasonably expect cross-sectional differences in the impact of the Global Settlement on self-selection bias.
10. Forecasts for current-year EPS are the forecasts in I/B/E/S with code FPI 1. Forecasts for subsequent-year EPS are the forecasts in I/B/E/S with code FPI 2.
11. We excluded forecasts in the I/B/E/S Excluded Estimates file and forecasts for which actual earnings figures were missing.
12. The I/B/E/S Summary file contains monthly snapshots of consensus-level data and corresponding stock prices. The snapshots are as of the Thursday before the third Friday of every month. The reported stock prices in this file are the last available prices before the Thursday. I/B/E/S's earnings-related data and stock prices are split adjusted.
13. Using stock price to normalize forecast bias is common (see, e.g., Richardson, Teoh, and Wysocki 2004). Later in the article, we discuss the robustness of our findings to alternative scaling of analyst forecast errors.
14. We defined extreme values as those in 1 percent of both tails of the distribution. Variables that took only positive (negative) values were trimmed only on the right (left) tail of the distribution.
15. Realized earnings and forecasts are scaled by the stock price, consistent with the scaling of the bias measure.
16. Other regulations, such as NASD Rule 2711, affect all analysts.
17. In this analysis, for each forecast month of each sample company-year, the mean forecast bias is calculated separately for Big-12 and other analysts.
18. This step also ruled out the possibility that such events as the decimalization of stock prices in August 2000–April 2001 affected our findings.
19. The sector classification for each company is from the I/B/E/S Identifier file.

References

Agrawal, Anup, and Mark A. Chen. 2005. "Analyst Conflicts and Research Quality." Working paper, University of Maryland (August).

Agrawal, Anup, Sahiba Chadha, and Mark A. Chen. 2006. "Who Is Afraid of Reg FD? The Behavior and Performance of Sell-Side Analysts following the SEC's Fair Disclosure Rules." *Journal of Business*, vol. 79, no. 6 (November):2811–2834.

Bailey, Warren, Haitao Li, Connie X. Mao, and Rui Zhong. 2003. "Regulation Fair Disclosure and Earnings Information: Market, Analyst, and Corporate Responses." *Journal of Finance*, vol. 58, no. 6 (December):2487–2514.

Barber, Brad M., Reuven Lehavy, Maureen McNichols, and Brett Trueman. 2006. "Buys, Holds, and Sells: The Distribution of Investment Banks' Stock Ratings and the Implications for the Profitability of Analysts' Recommendations." *Journal of Accounting and Economics*, vol. 41, no. 1–2 (April):87–117.

Beckers, Stan, Michael Steliaros, and Alexander Thomson. 2004. "Bias in European Analysts' Earnings Forecasts." *Financial Analysts Journal*, vol. 60, no. 2 (March/April):74–85.

Brown, Lawrence D. 1997. "Analyst Forecasting Errors: Additional Evidence." *Financial Analysts Journal*, vol. 53, no. 6 (November/December):81–88.

———. 2001. "A Temporal Analysis of Earnings Surprises: Profits versus Losses." *Journal of Accounting Research*, vol. 39, no. 2 (September):221–241.

Carleton, Willard T., Carl R. Chen, and Thomas L. Steiner. 1998. "Optimism Biases among Brokerage and Non-Brokerage Firms' Equity Recommendations: Agency Costs in the Investment Industry." *Financial Management*, vol. 27, no. 1 (Spring):17–30.

Chan, Louis K.C., Jason J. Karceski, and Josef Lakonishok. 2007. "Analysts' Conflicts of Interest and Biases in Earnings Forecasts." *Journal of Financial and Quantitative Analysis*, vol. 42, no. 4 (December):893–913.

Chopra, Vijay Kumar. 1998. "Why So Much Error in Analysts' Earnings Forecasts?" *Financial Analysts Journal*, vol. 54, no. 6 (November/December):35–42.

Clarke, Jonathan, Ajay Khorana, Ajay Patel, and Raghavendra Rau. 2006. "Independents' Day? Analyst Behavior Surrounding the Global Settlement." Working paper, Georgia Institute of Technology (March).

Clement, Michael B. 1999. "Analyst Forecast Accuracy: Do Ability, Resources, and Portfolio Complexity Matter?" *Journal of Accounting and Economics*, vol. 27, no. 3 (July):285–303.

Doukas, John A., Chansog Kim, and Christos Pantzalis. 2005. "The Two Faces of Analyst Coverage." *Financial Management*, vol. 34, no. 2 (Summer):99–125.

Easterwood, John C., and Stacey R. Nutt. 1999. "Inefficiency in Analysts' Earnings Forecasts: Systematic Misreaction or Systematic Optimism?" *Journal of Finance*, vol. 54, no. 5 (October):1777–1797.

Ferreira, Eurico J., and Stanley D. Smith. 2006. "Effect of Reg FD on Information in Analysts' Rating Changes." *Financial Analysts Journal*, vol. 62, no. 3 (May/June):44–57.

Herrmann, Donald, Ole-Kristian Hope, and Wayne B. Thomas. 2008. "International Diversification and Forecast Optimism: The Effects of Reg FD." *Accounting Horizons*, vol. 22, no. 2 (June): 179–197.

Hong, Harrison, and Jeffrey D. Kubik. 2003. "Analyzing the Analysts: Career Concerns and Biased Earnings Forecasts." *Journal of Finance*, vol. 58, no. 1 (February):313–351.

Jackson, Andrew R. 2005. "Trade Generation, Reputation, and Sell-Side Analysts." *Journal of Finance*, vol. 60, no. 2 (April): 673–717.

Kadan, Ohad, Leonardo Madureira, Rong Wang, and Tzachi Zach. 2009. "Conflicts of Interest and Stock Recommendations: The Effects of the Global Settlement and Related Regulations." *Review of Financial Studies*, vol. 22, no. 10 (October):4189–4217.

Kwag, Seung-Woog, and Ronald E. Shrieves. 2006. "Chronic Bias in Earnings Forecasts." *Financial Analysts Journal*, vol. 62, no. 1 (January/February):81–96.

Lee, Chun I., Leonard Rosenthal, and Kimberly Gleason. 2004. "Effect of Regulation FD on Asymmetric Information." *Financial Analysts Journal*, vol. 60, no. 3 (May/June):79–89.

Lim, Terence. 2001. "Rationality and Analysts' Forecast Bias." *Journal of Finance*, vol. 56, no. 1 (February):369–385.

Lin, Hsiou-wei, and Maureen F. McNichols. 1998. "Underwriting Relationships, Analysts' Earnings Forecasts and Investment Recommendations." *Journal of Accounting and Economics*, vol. 25, no. 1 (February):101–127.

Malmendier, Ulrike, and Devin Shanthikumar. 2007. "Are Small Investors Naive about Incentives?" *Journal of Financial Economics*, vol. 85, no. 2 (August):457–489.

Matsumoto, Dawn A. 2002. "Management's Incentives to Avoid Negative Earnings Surprises." *Accounting Review*, vol. 77, no. 3 (July):483–514.

McNichols, Maureen F., and Patricia C. O'Brien. 1997. "Self-Selection and Analyst Coverage." *Journal of Accounting Research*, vol. 35 (Supplement):167–199.

Mohanram, Partha S., and Shyam V. Sunder. 2006. "How Has Regulation FD Affected the Operations of Financial Analysts?" *Contemporary Accounting Research*, vol. 23, no. 2 (Summer):491–525.

NASD and NYSE. 2005. "Joint Report by NASD and the NYSE: On the Operation and Effectiveness of the Research Analyst Conflict of Interest Rule." National Association of Securities Dealers and the New York Stock Exchange (December): www.nyse.com/pdfs/rajointreport.pdf.

Nutt, Stacey R., John C. Easterwood, and Cintia M. Easterwood. 1999. "New Evidence on Serial Correlation in Analyst Forecast Errors." *Financial Management*, vol. 28, no. 4 (Winter):106–117.

Richardson, Scott, Siew Hong Teoh, and Peter D. Wysocki. 2004. "The Walk-Down to Beatable Analyst Forecasts: The Role of Equity Issuance and Insider Trading Incentives." *Contemporary Accounting Research*, vol. 21, no. 4 (Winter):885–924.

Scherbina, Anna. 2004. "Analyst Disagreement, Forecast Bias and Stock Returns." Working paper, Harvard University (June).

When Sell-Side Analysts Meet High-Volatility Stocks: An Alternative Explanation for the Low-Volatility Puzzle[1]

Jason C. Hsu[2] **Hideaki Kudo**[3] **Toru Yamada**[4]

Abstract

Empirically, high-volatility stocks tend to deliver low average returns; this result is robust globally and has been documented in various studies. We confirm this finding using a global equity dataset that includes emerging markets data. We also show that high-volatility stocks exhibit high analyst bias in earnings growth forecasts. Although sell-side analysts are predictably optimistic, the relationship between the degree of optimism and a stock's volatility has not been documented before. We hypothesize that analysts inflate earnings forecasts more aggressively for volatile stocks, in part because the inflation would be more difficult for investors to detect. Because investors are known to overreact to analyst forecasts (under-adjust to analyst bias), this can lead to systematic overvaluation and low returns for high-volatility stocks. Additionally, we find sell-side analysts' research informative despite the analysts' biases; stocks that have high forward E/P ratios based on analyst earnings forecasts tend to outperform and produce significantly positive Fama–French alphas. This evidence rejects the cynical view of some in our industry that sell-side analysts are unskilled. More interestingly, we find high forward E/P stocks also exhibit high analyst bias, which supports an interpretation that analysts are more willing to inflate earnings forecasts for stocks that they believe are likely to deliver high returns—or for which their inflated forecasts are likely to do no harm.

[1] We would like to thank Isao Uesaki and Vivek Vishwanathan for their comments and criticisms, and Katy Sherrerd for her editing assistance.

[2] Research Affiliates and UCLA Anderson School of Management.
[3] Nomura Asset Management.
[4] Nomura Asset Management.

1. Introduction

Somewhat counter to the general intuition, empirical research shows that high-volatility stocks tend to deliver lower average returns than low-volatility stocks. Various explanations of this "puzzle" have been hypothesized, but the topic remains an active area for theoretical research. This paper is empirical in nature and primarily aims to document a new pattern in analyst earnings growth forecast bias in the cross-section for stocks. We also seek to contribute to the low-volatility puzzle literature by arguing that analyst behavior may partially explain the low-volatility anomaly.

We extend the research in two ways. First, we replicate the low-volatility effect using a global dataset that includes emerging markets data. Our results show that the low-volatility effect is robust even after controlling for regions, industrial sectors, and various firm characteristics. Second, we explore a possible link between analyst forecasts and the performance of low- (or high-) volatility stocks and find that high-volatility stocks tend to experience high upward bias in analyst earnings growth forecasts; this cross-sectional relationship has not been identified before. Additionally, high bias (optimistic forecast) generally leads to low stock returns—an observation which suggests that investors underestimate the magnitude of the bias and therefore overreact to analyst growth forecasts.[5] These empirical facts and their interpretations fit neatly together to suggest a new linkage between analyst behaviors and the low-volatility puzzle. As we will discuss later, sell-side analysts have strategic reasons to prefer to inflate growth forecasts for volatile stocks. Because investors overreact to analyst growth forecasts, which creates excess demand for high-volatility stocks, this mechanism produces low returns for volatile stocks and can partially account for the low-volatility effect.

We also find that, despite the upward bias, analyst earnings forecasts are informative for trading. Our evidence suggests that sell-side analysts are likely more skilled than widespread industry cynicism would suggest, and their behaviors are not merely dictated by the incentive to

[5] See La Porta [1996], Dechow and Sloan [1997], Rajan and Servaes [1997], Dechow, Hutton, and Sloan [1999] and Hayes and Levine [2000] for evidence on and interpretation of investor overreaction to analyst growth forecasts.

maintain positive relationships with banking clients and prospects. Specifically, stocks with a high analyst-forecasted earnings-to-price (forward E/P) ratio tend to deliver significantly higher returns and positive Fama–French alphas—that is, *stocks that analysts find "cheap" based on their forecasts tend to subsequently outperform*.[6]

The outline of the paper is as follows. We first review the relevant literature on the low-volatility puzzle and sell-side analyst forecast bias. Next, we propose a simple model of analyst behavior, which can explain the low-volatility puzzle and predict a number of interesting equity return patterns. We then describe our global dataset that includes emerging countries. A key contribution of our research is in demonstrating that the low-volatility effect is robust globally and is not driven by country or sector effects or by firm characteristics. Using global equity data and the I/B/E/S database, we next document that high return volatilities are associated with high upward biases in analyst earnings growth forecasts. Finally, we document that analyst forecasts, although systematically biased upward, do indeed contain useful cross-sectional information regarding future stock returns. This last finding argues in favor of the skill and value of sell-side analyst research.

2. Literature Review

Low-Volatility Puzzle

The literature on the low volatility puzzle has typically examined the two components of volatility—systematic and idiosyncratic—separately. The earlier literature on the rejection of the CAPM found that low-beta stocks produce higher risk-adjusted returns than high-beta stocks.[7] These findings are related to the low-volatility effect because low- (high-) beta stocks are more likely to exhibit low (high) volatility. The low-beta effect does not, however, subsume

[6] Although secondary to the primary focus of our paper, our new findings suggest that not only do sell-side analysts express valuable information in their earnings forecasts, but that investors underreact to the information long (i.e., months) after the forecasts become available, allowing profitable trading strategies to be constructed based on clever manipulation of I/B/E/S data. This evidence is consistent with the findings of Womack [1996], Barber, Lehavy, McNichols, and Trueman [2001], Mikhail, Walther, and Willis [2004] and Li [2005] on investor underreaction to analyst recommendations.

[7] See Black, Jensen, and Scholes [1972], Miller and Scholes [1972], and Haugen and Heins [1975].

the low-volatility effect. More recent literature has focused on idiosyncratic volatility and has generally found that stocks with low idiosyncratic volatility tend to produce higher risk-adjusted returns than stocks with high idiosyncratic volatility.[8] This finding is also related to the low-volatility puzzle since stocks with low idiosyncratic volatility usually exhibit low total volatility. Using developed-country equity data from 1985 to 2006, Blitz and van Vliet [2007] reported that low-volatility stocks outperformed high-volatility stocks. Frazzini and Pedersen [2011] also documented similar results using an expanded time horizon (1984–2009).

Various conjectures have been presented for explaining the low-beta and/or the low-idiosyncratic-volatility effect. Excellent syntheses of the related theories and empirical evidence has been provided by Baker, Bradley, and Wurgler [2011] and Pedersen and Frazzini [2011]. Baker, Bradley, and Wurgler summarized and argued the behavioral explanation for the low-volatility effect: investors are assumed to have a "preference for lotteries" and views high volatility stocks as speculation/gambling tools, which inflates the price for high-volatility stocks and depresses their future returns.[9] Rational asset managers are unable to arbitrage away this behavioral anomaly because over-weighting low-volatility stocks creates too much tracking error against their benchmarks.[10] Pedersen and Frazzini [2011] advocated a rational model in which investors are leverage constrained. In this model, investors use high-beta stocks to improve portfolio expected returns even though leveraging low-volatility stocks would produce better results. This excess demand for high-volatility stocks results in high prices in the present day followed by low future returns for these securities.[11] Because all investors are leverage and shorting constrained to varying degrees, the low-volatility premium is not arbitraged away. In the rational model, high beta stocks would have lower returns than "fair" but would not be expected to actually have lower returns than low beta stocks, which is what has been documented in a number of empirical studies.

In this paper, we provide another explanation for the low-volatility effect based on sell-side analyst behavior and investor reactions to analyst forecasts. We find that volatility can be a proxy for analyst bias—high-volatility stocks tend to experience more analyst optimism.

---

8 See Malkiel and Xu [2002], Spiegel and Wang [2006], Ang et al. [2006, 2009], and Bali and Cakici [2008].

9 See Mitton and Vorkink [2007], Barberis and Huang [2008] and Kumar [2009] for more detailed discussions regarding the investor preference for lottery-like payoffs and for high-volatility stocks.

10 See Brennan [1993] and Brennan, Cheng, and Li [2012] for more detailed discussions of the theoretical motivation for and the empirical evidence that supports why benchmark-sensitive institutional equity managers are unwilling to take advantage of the low-volatility premium.

11 The original insight into the effect of leverage constraints was provided by Black [1972].

Since the market is fooled, partly by the rosy forecasts, this leads to high prices and low returns for high-volatility stocks.

Sell-Side Analyst Behavior

It is well known that sell-side analysts tend to issue upward-biased earnings forecasts; anecdotal evidence and theoretical research suggest that the optimism may be strategic rather than indicative of a lack of skill.[12,13] Interestingly, despite the strong evidence on sell-side analyst optimism, investors do not seem to properly adjust for this bias. For stocks that are associated with high analyst optimism, the literature documents initial price overreaction to the rosy forecasts, followed by mean-reversion when high growth fails to materialize.[14]

Because investors do not fully adjust for sell-side analyst optimism, the ability to forecast analyst bias for stocks can be a valuable tool for investors**.** Frankel and Lee [1998] hypothesized that analysts, like naïve investors, can exhibit the behavioral tendency to over-extrapolate recent firm growth in making their own forecasts. They also found that growth-oriented stocks—those with high P/B ratios, high past sales growth, and high long-term earnings forecasts and ROE forecasts—tend to experience high analyst optimism. In this paper, we identify two additional stock characteristics—high volatility and high forward E/P—that predict analyst optimism. Our variables, however, are motivated by rational and strategic analyst behaviors and not by analysts' mistakes.

Although analysts are encouraged to produce rosy forecasts, they are also incentivized to provide high-quality research and profitable stock recommendations. Research finds that analyst reputation drives brokerage order flows.[15] Research also supports that analyst promotions are related to their relative forecast accuracy and the profitability of their stock picks.[16] This finding, according to Francis and Philbrick [1993], suggests a complex optimization problem for sell-side analysts. Jackson [2005] claimed that an equilibrium can exist in which sell-side analysts inflate earnings growth forecasts, but these forecasts are still informative. Empirical evidence seems to

---

12 See Ramnath, Rock, and Shane [2008] for a comprehensive review of the analyst forecast literature as well as a suggested list of the unexplored questions in the literature.

13 See Francis and Philbrick [1993], Kang, O'Brien and Sivaramakrishnan [1994], Dugar and Nathan [1995], Lin and McNichols [1998], Michaely and Womack [1999], and Dechow, Hutton and Sloan [2000].

14 See Dechow and Sloan [1997], Rajan and Servaes [1997], Dechow, Hutton and Sloan [1999], and Purnanandam and Swaminathan [2004].

15 See Irvine [2004], Jackson [2005], and Cheng, Liu, and Qian [2006].

16 See Dechow, Hutton, and Sloan [2000] and Hong, Kubik, and Solomon [2000].

support the informativeness of analyst research in spite of the observed bias: Kim, Lin, and Slovin [1997] and Green [2006] found that early access to sell-side analyst stock picks leads to abnormal profits.

It is an interesting question to explore whether sell-side analyst stock recommendations are valuable when investors do not have privileged early access. In our paper, we are able to extract information from analyst forecasts by examining the forward E/P for stocks based on the sell-side analyst earnings forecast. We found that stocks with high forward E/P ratios based on publicly available I/B/E/S analyst 12-month earnings forecasts produced higher subsequent 12-month returns. This is a new finding in the sell-side analyst literature and is consistent with earlier results supporting market under-reaction to analyst recommendations.[17]

3. A Model of Analyst Behavior and an Explanation for the Low-Volatility Puzzle

We propose a simple model to reconcile the empirical observation that sell-side analyst earnings forecasts are upward biased and unreliable on the one hand, yet are informative in producing abnormal profits for investors on the other. Although sell-side analysts have been shown to display over-optimism regarding firm earnings growth, it is hard to believe that analyst forecasts are arbitrarily positive. Analysts are presumably skilled and rational economic agents who optimize their behaviors to satisfy competing objectives.[18] Sell-side research, considered by some to be valuable, can drive significant brokerage trade flows.[19] Thus, because sell-side research can influence client investment activities, analysts are rated and the rankings are publicized. Presumably, research quality rankings matter to the employer investment banks.

[17] Frankel and Lee [1998], using an accounting valuation method (the residual income model) based on analyst forecasts, found that analyst forecasts are informative for predicting long-term returns. Barber, Lehavy, McNichols and Trueman [2001] and Loh and Mian [2006] formed trading portfolios based on published analyst recommendations and produced abnormal profits.

[18] See Francis and Philbrick [1993].

[19] See Brennan and Chordia [1993], Hayes [1998], Conrad, Johnston, and Wahal [2001] and Irvine [2000].

Theoretical and empirical research support the thesis that forecast accuracy and stock recommendations are linked with analysts' promotions and turnover.[20]

On the flip side, theories and empirical evidence also suggest that relationships with investment banking clients and prospects could influence analysts to bias their earnings growth forecasts upward and to set target stock prices higher than they otherwise would.[21] So, how might a skilled sell-side analyst achieve the complex objective of producing rosy earnings growth forecasts without appearing obviously biased and, at the same time, providing profitable trading recommendations to clients?

We propose a simple model of analyst behavior that produces both (1) the observed cross-sectional pattern in which high-volatility stocks experience high analyst forecast bias and (2) forecasts that are informative for trading. Imagine that analysts are skilled at ascertaining the mean and standard deviation of earnings growth for the stocks they cover. These analysts need to produce quality research and profitable recommendations to further their careers and reputations, while at the same time remaining sensitive to senior management's desire to maintain investment banking relationships. We posit that there is an equilibrium behavior such that all analysts inflate their reported growth estimates upward by, say, half a standard deviation in order to (1) be investment banking business friendly[22] and (2) avoid detection for inflating growth forecasts in certain situations.

This equilibrium behavior would predict higher growth forecast bias for firms with higher earnings growth variability and would, in turn, predict higher return volatility for these firms. This prediction is consistent with our empirical finding that high-volatility stocks are associated with high analyst forecast bias. Further, because evidence suggests that investors do not fully appreciate the upward bias, and thus overreact to analyst optimism in the short run, volatile stocks tend to be overvalued and experience low subsequent returns. This could then explain, in part, the documented underperformance of high-volatility stocks.

Our simple model also posits that analysts express valuable information in their forecasts in order to signal their skill to clients and management, but they strategically obfuscate the

[20] See Mikhail, Walther, and Willis [1999], Hong, Kubik, and Solomon [2000], and Clarke and Subramanian [2006].

[21] See Dugar and Nathan [1995], Lin and McNichols [1998] and Clarke, Khorana, and Rau [2004].

[22] The literature primarily focuses on the relationship between analyst earnings forecast inflation and the investment banking client relationship. Evidence also exists, however, that investment banks use inflated earnings growth to justify high price targets and strong buy recommendations in order to encourage more trading for their brokerage businesses (see Irvine [2000]).

information in an attempt to provide client-friendly inflated forecasts. If true, this suggests that profitable trading information can be potentially backed out of biased analyst forecasts; investors simply need to decode the analyst signal more effectively. We know that analysts overwhelmingly prefer to communicate equity attractiveness using E/P ratios,[23] so we can interpret the forward E/P ratio as a proxy for the analyst's private information on the attractiveness of a stock.

In our research, we find that stocks with high forward E/P forecasts outperform stocks with low forward E/P forecasts. Thus, while the complex strategic behavior of analysts leads to persistent upward bias and poor reliability in analysts' published growth forecasts, we find evidence that analysts are still able to communicate valuable recommendations through forward E/P forecasts. Our new evidence that analysts are more skilled than would be suggested by their lack of forecasting accuracy is, if anything, a vindicating discovery for sell-side analysts, given the prevailing industry wisdom regarding the value of their research.

4. Data

Our global equity dataset represents a broader dataset than has been used in previous research on the low-volatility premium puzzle; specifically, we expand the global dataset to include emerging markets. We use the I/B/E/S database to gather consensus analyst earnings forecasts. For each stock in the I/B/E/S database, the consensus earnings forecast is generally provided for at least the next two fiscal years. At the start of each fiscal year, the database records the reported previous fiscal year earnings per share (EPS) and also reports the consensus fiscal year-end EPS forecast for the current fiscal year and the following fiscal year. Table 1 shows the I/B/E/S monthly data structure for Company A, which has a fiscal year ending in September. At month-end October 2000, the database records realized EPS for the prior fiscal year (1999) as well as the consensus forecast for the current fiscal year (2000), which ends September 2001, and the next fiscal year (2001), which ends September 2002. We denote the prior fiscal year as FY0, the current fiscal year as FY1, and the next fiscal year as FY2.

[23] See Block [1999], Bradshaw [2004] and Demirakos, Strong, and Walker [2004].

Table 1. I/B/E/S Consensus Forecast for Company A

Date	Previous Fiscal Year End (FY0)	Current Fiscal Year End (FY1)	Next Fiscal Year End (FY2)	FY0 Realized EPS	FY1 Consensus EPS Forecast	FY2 Consensus EPS Forecast
⋮	⋮	⋮	⋮	⋮	⋮	⋮
Jun-2000	Sep-1999	Sep-2000	Sep-2001	1.35	1.93	2.18
Jul-2000	Sep-1999	Sep-2000	Sep-2001	1.35	1.95	2.25
Aug-2000	Sep-1999	Sep-2000	Sep-2001	1.35	1.95	2.25
Sep-2000	Sep-1999	Sep-2000	Sep-2001	1.35	1.95	2.25
Oct-2000	Sep-2000	Sep-2001	Sep-2002	1.76	1.21	1.39
Nov-2000	Sep-2000	Sep-2001	Sep-2002	1.76	1.19	1.39

A key variable of interest is the analyst forecast bias for current fiscal-year EPS. Analyst forecast bias is simply the time-series average of the forecast errors or the differences between the consensus EPS estimates and the subsequent realized EPS numbers. Operationally, we define the forecast error for Company A associated with the month of October 2000 as the 12-Month-Forward Realized EPS minus the 12-Month-Forward Consensus EPS Forecast. The forward consensus EPS is the time-weighted average of the current and next year's consensus EPS, and the forward realized EPS is also the time-weighted average. Because EPS_t is neither standardized (EPS_t gives no information for making cross-sectional comparisons) nor stationary (EPS_t generally grows over time and is unbounded), we elect to work with a transformed variable, EPS_t/BPS_{t-1}. Dividing earnings per share by book value per share creates a variable that is standardized across stocks and is stationary. EPS_t/BPS_{t-1} is also referred to as the return on shareholder equity, or ROE_t.[24]

We do not have an explicit interest in ROE. We are merely interested in standardizing the EPS variable so that it can be more meaningfully compared on a cross-sectional and inter-temporal basis. Other transformations, such as EPS/Asset or EPS/Sales, would accomplish the same goal and produce similar analyses. We then define earnings growth as $(EPS_{\text{12 months forward}} - EPS_{\text{past 12 months}})/BPS$. We do not use the traditional definition of earnings growth, $EPS_{\text{12 months forward}}/EPS_{\text{past 12 months}}$, because EPS can often be negative and can switch signs from year to

[24] Here and hereafter, all subindex t are not necessary because the context makes the interpretation obvious. Incidentally, $t-1$ means the prior fiscal year, not the previous month.

year, so that the resulting growth rate measurement can become difficult to interpret.[25] For example, two extremely opposite earnings growth profiles—$2 per share last year declining to –$2 per share versus –$2 per share growing to $2 per share—would result in the same growth rate, which is clearly undesirable for our econometric examination.

Corporate accounting data are sourced from Worldscope and total return data are from IDC Exshares. The sample period for our study ranges from January 1987 through December 2011 for developed countries and from December 1994 through December 2011 for emerging countries.[26,27] All return-related statistics are computed using excess returns, which are calculated as the net return in excess of local three-month interest rates. Our universe of stocks draws from the union of the MSCI and FTSE index memberships across all developed and emerging market countries.[28]

Because we use I/B/E/S consensus and reported EPS in our study, our universe is restricted to stocks for which both variables are available. The average number of stocks in the unrestricted universe is 3,308 and 910 for the developed and emerging markets, respectively. After eliminating stocks without consensus EPS, the universe reduces to 2,846 for the developed markets[29] and 537 for the emerging markets. We examine the effect of the sample selection rules and conclude that they do not adversely influence our results. We do not report these tests for the brevity of exposition. For robustness, we have repeated the tests with "winsorized" outlier observations. We do not separately report these results as our research appears to be unaffected by outliers.

5. Portfolios Sorted on Volatility

Low-Volatility Premium in Developed and Emerging Markets

We begin our analysis by examining the pattern of returns in the cross-section of global stocks,

25 In very rare situations, book value per share can also be negative. We discard data points with negative book value per share.

26 Before January 1987 and December 1994, the numbers of stocks are too small.

27 For the study of analyst forecast biases, however, we need the next fiscal year realized earnings. This would reduce the sample range up to December 2009.

28 We follow the definition of countries used by the MSCI World (Developed Countries) Index and Emerging Markets Index.

29 The mean numbers of stocks are 1,138 for North America; 898 for Europe; 596 for Japan; and 214 for Asia Pacific ex-Japan.

sorted by volatility. At the end of each month, we rank stocks based on their volatility using the past five years of monthly data. We then report the annualized buy-and-hold return for each decile portfolio. We note, however, that in a simple global sort, the constituents for each volatility decile could be dominated by a particular country or global sector because stocks from a particular country or industry sector may share a similar level of volatility. As a result, country and/or sector effects can become indistinguishable from the volatility effect. Additionally, we observe that small-capitalization stocks tend to be more volatile than average. To adjust for the impact of country, sector, and firm characteristics, we perform a global volatility portfolio sort neutralizing these effects. Specifically, we sort on adjusted volatility using the following equation:

$$\log(Vol_i) = \beta_1 \cdot Size_i + \beta_2 \cdot BP_i + \sum_j \gamma_j \cdot SD_{i,j} + \sum_k \delta_k \cdot Ctry_{i,k} + \varepsilon_i, \qquad (1)$$

where Vol_i is the total volatility of stock i measured from the previous 60 months, $Size_i$ is the market capitalization at the end of the preceding month, $SD_{i,j}$ is a dummy variable for industrial sector j (as classified by GICS 10 sectors), $Ctry_{i,k}$ is a dummy for country k, and ε_i is the adjusted volatility residual net of the influences of country, sector, and firm characteristics. Using Equation (1), we compute the adjusted volatility for each stock in our global universe and then sort stocks into decile portfolios based on this adjusted measure.

We report the returns and characteristics of the adjusted volatility portfolios in Table 2. The decile portfolios D1 and D10, in the top panel, contain firms with the lowest and highest adjusted volatilities, respectively, for the developed markets. The quintile portfolios follow the same format and report results for the emerging markets. For the developed markets, the returns of the low-volatility portfolios are higher than those of the high-volatility portfolios, and the pattern is nearly monotonic. For the emerging markets, the low-volatility effect is not present when we only examine the quintile returns. When we include the Sharpe ratio term, the low-volatility puzzle is strong for both the developed and emerging market countries. We also note that when we eliminate the 1994–1998 sample period, which was characterized by unprecedented EM currency fluctuations, the low-volatility effects are statistically stronger. This pattern holds true for the global portfolios sorted using raw (unadjusted) volatilities, which we do not separately report. These results are consistent with what was reported by Blitz and van Vliet

[2007] and Frazinni and Pedersen [2011]. These results confirm that the low-volatility effect is robust globally and is not subsumed by the standard size and value anomalies or driven by country or industry differences.

Analyst Forecast Bias and Stock Volatility

In this section, we examine the portfolio characteristics associated with the various volatility decile portfolios. Table 3 reports the descriptive statistics such as book-to-price (B/P), earnings growth variability, average market capitalization, and so forth for the stocks in the decile portfolios. In addition, we report statistics on analyst earnings growth forecasts, subsequent realized growth, and analyst forecast bias. Again, we only report the statistics of portfolios formed on adjusted volatility, noting that the results are similar using raw volatilities.

Because the influences from countries, sectors, and firm characteristics are neutralized in the portfolio construction process, it is not surprising that the average market-cap and B/P characteristics are similar across the decile portfolios. The country and industry allocations are similar as well, but are not displayed in Table 3 for brevity. First, we observe that the earnings growth forecast biases, as measured by ($EPS_{\text{12-months-forward forecast}} - EPS_{\text{12-months-forward realized}}$)/BPS, are positive on average for stocks, meaning that analysts are systematically over-optimistic regarding future corporate earnings growth. This is consistent with the literature on upward bias in sell-side analyst forecasts. Additionally, we observe that the low-volatility portfolios generally have lower forecasted earnings growth as measured by ($EPS_{\text{12-months-forward forecast}} - EPS_{\text{past-12-months realized}}$)/BPS, but do not generally display lower realized earnings growth as measured by ($EPS_{\text{12-months-forward realized}} - EPS_{\text{past-12-months realized}}$)/BPS. This observation suggests an interesting pattern of analyst bias in the cross-section—analysts seem to be more optimistic on the more volatile stocks!

A Model of Sell-Side Analyst Behavior

The observation that return volatility is cross-sectionally correlated with analyst bias in earnings growth forecasts is a new empirical finding, which contributes to the literature on analyst forecast bias as well as to the literature on the low-volatility premium. Because this paper is empirical in nature, we propose a plausible story to rationalize this finding, but do not propose testable implications of the story to ascertain its validity against competing hypotheses.

As we discussed earlier, sell-side analyst behaviors are thought to be influenced by their desire (1) to maintain good relationships with investment banking clients and prospects, (2) to avoid damaging their reputation with brokerage clients who subscribe to analyst research reports, and (3) to achieve high rankings against other analysts in published quality rankings.

Empirical evidence supports the fact that sell-side analysts have superior abilities to analyze public information and are adept at producing valuable private information on companies. It is not unreasonable to model analysts as skilled at estimating the distribution of next-period earnings growth, $\tilde{g}_t$, for firms they cover. Note that realized earnings growth, $\tilde{g}_t$, is a random variable drawn from a distribution with mean g_t and standard deviation σ_t. More formally, each analyst *i* produces a forecast of $\hat{g}_{t,i}$ and $\hat{\sigma}_{t,i}$. The true skill of an analyst is determined by the deviation over time between $\hat{g}_{t,i}$ and the unobserved true mean $g_{t,i}$. Since $g_{t,i}$ cannot be observed, the skill of analyst *i* can only be estimated by the average difference between his forecast $\hat{g}_{t,i}$ and the realized $\tilde{g}_{t,i}$ over time.[30] Finally, analysts report a biased forecast, $G_{t,i}$, instead of their true private information, $\hat{g}_{t,i}$.

We assume that the utility function of the analysts is (1) increasing in the "optimism of the reported growth forecast," or $G_{t,i}-\hat{g}_{t,i}$; (2) decreasing in the "detectability of the forecast bias," or $(G_{t,i}-\hat{g}_{t,i})/\hat{\sigma}_{t,i}$; and (3) decreasing in distortion in valuation accuracy of the forecast, or $|EPS(G_{t,i})/P_t - EPS(\hat{g}_{t,i})/P_t|$, where $EPS(G_{t,i})/P_t$ is the forward E/P based on the reported forecast $G_{t,i}$, and $EPS(\hat{g}_{t,i})/P_t$ is the forward E/P based on the true forecast $\hat{g}_{t,i}$. Although these assumptions are naïve and incomplete as descriptions of reality, they are consistent with the empirical evidence on analysts' behaviors and incentives.

If the variability of earnings growth, σ_t, for firm *i* is extremely low, then large bias, $G_{t,i}-\hat{g}_{t,i}$, would be easy for brokerage clients to detect. An econometrically savvy investor can detect whether an analyst has been "pumping" stock prices through highly inflated forecasts (over the last *T* periods) by testing if $\frac{1}{T}\sum\left(G_{t,i}-\tilde{g}_t\right)/\tilde{\sigma}_T$ is significantly larger than zero, where $\tilde{g}_t$ and $\tilde{\sigma}_T$ are the realized earnings growth and variability. Analyst stock recommendations are usually justified by valuation multiples based on forward earnings. As a result, analysts would not want to inflate reported $G_{t,i}$ and next year's earnings $EPS(G_{t,i})$ so significantly that an unattractive stock (with low $EPS(\hat{g}_{t,i})/P_t$ based on the analyst's true forecast) appears attractive.

Without writing a formal mathematical model, we simply state that a repeated game

30 For simplicity, we assume that each analyst covers only one firm.

equilibrium exists whereby all analysts inflate their reported earnings growth forecasts relative to their private unbiased growth estimates by k times earnings growth variability. The scalar k is determined by (1) the benefit to the analyst from improving/maintaining investment banking client/prospect relationships through "friendly" outlooks, (2) the risk of being accused of "pump and dump" by brokerage clients, and (3) the benefit from providing quality stock recommendations to brokerage clients. Intuitively, in this equilibrium, analysts inflate growth forecasts by a careful amount to avoid losing credibility outright and to ensure that their forecasts can still result in forward E/P ratios, which lead to good buy/hold/sell recommendations.

Theoretically, return volatility has a positive relationship with earnings growth variability, which we confirm empirically in Table 3. This then suggests that more volatile stocks are more likely to receive greater analyst inflation in earnings growth forecasts. Since investors are documented to overreact to analyst growth forecasts, our model predicts low returns for high-volatility stocks.

6. Forward E/P and Stock Returns

High Forward E/P = High Returns

Another prediction of our simple model is that stocks with analyst-forecasted high forward E/P ratios will outperform stocks with low forward E/P ratios. In Table 4a, we show that developed market stocks in the top decile, as sorted by analyst-forecasted forward E/P ratios, produce a 6% higher annualized return than those in the bottom decile. The Sharpe ratios for the top and bottom deciles are 0.48 and 0.19, respectively. Similarly, for emerging market stocks, the top quintile stocks outperform the bottom quintile by nearly 10% per annum (a Sharpe ratio of 0.73 versus 0.35).[31]

The forward E/P ratio can be interpreted as a tool for analysts to communicate the attractiveness of stocks.[32] In the bottom panel of Tables 4a and 4b, we show that the information contained in an analyst's forward E/P is not subsumed by the Fama–French return model; specifically, stocks that analysts find attractive (in three of the top four deciles for developed

[31] The emerging markets data are likely significantly more noisy than the developed markets data. This might contribute to the lack of monotonicity in the returns and the Sharpe ratios of the sorted portfolios.

[32] See Demirakos, Strong, and Walker [2004].

markets and in the top quintiles for emerging markets) display significant Fama–French alphas. Brokerage clients with advanced access to analyst research and recommendations appear to achieve better investment performance.

Tables 4a and 4b show that the analyst-earnings-growth-forecast bias is increasing in the forward E/P. This is another novel empirical fact that we introduce into the literature. This observation suggests that analysts inflate the earnings growth forecasts more aggressively for stocks that they find attractive from a forward E/P perspective and do not tend to inflate the earnings as aggressively for stocks they find to be less attractive. On average, for stocks that analysts find most attractive in the developed markets (top decile by forward E/P), the upward growth bias is 7%, and in the emerging markets (top quintile), the bias is 6%. This behavior is consistent with our simple model in which the analyst prefers to inflate earnings as much as possible without losing credibility with clients. For stocks that analysts believe are likely to produce great returns, inflating earnings aggressively is less likely to create a poor experience for clients who trade on analyst forecasts.

Volatility and Forward E/P Double-Sorted Portfolios

To summarize our findings and to explore any potential interactions, we perform an unconditional double sort on volatility and forward E/P. We report the portfolio statistics in Table 5a for developed markets and in Table 5b for emerging markets. The new discovery that we make is that the low-volatility effect is much more pronounced for the low forward E/P stocks. In the developed markets, for low forward E/P stocks, the lowest volatility portfolio has a Sharpe ratio of 0.42 and the highest volatility portfolio has a Sharpe ratio of 0.11, a difference of 74%. For high forward E/P stocks, the Sharpe ratios for the lowest and highest volatility portfolios are 0.63 and 0.45, respectively, a difference of 28%. In the emerging markets, we observe the same pattern. For low forward E/P stocks, the low volatility portfolio has s Sharpe ratio of 0.39 compared to a Sharpe ratio of 0.26 for the high-volatility portfolio, which is a 33% difference, and for high forward E/P stocks, the corresponding Sharpe ratios are 0.61 and 0.55, respectively, a 9% difference.

Table 6 reports the corresponding Fama–French alphas for the double-sorted portfolios. The results show a general pattern in which alphas are large for high forward E/P stocks and low-volatility stocks and are small for low forward E/P stocks and high-volatility stocks. This result can be interpreted in the following way. Forward E/P is a proxy for analysts' valuable private information, which is communicated only to their brokerage firm's clients. Empirical evidence also shows that investors underreact to analysts' stock recommendations, and this makes the forward E/P information from the I/B/E/S database valuable for creating outperformance.

Volatility is a proxy for analyst bias. Conventional wisdom indicates that investors have some awareness of the sell-side analyst bias, yet empirical evidence suggests that investors still substantially overreact to analyst optimism (or under-appreciate the size of the analyst bias). The degree to which investors over- or underreact to different aspects of the analyst research report is succinctly captured in the cross-sectional pattern of the Fama–French alphas presented in Table 6. We believe this particular finding is novel and contributes to the empirical literature on investor over/under-reaction to the release of analyst research.

5. Conclusions

The contributions of this paper are mainly empirical; we want to be careful not to overstate the significance of our theoretical contribution. Given our emphasis on the empirical results, we attempt to contribute to the literature by offering plausible explanations for the low-volatility puzzle and the sell-side analyst behaviors discussed throughout the paper.

Our empirical results both confirm and extend the work of other researchers. We confirm the findings of low-volatility returns in global developed and emerging markets. When we explore possible linkages between the low-volatility findings and analyst forecasts, we find several interesting results. We find evidence that sell-side analysts are strategic in how they inflate earnings growth forecasts for stocks. It is well accepted that sell-side analysts have incentives to provide optimistic forecasts, and their positive bias has a very specific cross-sectional pattern. First, they tend to inflate earnings growth forecasts for more volatile stocks. We hypothesize that this is because it is harder for clients to detect inflation in growth forecasts for stocks that have highly volatile growth. Second, analysts tend to more aggressively inflate growth forecasts for stocks that they have strong positive information on. We suspect that this is because clients are less likely to complain about overly optimistic growth forecasts for stock recommendations that prove to be profitable.

These strategic behaviors by analysts can explain, partially, the low-volatility premium. High-volatility stocks are more likely to receive more inflated earnings forecasts. Because investors are tend to overreact to analyst optimism and are generally willing to overpay for stocks with high analyst bias, this would predict low returns for high-volatility stocks. More interestingly, we find that analyst forecasts, while biased upward, do result on average in the correct stock picks for their clients. Specifically, stocks with forecasted high forward E/P ratios tend to outperform stocks with forecasted low forward E/P ratios. The high E/P stocks also produce sizeable positive Fama–French alphas. Finally, we document that the low-volatility effect is significantly stronger for low forward E/P stocks than for high forward E/P stocks.

Our empirical findings are novel and add to the literature on analyst behavior. They also provide greater richness to and expand on the known cross-sectional pattern of volatility premia. Finally, they provide insights into a plausible new mechanism that uses sell-side analyst behaviors to explain the low-volatility premium. .

REFERENCES

Ang, A., R. Hodrick, Y. Xing, and X. Zhang. "The Cross-Section of Volatility and Expected Returns." *Journal of Finance*, vol. 61, no. 1 (2006), pp. 259–299.

——. "High Idiosyncratic Volatility and Low Returns: International and Further U.S. Evidence." *Journal of Financial Economics*, 91 (2009), pp. 1–23.

Arnott, R., F. Li, and K. Sherrerd. "Clairvoyant Value and the Value Effect." *Journal of Portfolio Management*, vol. 35, no. 3 (2009), pp. 12–26.

Baker, M., B. Bradley, and J. Wurgler. "Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly." NYU Working Paper No. FIN-10-002, 2010.

Barber, B., R. Lehavy, M. McNichols, and B. Trueman. "Can Investors Profit from the Prophets? Consensus Analyst Recommendations and Stock Returns." *Journal of Finance*, vol. 56, no. 2 (2001), pp. 531–563.

Barberis, N., and M. Huang. "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices." *American Economic Review*, vol. 98, no. 5 (2008), pp. 2066–2100.

Bali, T., and N. Cakici. "Idiosyncratic Volatility and the Cross-Section of Expected Returns." *Journal of Financial and Quantitative Analysis*, 43 (2008), pp. 29–58.

Blitz, D., and P. van Vliet. "The Volatility Effect: Lower Risk without Lower Return." *Journal of Portfolio Management*, 34 (2007), pp. 102–113.

Black, F. "Capital Market Equilibrium with Restricted Borrowing." *Journal of Business*, 45 (1972), pp. 444–454.

Black, F., N. Jensen, and M. Scholes. "The Capital Asset Pricing Model: Some Empirical Tests." In *Studies in the Theories of Capital Markets*, M. Jensen, ed. New York: Praeger, 1972.

Boyer, B., T. Mitton, and K. Vorkink. "Expected Idiosyncratic Skewness." *Review of Financial Studies*, vol. 23, no. 1 (2010), pp. 169–202.

Bradshaw, M. "The Use of Target Prices to Justify Sell-Side Analysts' Stock Recommendations." *Accounting Horizons*, 16 (2002), pp. 27–40.

——. "How Do Analysts Use Their Earnings Forecasts in Generating Stock Recommendations?" *Accounting Review*, 79 (2004), pp. 25–50.

Brav, A., and R. Lehavy. "An Empirical Analysis of Analysts' Target Prices: Short-Term Informativeness and Long-Term Dynamics." *Journal of Finance*, 58 (2003), pp. 1933–1968.

Brennan, M., X. Cheng, and F. Li. “Agency and Institutional Investment.” *European Financial Management*, vol. 18, no. 1 (2012), pp. 1–27.

Brennan, M., and T. Chordia. “Brokerage Commission Schedules.” *Journal of Finance*, 48 (1993), pp. 1379–1402.

Cao, X., and Y. Xu. “Long-Run Idiosyncratic Volatilities and Cross-Sectional Stock Returns.” Working paper, February 2010.

Clarke, R., H. de Silva, and S. Thorley. “Know Your VMS Exposure.” *Journal of Portfolio Management*, 36 (2010), pp. 52–59.

Clarke, J., A. Khorana, A. Patel, and P. Rau. “Analyst Behavior at Independent Research Firms, Brokerages, and Investment Banks: Conflicts of Interest or Better Information?” Working paper, Georgia Institute of Technology, Wake Forest University, and Purdue University, 2004.

Clarke, J., and A. Subramanian. “Dynamic Forecasting Behavior by Analysts: Theory and Evidence.” *Journal of Financial Economics*, vol. 81, no. 1 (2006), pp. 81–113.

Conrad, J., K. Johnston, and S. Wahal. “Institutional Trading and Soft Dollars.” *Journal of Finance*, vol. 56, no. 1 (2001), pp. 397–416.

Dechow, P., A. Hutton, and R. Sloan. “An Empirical Assessment of the Residual Income Valuation Model.” *Journal of Accounting and Economics*, 26 (1999), pp. 1–34.

Dechow, P., and R. Sloan. “Returns to Contrarian Investment Strategies: Tests of Naïve Expectations Hypotheses.” *Journal of Financial Economics*, 43 (1997), pp. 3–27.

Demirakos, E., N. Strong, and M. Walker. “What Valuation Models Do Analysts Use?” *Accounting Horizons*, vol. 18, no. 4 (2004), pp. 221–240.

Dugar, A. and S. Nathan. “The Effect of Investment Banking Relationships on Financial Analysts’ Earnings Forecasts and Investments Recommendations.” *Contemporary Accounting Research* vol. 12, no. 1 (1995), pp. 131–160.

Fama, E., and K. French. “Common Risk Factors in the Returns on Stocks and Bonds.” *Journal of Financial Economics*, 33 (1993), pp. 2–56.

Francis, J., and D. Philbrick. “Analysts’ Decisions as Products of a Multi-Task Environment.” *Journal of Accounting Research*, 31 (1993), pp. 216–230.

Francis, J., and L. Soffer. “The Relative Informativeness of Analysts’ Stock Recommendations and Earnings Forecast Revisions.” *Journal of Accounting Research*, 35 (1997), pp. 193–211.

Frankel, R., and C. Lee. “Accounting Valuation, Market Expectation, and Cross-Sectional Stock Returns.” *Journal of Accounting and Economics*, 9 (1998), pp. 195–228.

Frazzini, A., and L. Pedersen. “Betting Against Beta.” NBER Working Paper Series No.16601, 2011.

Fu, F. “Idiosyncratic Risk and the Cross-Section of Expected Stock Returns.” *Journal of Financial Economics*, vol. 91, no. 1 (2009), pp. 24–37.

Haugen, R., and J. Heins. “Risk and the Rate of Return on Financial Assets: Some Old Wine in New Bottles.” *Journal of Financial and Quantitative Analysis*, vol. 10, no. 5 (1975), pp. 775–784.

Hayes, R. “The Impact of Trading Commission Incentives on Analysts' Stock Coverage Decisions and Earnings Forecasts.” *Journal of Accounting Research*, 36 (1998), pp. 299–320.

Hayes, R., and C. Levine. “An Approach to Adjusting Analysts’ Consensus Forecasts for Selection Bias.” *Contemporary Accounting Research*, vol. 17, no. 1 (2000), pp. 61–83.

Hong, H., J. Kubik, and D. Solomon. “Security Analysts' Career Concerns and Herding of Earnings Forecasts.” *RAND Journal of Economics*, 31 (2000), pp. 121–144.

Huang, W., Q. Liu, S. Rhee, and L. Zhang. “Return Reversals, Idiosyncratic Risk, and Expected Returns.’’ *Review of Financial Studies*, vol. 23, no. 1 (2010), pp. 147–168.

Irvine, P. “Do Analysts Generate Trade for Their Firms? Evidence from the Toronto Stock Exchange.” *Journal of Accounting and Economics*, 30 (2000), pp. 209–226.

Irvine, P. “Analysts’ Forecasts and Brokerage-Firm Trading.” *The Accounting Review*, 79:1 (2004), pp. 125–149.

Jiang, G., D. Xu, and T. Yao. “The Information Content of Idiosyncratic Volatility.” *Journal of Financial and Quantitative Analysis*, vol. 44, no. 1 (2009), pp. 1–28.

Kang, S., J. O’Brien, and K. Sivaramakrishnan. “Analysts’ Interim Earnings Forecasts: Evidence on the Forecasting Process.” *Journal of Accounting Research*, 32 (1994), pp. 103–112.

Kumar, A. “Who Gambles in the Stock Market?” *Journal of Finance*, vol. 64, no. 4 (2009), pp. 1–28.

La Porta, R. “Expectations and the Cross-Section of Stock Returns.” *Journal of Finance*, 51 (1996), pp. 1715–1742.

Lin, H., and M. McNichols. “Underwriting Relationship, Analysts’ Earnings Forecasts and Investment Recommendations.” *Journal of Accounting and Economics*, vol. 25, no. 1 (1998), pp. 101–127.

Loh, R., and M. Mian. "Do Accurate Earnings Forecasts Facilitate Superior Investment Recommendations?" *Journal of Financial Economics*, vol. 80, no. 2 (2006), pp. 455–483.

Malkiel, B., and Y. Xu. “Idiosyncratic Risk and Security Returns.” Working Paper, University of Texas at Dallas, 2002.

Michaely, R., and K. Womack. “Conflict of Interest and the Credibility of Underwriter Analyst Recommendations.” *Review of Financial Studies*, 12 (1999), pp. 653–686.

Mikhail, M., B. Walther, and R. Willis. “Do Security Analysts Exhibit Persistent Differences in Stock Picking Ability?” *Journal of Financial Economics*, 74 (2004), pp. 67–91.

Miller, M., and M. Scholes. “Rate of Return in Relation to Risk: A Reexamination of Some Recent Findings.” In *Studies in the Theory of Capital Markets*, M. Jensen, ed. New York: Praeger, 1972, pp. 47–78.

Mitton, T., and K. Vorkink. “Equilibrium Underdiversification and the Preference for Skewness.” *Review of Financial Studies*, vol. 20, no. 4 (2007), pp. 1255–1288.

Morgan, J., and P. Stocken. “An Analysis of Stock Recommendations.” *RAND Journal of Economics*, vol. 34, no. 1 (2003), pp. 183–203.

Purnanandam, A., and B. Swaminathan. “Are IPOs Really Underpriced?” *Review of Financial Studies*, 17 (2004), pp. 811–848.

Rajan, R., and H. Servaes. "Analyst Following of Initial Public Offerings." *Journal of Finance,* vol. 52, no. 2 (1997), pp. 507–529.

Ramnath, S., S. Rock, and P. Shane. “The Financial Analyst Forecasting Literature: A Taxonomy with Suggestions for Further Research.” *International Journal of Forecasting*, 24 (2008), pp. 34–75.

Spiegel, M., and X. Wang. “Cross-Sectional Variation in Stock Returns: Liquidity and Idiosyncratic Risk.” Working paper, Yale University, 2006.

Tversky, A., and D. Kahneman. “Advances in Prospect Theory: Cumulative Representation of Uncertainty.” *Journal of Risk and Uncertainty*, vol. 5, no. 4 (1992), pp. 297–323.

Womack, K. “Do Brokerage Analysts’ Recommendations Have Investment Value?” Working paper, University of Toronto, 1996.

Long-Run Stock Returns: Participating in the Real Economy

Roger G. Ibbotson and Peng Chen

In the study reported here, we estimated the forward-looking long-term equity risk premium by extrapolating the way it has participated in the real economy. We decomposed the 1926–2000 historical equity returns into supply factors—inflation, earnings, dividends, the P/E, the dividend-payout ratio, book value, return on equity, and GDP per capita. Key findings are the following. First, the growth in corporate productivity measured by earnings is in line with the growth of overall economic productivity. Second, P/E increases account for only a small portion of the total return of equity. The bulk of the return is attributable to dividend payments and nominal earnings growth (including inflation and real earnings growth). Third, the increase in the equity market relative to economic productivity can be more than fully attributed to the increase in the P/E. Fourth, a secular decline has occurred in the dividend yield and payout ratio, rendering dividend growth alone a poor measure of corporate profitability and future growth. Our forecast of the equity risk premium is only slightly lower than the pure historical return estimate. We estimate the expected long-term equity risk premium (relative to the long-term government bond yield) to be about 6 percentage points arithmetically and 4 percentage points geometrically.

Numerous authors are directing their efforts toward estimating expected returns on stocks incremental to bonds.[1] These equity risk premium studies can be categorized into four groups based on the approaches the authors took. The first group of studies has attempted to derive the equity risk premium from the historical returns of stocks and bonds; an example is Ibbotson and Sinquefield (1976a, 1976b). The second group, which includes our current work, has used fundamental information—such as earnings, dividends, or overall economic productivity—to measure the expected equity risk premium. The third group has adopted demand-side models that derive expected equity returns through the payoff demanded by investors for bearing the risk of equity investments, as in the Ibbotson, Diermeier, and Siegel (1984) demand framework and, especially, in the large body of literature following the seminal work of Mehra and Prescott (1985).[2] The fourth group has relied on opinions of investors and financial professionals garnered from broad surveys.

In the work reported here, we used supply-side models. We first used this type of model in Diermeier, Ibbotson, and Siegel (1984). Numerous other authors have used supply-side models, usually with a focus on the Gordon (1962) constant-dividend-growth model. For example, Siegel (1999) predicted that the equity risk premium will shrink in the future because of low current dividend yields and high equity valuations. Fama and French (2002), studying a longer time period (1872–1999), estimated a historical expected geometric equity risk premium of 2.55 percentage points when they used dividend growth rates and a premium of 4.32 percentage points when they used earnings growth rates.[3] They argued that the increase in the P/E has resulted in a realized equity risk premium that is higher than the *ex ante* (expected) premium. Campbell and Shiller (2001) forecasted low returns because they believe the current market is overvalued. Arnott and Ryan (2001) argued that the forward-looking equity risk premium is actually negative. This conclusion was based on the low

Roger G. Ibbotson is professor of finance at Yale School of Management, New Haven, Connecticut. Peng Chen, CFA, is vice president and director of research at Ibbotson Associates, Chicago.

current dividend yield plus their forecast for very low dividend growth. Arnott and Bernstein (2002) argued similarly that the forward-looking equity risk premium is near zero or negative (see also Arnott and Asness 2003).

The survey results generally support somewhat higher equity risk premiums. For example, Welch (2000) conducted a survey of 226 academic financial economists about their expectations for the equity risk premium. The survey showed that they forecasted a geometric long-horizon equity risk premium of almost 4 pps.[4] Graham and Harvey (2001) conducted a multiyear survey of chief financial officers of U.S. corporations and found their expected 10-year geometric average equity risk premium to range from 3.9 pps to 4.7 pps.[5]

In this study, we linked historical equity returns with factors commonly used to describe the aggregate equity market and overall economic productivity. Unlike some studies, ours portrays results on a per share basis (per capita in the case of GDP). The factors include inflation, EPS, dividends per share, P/E, the dividend-payout ratio, book value per share, return on equity, and GDP per capita.[6]

We first decomposed historical equity returns into various sets of components based on six methods. Then, we used each method to examine each of the components. Finally, we forecasted the equity risk premium through supply-side models using historical data.

Our long-term forecasts are consistent with the historical supply of U.S. capital market earnings and GDP per capita growth over the 1926–2000 period. In an important distinction from the forecasts of many others, our forecasts assume market efficiency and a constant equity risk premium.[7] Thus, the current high P/E represents the market's forecast of higher earnings growth rates. Furthermore, our forecasts are consistent with Miller and Modigliani (1961) theory, in that dividend-payout ratios do not affect P/Es and high earnings-retention rates (usually associated with low yields) imply higher per share future growth. To the extent that corporate cash is not used for reinvestment, we assumed it to be used to repurchase a company's own shares or, perhaps more frequently, to purchase other companies' shares. Finally, our forecasts treat inflation as a pass-through, so the entire analysis can be done in real terms.

Six Methods for Decomposing Returns

We present six different methods for decomposing historical equity returns. The first two methods (especially Method 1) are based entirely on historical returns. The other four methods are methods of the supply side. We evaluated each method and its components by applying historical data for 1926–2000. The historical equity return and EPS data used in this study were obtained from Wilson and Jones (2002).[8] The average compound annual return for the stock market over the 1926–2000 period was 10.70 percent. The arithmetic annual average return was 12.56 percent, and the standard deviation was 19.67 percent. Because our methods used geometric averages, we focus on the components of the 10.70 percent geometric return. When we present our forecasts, we convert the geometric average returns to arithmetic average returns.

Method 1. Building Blocks. Ibbotson and Sinquefield developed a "building blocks" model to explain equity returns. The three building blocks are inflation, the real risk-free rate, and the equity risk premium. Inflation is represented by changes in the U.S. Consumer Price Index (*CPI*). The equity risk premium for year t, ERP_t, and the real risk-free rate for year t, RRf_t, are given by, respectively,

$$ERP_t = \frac{1+R_t}{1+Rf_t} - 1 = \frac{R_t - Rf_t}{1+Rf_t} \tag{1}$$

and

$$RRf_t = \frac{1+Rf_t}{1+CPI_t} - 1 = \frac{Rf_t - CPI_t}{1+CPI_t}, \tag{2}$$

where R_t, the return of the U.S. stock market, represented by the S&P 500 Index, is

$$R_t = (1+CPI_t)(1+RRf_t)(1+ERP_t) - 1 \tag{3}$$

and Rf_t is the return of risk-free assets, represented by the income return of long-term U.S. government bonds.

The compound average for equity return was 10.70 percent for 1926–2000. For the equity risk premium, we can interpret that investors were compensated 5.24 pps a year for investing in common stocks rather than long-term risk-free assets (such as long-term U.S. government bonds). This calculation also shows that roughly half of the total historical equity return has come from the equity risk premium; the other half is from inflation and the long-term real risk-free rate. Average U.S. equity returns from 1926 through 2000 can be reconstructed as follows:[9]

$$\bar{R} = (1 + \overline{CPI})(1 + \overline{RRf})(1 + \overline{ERP}) - 1$$
$$10.70\% = (1 + 3.08\%) \times (1 + 2.05\%) \times (1 + 5.24\%) - 1.$$

The first column in **Figure 1** shows the decomposition of historical equity returns for 1926–2000 according to the building blocks method.

Method 2. Capital Gain and Income. The equity return, based on the form in which the return is distributed, can be broken into capital gain, cg, and income return, Inc. Income return of common stock is distributed to investors through dividends, whereas capital gain is distributed through price appreciation. Real capital gain, Rcg, can be computed by subtracting inflation from capital gain. The equity return in period t can then be decomposed as follows:

$$R_t = [(1 + CPI_t)(1 + Rcg_t) - 1] + Inc_t + Rinv_t, \quad (4)$$

where $Rinv$ is reinvestment return.

The average income return was calculated to be 4.28 percent in the study period, the average capital gain was 6.19 percent, and the average real capital gain was 3.02 percent. The reinvestment return averaged 0.20 percent from 1926 through 2000. For Method 2, the average U.S. equity return for 1926–2000 can thus be computed according to

$$\bar{R} = [(1 + \overline{CPI})(1 + \overline{Rcg}) - 1] + \overline{Inc} + \overline{Rinv}$$
$$10.70\% = [(1 + 3.08\%) \times (1 + 3.02\%) - 1] + 4.28\% + 0.20\%.$$

The second column in Figure 1 shows the decomposition of historical equity returns for 1926–2000 according to the capital gain and income method.

Method 3. Earnings. The real-capital-gain portion of the return in the capital gain and income method can be broken into growth in real EPS, g_{REPS}, and growth in P/E, $g_{P/E}$:

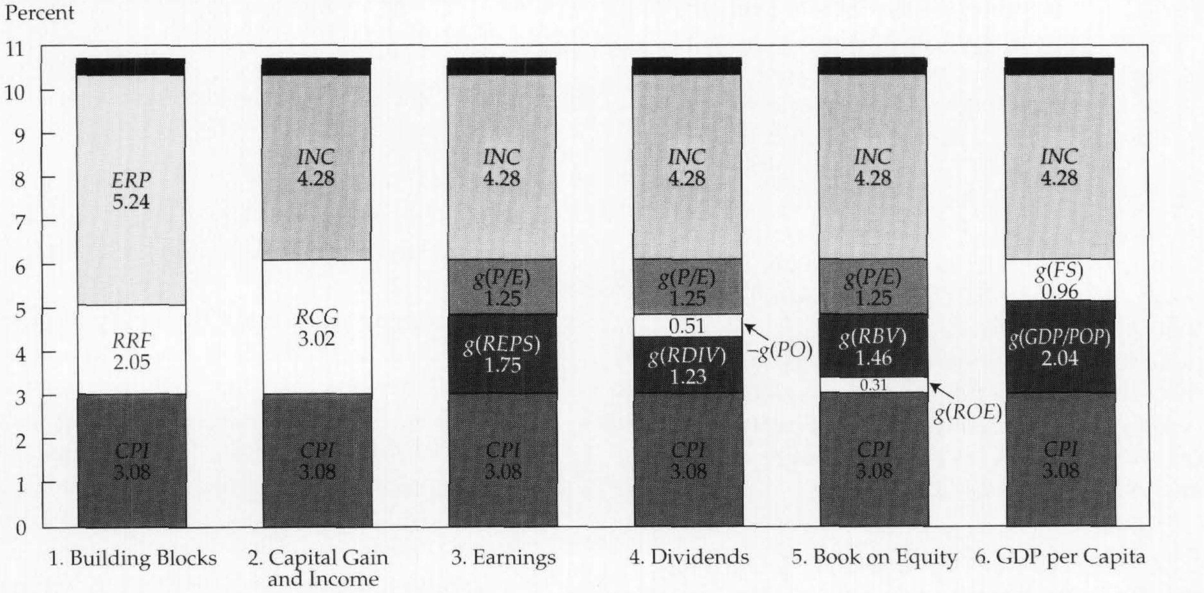
$$\begin{aligned} Rcg_t &= \frac{P_t}{P_{t-1}} - 1 \\ &= \frac{P_t/E_t}{P_{t-1}/E_{t-1}}\left(\frac{E_t}{E_{t-1}}\right) - 1 \\ &= (1 + g_{P/E,t})(1 + g_{REPS,t}) - 1. \end{aligned} \quad (5)$$

Therefore, equity's total return can be broken into four components—inflation, growth in real EPS, growth in P/E, and income return:

$$\begin{aligned} R_t = {} & [(1 + CPI_t)(1 + g_{REPS,t})(1 + g_{P/E,t}) - 1] \\ & + Inc_t + Rinv_t. \end{aligned} \quad (6)$$

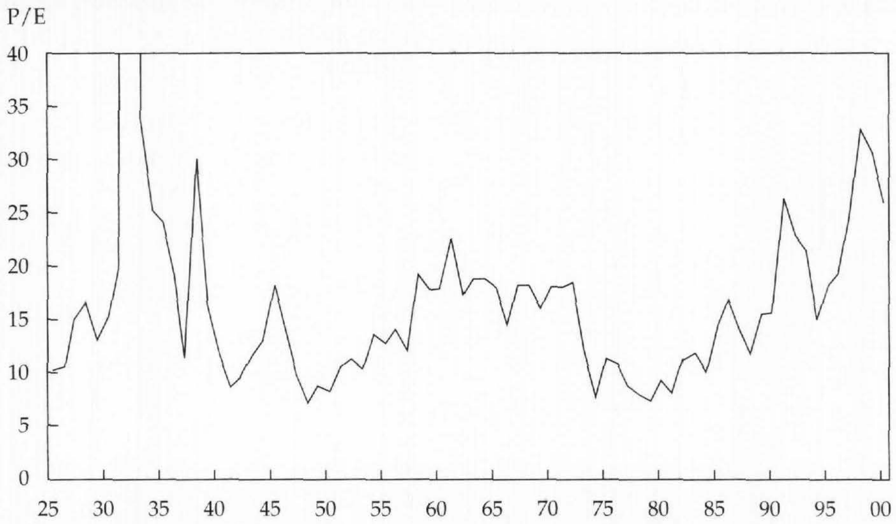
The real earnings of U.S. equity increased 1.75 percent annually between 1926 and 2000. The P/E, as **Figure 2** illustrates, was 10.22 at the beginning of 1926 and 25.96 at the end of 2000. The highest P/E (136.50 and off the chart in Figure 2) was recorded during the Great Depression, in December 1932, when earnings were near zero, and the lowest in the period (7.07) was recorded in 1948. The average year-end P/E was 13.76.[10]

Figure 1. Decomposition of Historical Equity Returns by Six Methods, 1926–2000



Notes: The block on the top of each column is the reinvestment return plus the geometric interactions among the components. Including the geometric interactions ensured that the components summed to 10.70 percent in this and subsequent figures. The table that constitutes Appendix A gives detailed information on the reinvestment and geometric interaction for all the methods.

Figure 2. P/E, 1926–2000



The U.S. equity returns from 1926 and 2000 can be computed according to the earnings method as follows:

$$\bar{R} = [(1+\overline{CPI})(1+\overline{g_{REPS}})(1+\overline{g_{P/E}})-1] + \overline{Inc} + \overline{Rinv}$$

$$10.70\% = [(1+3.08\%) \times (1+1.75\%) \times (1+1.25\%) - 1] + 4.28\% + 0.20\%.$$

The third column in Figure 1 shows the decomposition of historical equity returns for 1926–2000 according to the earnings method.

Method 4. Dividends. In this method, real dividends, *RDiv*, equal the real earnings times the dividend-payout ratio, *PO*, or

$$REPS_t = \frac{RDiv_t}{PO_t}; \tag{7}$$

therefore, the growth rate of earnings can be calculated by the difference between the growth rate of real dividends, g_{RDiv}, and the growth rate of the payout ratio, g_{PO}:

$$(1+g_{REPS,t}) = \frac{(1+g_{RDiv,t})}{(1+g_{PO,t})}. \tag{8}$$

If dividend growth and payout-ratio growth are substituted for the earnings growth in Equation 6, equity total return in period t can be broken into (1) inflation, (2) the growth rate of P/E, (3) the growth rate of the dollar amount of dividends after inflation, (4) the growth rate of the payout ratio, and (5) the dividend yield:

$$R_t = \left[(1+CPI_t)(1+g_{P/E,t})\left(\frac{1+g_{RDiv,t}}{1+g_{PO,t}}\right)-1\right] + Inc_t + Rinv_t. \tag{9}$$

Figure 3 shows the annual income return (dividend yield) of U.S. equity for 1926–2000. The dividend yield dropped from 5.15 percent at the beginning of 1926 to only 1.10 percent at the end of 2000. **Figure 4** shows the year-end dividend-payout ratio for 1926–2000. On average, the dollar amount of dividends after inflation grew 1.23 percent a year, while the dividend-payout ratio decreased 0.51 percent a year. The dividend-payout ratio was 46.68 percent at the beginning of 1926. It had decreased to 31.78 percent at the end of 2000. The highest dividend-payout ratio was recorded in 1932, and the lowest was the 31.78 percent recorded in 2000.

The U.S. equity returns from 1926 through 2000 can be computed in the dividends method according to

$$\bar{R} = \left[(1+\overline{CPI})(1+\overline{g_{P/E}})\left(\frac{1+\overline{g_{RDiv}}}{1+\overline{g_{PO}}}\right)-1\right] + \overline{Inc} + \overline{Rinv}$$

$$10.70\% = \left[(1+3.08\%) \times (1+1.25\%) \times \left(\frac{1+1.23\%}{1-0.51\%}\right)-1\right] + 4.28\% + 0.20\%.$$

The decomposition of equity return according to the dividends method is given in the fourth column of Figure 1.

Method 5. Return on Book Equity. Earnings can be broken into the book value of equity, *BV*, and return on the book value of equity, *ROE*:

$$EPS_t = BV_t(ROE_t). \tag{10}$$

The growth rate of earnings can be calculated from the combined growth rates of real book value, g_{RBV}, and of *ROE*:

$$1+g_{REPS,t} = (1+g_{RBV,t})(1+g_{ROE,t}). \tag{11}$$

Figure 3. Income Return (Dividend Yield), 1926–2000

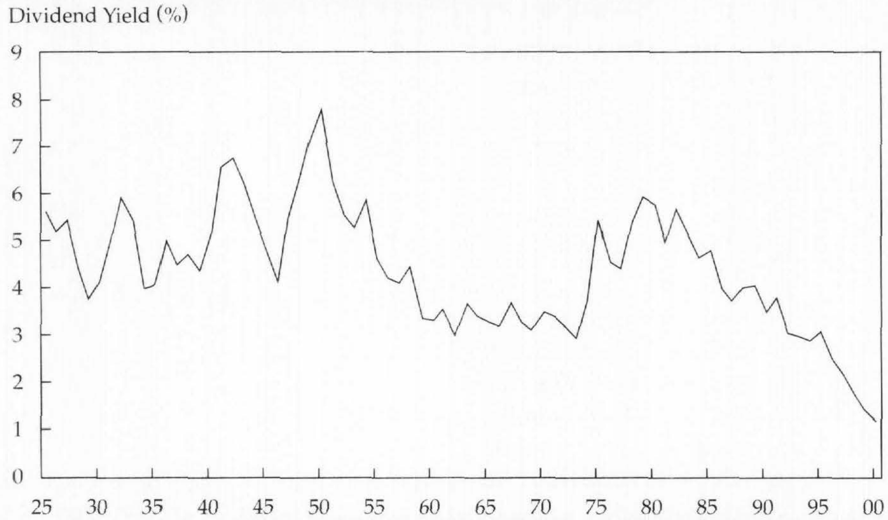
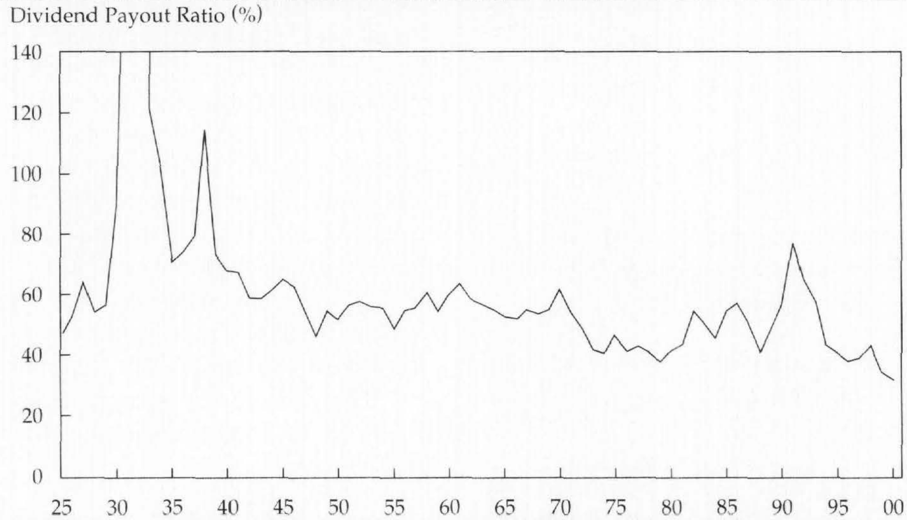


Figure 4. Dividend-Payout Ratio, Year-End 1926–2000



Note: The dividend-payout ratio was 190.52 percent in December 1931 and 929.12 percent in December 1932.

In this method, BV growth and ROE growth are substituted for earnings growth in the equity return decomposition, as shown in the fifth column of Figure 1. Then, equity's total return in period t can be computed by

$$R_t = [(1 + CPI_t)(1 + g_{P/E,t})(1 + g_{RBV,t})(1 + g_{ROE,t}) - 1] + Inc_t + Rinv_t. \quad (12)$$

We estimated that the average growth rate of the book value after inflation was 1.46 percent for 1926–2000.[11] The average ROE growth a year during the same time period was calculated to be 0.31 percent:

$$\bar{R} = [(1 + \overline{CPI})(1 + \overline{g_{P/E}})(1 + \overline{g_{BV}})(1 + \overline{g_{ROE}}) - 1] + \overline{Inc} + \overline{Rinv}$$

$$10.70\% = [(1 + 3.08\%)(1 + 1.25\%)(1 + 1.46\%)(1 + 0.31\%) - 1] + 4.28\% + 0.20\%.$$

Method 6. GDP per Capita. Diermeier et al. proposed a framework to analyze the aggregate supply of financial asset returns. Because we were interested only in the supply model of the equity returns in this study, we developed a slightly different supply model based on the growth of economic productivity. In this method, the market return over the long run is decomposed into (1) inflation, (2) the real growth rate of overall economic productivity (GDP per capita, $g_{GDP/POP}$), (3) the increase in the equity market relative to overall economic productivity (the increase in the factor share of equities in the overall economy, g_{FS}), and (4) dividend yields.[12] This model is expressed by the following equation:

$$R_t = [(1 + CPI_t)(1 + g_{GPD/POP,t})(1 + g_{FS,t}) - 1] + Inc_t + Rinv_t. \tag{13}$$

Figure 5 shows the growth of the U.S. stock market, GDP per capita, earnings, and dividends initialized to unity ($1.00) at the end of 1925. The level of all four factors dropped significantly in the early 1930s. For the whole period, GDP per capita slightly outgrew earnings and dividends, but all four factors grew at approximately the same rate. In other words, overall economic productivity increased slightly faster than corporate earnings or dividends over the past 75 years. Although GDP per capita outgrew earnings and dividends, the overall stock market price grew faster than GDP per capita. The primary reason is that the market P/E increased 2.54 times during the same time period.

Average equity market return can be calculated according to this model as follows:

$$\begin{aligned}\bar{R} &= [(1 + \overline{CPI})(1 + \overline{g_{GDP/POP}})(1 + \overline{g_{FS}}) - 1] + \overline{Inc} + \overline{Rinv} \\ 10.70\% &= [(1 + 3.08\%)(1 + 2.04\%)(1 + 0.96\%) - 1] + 4.28\% + 0.20\%.\end{aligned}$$

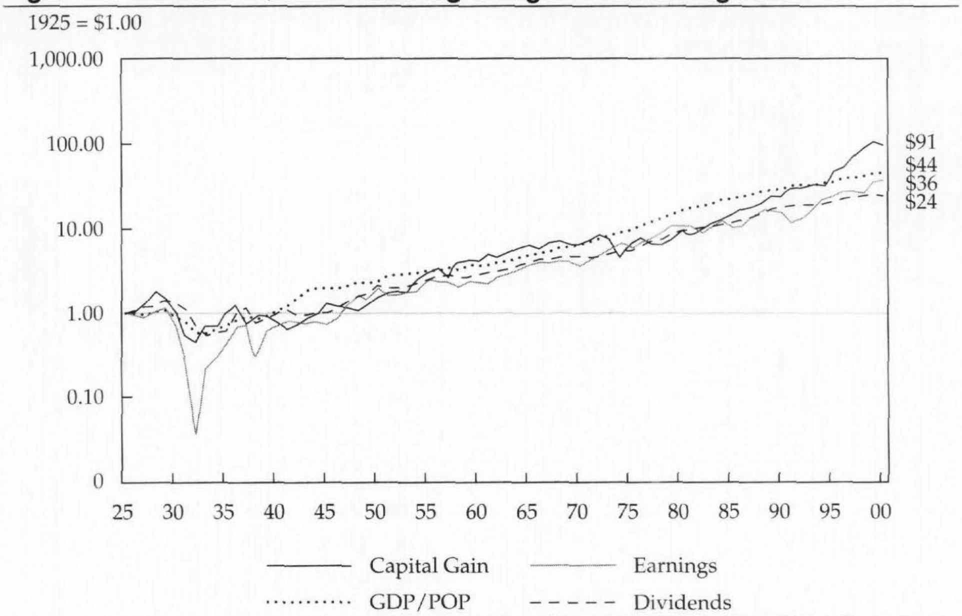
We calculated the average annual increase in the factor share of the equity market relative to the overall economy to be 0.96 percent. The increase in this factor share is less than the annual increase of the P/E (1.25 percent) over the same time period. This finding suggests that the increase in the equity market share relative to the overall economy can be fully attributed to the increase in its P/E.

The decomposition of historical equity returns by the GDP per capita model is given in the last column of Figure 1.

Summary of Equity Returns and Components. The decomposition of the six models into their components can be compared by looking at Figure 1. The differences among the five models arise from the different components that represent the capital gain portion of the equity returns.

This analysis produced several important findings. First, as Figure 5 shows, the growth in corporate earnings has been in line with the growth of overall economic productivity. Second, P/E increases accounted for only 1.25 pps of the 10.70 percent total equity return. Most of the return has been attributable to dividend payments and nominal earnings growth (including inflation and real earnings growth). Third, the increase in the relative factor share of equity can be fully attributed to the increase in P/E. Overall, economic productivity outgrew both corporate earnings and dividends from 1926 through 2000. Fourth, despite the record earnings growth in the 1990s, the dividend yield and the payout ratio declined sharply, which renders dividends alone a poor measure for corporate profitability and future earnings growth.

Figure 5. Growth of $1 from the Beginning of 1926 through 2000



Long-Term Forecast of Equity Returns

Supply-side models can be used to forecast the long-term expected equity return. The supply of stock market returns is generated by the productivity of the corporations in the real economy. Over the long run, the equity return should be close to the long-run supply estimate. In other words, investors should not expect a much higher or a much lower return than that produced by the companies in the real economy. Therefore, we believe investors' expectations for long-term equity performance should be based on the supply of equity returns produced by corporations.

The supply of equity returns consists of two main components—current returns in the form of dividends and long-term productivity growth in the form of capital gains. In this section, we focus on two of the supply-side models—the earnings model and the dividends model (Methods 3 and 4).[13] We studied the components of these two models by identifying which components are tied to the supply of equity returns and which components are not. Then, we estimated the long-term, sustainable return based on historical information about these supply components.

Model 3F. Forward-Looking Earnings. According to the earnings model (Equation 6), the historical equity return can be broken into four components—the income return, inflation, the growth in real EPS, and the growth in P/E. Only the first three of these components are historically supplied by companies. The growth in P/E reflects investors' changing predictions of future earnings growth. Although we forecasted that the past supply of corporate growth will continue, we did not forecast any change in investor predictions. Thus, the supply side of equity return, *SR*, includes only inflation, the growth in real EPS, and income return:[14]

$$SR_t = [(1 + CPI_t)(1 + g_{REPS,t}) - 1] + Inc_t + Rinv_t. \quad (14)$$

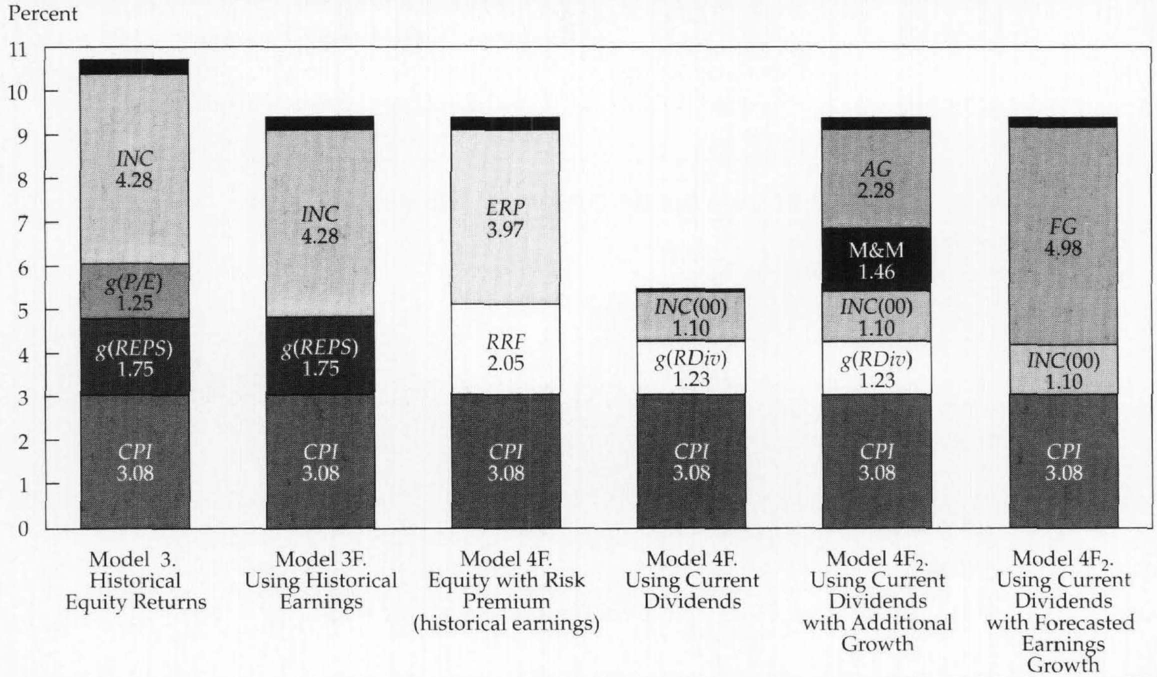
The long-term supply of U.S. equity returns based on the earnings model is 9.37 percent, calculated as follows:

$$\overline{SR} = [(1 + \overline{CPI})(1 + \overline{g_{REPS}}) - 1] + \overline{Inc} + \overline{Rinv}$$

$$9.37\% = [(1 + 3.08\%)(1 + 1.75\%) - 1] + 4.28\% + 0.20\%.$$

The decomposition according to Model 3F is compared with that of Method 3 (based on historical data plus the estimated equity risk premium) in the first two columns of **Figure 6**.

Figure 6. Historical vs. Current Dividend-Yield Forecasts Based on Earnings and Dividends Models



Notes: *Inc*(00) is the dividend yield in year 2000. *FG* is the real earnings growth rate, forecasted to be 4.98 percent. Model 4F$_2$ corrects Model 4F as follows: add 1.46 pps for M&M consistency and add 2.24 pps for the additional growth, *AG*, implied by the high current market P/E.

The supply-side equity risk premium, *ERP*, based on the earnings model is calculated to be 3.97 pps:

$$\begin{aligned}\overline{ERP} &= \frac{(1+\overline{SR})}{(1+\overline{CPI})(1+\overline{RRf})} - 1 \\ &= \frac{1+9.37\%}{(1+3.08\%)(1+2.05\%)} - 1 \\ &= 3.97\%.\end{aligned}$$

The *ERP* is taken into account in the third column of Figure 6.

Model 4F. Forward-Looking Dividends. The forward-looking dividends model is also referred to as the constant-dividend-growth model (or the Gordon model). In it, the expected equity return equals the dividend yield plus the expected dividend growth rate. The supply of the equity return in the Gordon model includes inflation, the growth in real dividends, and dividend yield.

As is commonly done with the constant-dividend-growth model, we used the current dividend yield of 1.10 percent instead of the historical dividend yield of 4.28 percent. This decision reduced the estimate of the supply of equity returns to 5.44 percent:

$$\begin{aligned}\overline{SR} &= [(1+\overline{CPI})(1+\overline{g_{RDiv}})-1] + Inc(00) + \overline{Rinv} \\ 5.54\% &= [(1+3.08\%)(1+1.23\%)-1] + 1.10\% + 0.20\%,\end{aligned}$$

where $Inc(00)$ is the dividend yield in year 2000. The equity risk premium was estimated to be 0.24 pps:

$$\begin{aligned}\overline{ERP} &= \frac{(1+\overline{SR})}{(1+\overline{CPI})(1+\overline{RRf})} - 1 \\ &= \frac{1+5.54\%}{(1+3.08\%)+(1+2.05\%)} - 1 \\ &= 0.24\%.\end{aligned}$$

Figure 6 allows a comparison of forecasted equity returns including the equity risk premium estimates based on the earnings model and the dividends model. In the next section, we show why we disagree with the dividends model and prefer to use the earnings model to estimate the supply-side equity risk premium.

Differences between the Earnings Model and the Dividends Model. The earnings model (3F) and the dividends model (4F) differ in essentially two ways. The differences relate to the low current payout ratio and the high current P/E. These two differences are reconciled in what we will call Model $4F_2$ shown in the two right-hand columns of Figure 6. First, to reflect growth in productivity, the earnings model uses historical earnings growth whereas the dividend model uses historical dividend growth. Historical dividend growth underestimates historical earnings growth, however, because of the decrease in the payout ratio. Overall, the dividend growth underestimated the increase in earnings productivity by 0.51 pps a year for 1926–2000. Today's low dividend yield also reflects the current payout ratio, which is at a historical low of 31.8 percent (compared with the historical average of 59.2 percent). Applying such a low rate to the future would mean that even more earnings would be retained in the future than in the historical period studied. But had more earnings been retained, the historical earnings growth would have been 0.95 pps a year higher, so (assuming the historical average dividend-payout ratio) the current yield of 1.10 percent would need to be adjusted upward by 0.95 pps.

By using the current dividend-payout ratio in the dividend model, Model 4F creates two errors, both of which violate Miller and Modigliani theory. A company's dividend-payout ratio affects only the form in which shareholders receive their returns (i.e., dividends versus capital gains), not their total returns. The *current* low dividend-payout ratio should not affect our forecast. Companies today probably have such low payout ratios to reduce the tax burden on their investors. Instead of paying dividends, many companies reinvest earnings, buy back shares, or use the cash to purchase other companies.[15] Therefore, the dividend growth model has to be upwardly adjusted by 1.46 pps (0.51 pp plus 0.95 pp) so as not to violate M&M theory.

The second difference between Model 3F and Model 4F is related to the fact that the current P/E (25.96) is much higher than the historical average (13.76). The current yield (1.10 percent) is at a historic low—because of the previously mentioned low payout ratio and because of the high P/E. Even assuming the historical average payout ratio, the current dividend yield would be much lower than its historical average (2.05 percent versus 4.28 percent). This difference is geometrically estimated to be 2.28 pps a year. In Figure 6, the additional growth, *AG*, accounts for 2.28 pps of the return; in the last column, the forecasted real earnings growth rate, *FG*, accounts for 4.98 pps. The high P/E could be caused by (1) mispricing, (2) a low required rate of return, and/or (3) a high expected future earnings growth rate. Mispricing as a cause is eliminated by our assumption of market efficiency, and a low required rate of return is eliminated by our assumption of a constant equity risk premium through the past and future periods that we are trying to estimate. Thus, we interpret the high P/E as the market expectation of higher earnings growth and the following equation is the model for

Model $4F_2$, which reconciles the differences between the earnings model and the dividends model:[16]

$$\overline{SR} = [(1+\overline{CPI})(1+\overline{g_{RDiv}})(1-\overline{g_{PO}})-1] + Inc(00) + AY + AG + \overline{Rinv}$$

$$9.67\% = [(1+3.08\%)(1+1.23\%)(1+0.51\%)-1] + 1.10\% + 0.95\% + 2.28\% + 0.20\%.$$

To summarize, the earnings model and the dividends model have three differences. The first two differences relate to the dividend-payout ratio and are direct violations of M&M. The third difference results from the expectation of higher-than-average earnings growth, which is predicted by the high current P/E. Reconciling these differences reconciles the earnings and dividends models.

Geometric vs. Arithmetic. The estimated equity return (9.37 percent) and equity risk premium (3.97 pps) are geometric averages. The arithmetic average, however, is often used in portfolio optimization. One way to convert the geometric average into an arithmetic average is to assume the returns are independently lognormally distributed over time. Then, the arithmetic average, R_A, and geometric average, R_G, have roughly the following relationship:

$$R_A = R_G + \frac{\sigma^2}{2}, \tag{15}$$

where σ^2 is the variance.

The standard deviation of equity returns is 19.67 percent. Because almost all the variation in equity returns is from the equity risk premium, rather than the risk-free rate, we need to add 1.93 pps to the geometric estimate of the equity risk premium to convert the returns into arithmetic form, so $R_A = R_G + 1.93$ pps. The arithmetic average equity risk premium then becomes 5.90 pps for the earnings model.

To summarize, the long-term supply of equity return is estimated to be 9.37 percent (6.09 percent after inflation), conditional on the historical average risk-free rate. The supply-side equity risk premium is estimated to be 3.97 pps geometrically and 5.90 pps arithmetically.[17]

Conclusions

We adopted a supply-side approach to estimate the forward-looking, long-term, sustainable equity return and equity risk premium. We analyzed historical equity returns by decomposing returns into factors commonly used to describe the aggregate equity market and overall economic productivity—inflation, earnings, dividends, P/E, the dividend-payout ratio, BV, ROE, and GDP per capita. We examined each factor and its relationship to the long-term supply-side framework. We used historical information in our supply-side models to forecast the equity risk premium. A complete tabulation of all the numbers from all models and methods is presented in Appendix A.

Contrary to several recent studies on the equity risk premium declaring the forward-looking premium to be close to zero or negative, we found

Appendix A. Summary Tabulations for Forecasted Equity Return

Method/Model	Sum	Inflation	Real Risk-Free Rate	Equity Risk Premium	Real Capital Gain	g(Real *EPS*)	g(Real *Div*)	$-g$(Payout Ratio)
A. Historical								
Method 1	10.70	3.08	2.05	5.24				
Method 2	10.70	3.08			3.02			
Method 3	10.70	3.08				1.75		
Method 4	10.70	3.08					1.23	0.51
Method 5	10.70	3.08						
Method 6	10.70	3.08						
B. Forecast with historical dividend yield								
Model 3F	9.37	3.08				1.75		
Model 3F (*ERP*)	9.37	3.08	2.05	3.97				
C. Forecast with current dividend yield								
Model 4F	5.44	3.08					1.23	
Model 4F (*ERP*)	5.44	3.08	2.05	0.24				
Model $4F_2$	9.37	3.08					1.23	0.51
Model $4F_2$ (*FG*)	9.37	3.08						

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

the long-term supply of the equity risk premium to be only slightly lower than the straight historical estimate. We estimated the equity risk premium to be 3.97 pps in geometric terms and 5.90 pps on an arithmetic basis. These estimates are about 1.25 pps lower than the historical estimates. The differences between our estimates and the ones provided by several other recent studies result principally from the inappropriate assumptions those authors used, which violate the M&M theorem. Also, our models interpret the current high P/E as the market forecasting high future growth rather than a low discount rate or an overvaluation. Our estimate is in line with both the historical supply measures of public corporations (i.e., earnings) and overall economic productivity (GDP per capita).

The implication of an estimated equity risk premium being far closer to the historical premium than zero or negative is that stocks are expected to outperform bonds over the long run. For long-term investors, such as pension funds and individuals saving for retirement, stocks should continue to be a favored asset class in a diversified portfolio. Because our estimate of the equity risk premium is lower than historical performance, however, some investors should lower their equity allocations and/or increase their savings rate to meet future liabilities.

Notes

1. In our study, we defined the equity risk premium as the difference between the long-run expected return on stocks and the long-term risk-free (U.S. Treasury) yield. [Some other studies, including Ibbotson and Sinquefield (1976a, 1976b) used short-term U.S. T-bills as the risk-free rate.] We did all of our analysis in geometric form, then converted to arithmetic data at the end, so the estimate is expressed in both arithmetic and geometric forms.
2. See also Mehra (2003).
3. Comparing estimates from one study with another is sometimes difficult because of changing points of reference. The equity risk premium estimate can be significantly different simply because the authors used arithmetic versus geometric returns, a long-term risk-free rate versus a short-term risk-free rate, bond income return (yield) versus bond total return, or long-term strategic forecasting versus short-term market-timing estimates. We provide a detailed discussion of arithmetic versus geometric returns in the section "The Long-Term Forecast."
4. Welch's survey reported a 7 pp equity risk premium measured as the arithmetic difference between equity and T-bill returns. To make an apples-to-apples comparison, we converted the 7 pp number into a geometric equity risk premium relative to the long-term U.S. government bond income return, which produced an estimate of almost 4 pps.
5. For further discussion of approaches to estimating the equity risk premium, see the presentations and discussions at www.aimrpubs.org/ap/home.html from AIMR's *Equity Risk Premium Forum*.
6. Each per share quantity is per share of the S&P 500 portfolio. Hereafter, we will merely refer to each factor without always mentioning "per share"—for example, "dividends" instead of "dividends per share."
7. Many theoretical models suggest that the equity risk premium is dynamic over time. Recent empirical studies (e.g., Goyal and Welch 2001; Ang and Bekaert 2001) found no evidence, however, of long-horizon return predictability by using either earnings or dividend yields. Therefore, instead

g(*BV*)	*g*(*ROE*)	*g*(*P/E*)	*g*(Real *GDP/POP*)	*g*(*FS-GDP/POP*)	Income Return	Reinvestment + Interaction	Additional Growth	Forecasted Earnings Growth
						0.33		
					4.28	0.32		
		1.25			4.28	0.34		
		1.25			4.28	0.35		
1.25	0.31	1.25			4.28	0.31		
			2.04	0.96	4.28	0.32		
					4.28	0.26		
						0.27		
					1.10[a]	0.03		
						0.07		
					2.05[b]	0.21	2.28	
					1.10[a]	0.21		4.98

of trying to build a model for a dynamic equity risk premium, we assumed that the long-term equity risk premium is constant. This assumption provided a benchmark for analysis and discussion.

8. We updated the series with data from Standard and Poor's to include the year 2000.
9. Appendix A summarizes all the tabulations we discuss.
10. The average P/E was calculated by reversing the average earnings-to-price ratio for 1926–2000.
11. Book values were calculated from the book-to-market ratios reported in Vuolenteenaho (2000). The aggregate book-to-market ratio was 2.0 in 1928 and 4.1 in 1999. We used the growth rate in book value calculated for 1928–1999 as the proxy for the growth rate for 1926–2000. The average ROE growth rate was calculated from the derived book value and the earnings data.
12. Instead of assuming a constant equity factor share, we examined the historical growth rate of the equity factor share relative to the overall growth of the economy.
13. We did not use Methods 1, 2, and 5 in forecasting because the forecasts of Methods 1 and 2 would be identical to the historical estimate reported in the previous section and because the forecast of Method 5 would require more complete BV and ROE data than we currently have available. We did use Method 6 to forecast future stock returns but found the results to be very similar to those for the earnings model; therefore, we do not report the results here.
14. This model uses historical income return as an input for reasons that are discussed in the section "Differences between the Earnings Model and the Dividends Model."
15. The current tax code provides incentives for companies to distribute cash through share repurchases rather than through dividends. Green and Hollifield (2001) found that the tax savings through repurchases are on the order of 40–50 percent of the taxes that investors would have paid if dividends were distributed.
16. Contrary to efficient market models, Shiller (2000) and Campbell and Shiller argued that the P/E appears to forecast future stock price change.
17. We could also use the GDP per capita model to estimate the long-term equity risk premium. This model implies long-run stock returns should be in line with the productivity of the overall economy. The equity risk premium estimated by using the GDP per capita model would be slightly higher than the *ERP* estimate from the earnings model because GDP per capita grew slightly faster than corporate earnings in the study period. A similar approach can be found in Diermeier et al., who proposed using the growth rate of the overall economy as a proxy for the growth rate in aggregate wealth in the long run.

References

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. 2003. "Surprise! Higher Dividends = Higher Earnings Growth." *Financial Analysts Journal*, vol. 59, no. 1 (January/February):70–87.

Arnott, Robert D., and Peter L. Bernstein. 2002. "What Risk Premium Is 'Normal'?" *Financial Analysts Journal*, vol. 58, no. 2 (March/April):64–84.

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.

Campbell, John Y., and Robert J. Shiller. 2001. "Valuation Ratios and the Long-Run Stock Market Outlook: An Update." *NBER* Working Paper No. 8221.

Diermeier, Jeffrey J., Roger G. Ibbotson, and Laurance B. Siegel. 1984. "The Supply for Capital Market Returns." *Financial Analysts Journal*, vol. 40, no. 2 (March/April):2–8.

Fama, Eugene F., and Kenneth R. French. 2001. "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay?" *Journal of Financial Economics*, vol. 60, no. 1 (April):3–43.

———. 2002. "The Equity Risk Premium." *Journal of Finance*, vol. 57, no. 2 (April):637–659.

Gordon, Myron. 1962. *Investment, Financing, and Valuation of the Corporation*. Homewood, IL: Irwin.

Goyal, Amit, and Ivo Welch. 2001. "Predicting the Equity Premium with Dividend Ratios." Working paper. Yale School of Management and UCLA.

Graham, John R., and Campbell R. Harvey. 2001. "Expectations of Equity Risk Premia, Volatility and Asymmetry from a Corporate Finance Perspective." Working paper, Fuqua School of Business, Duke University (August).

Green, Richard C., and Burton Hollifield. 2001. "The Personal-Tax Advantages of Equity." Working paper, Carnegie Mellon University (January).

Ibbotson Associates. 2001. *Stocks, Bonds, Bills, and Inflation: 2001 Yearbook*. Chicago, IL: Ibbotson Associates.

Ibbotson, Roger G., and Rex A. Sinquefield. 1976a. "Stocks, Bonds, Bills, and Inflation: Year-by-Year Historical Returns (1926–1974)." *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, Roger G., Jeffrey J. Diermeier, and Laurance B. Siegel. 1984. "The Demand for Capital Market Returns: A New Equilibrium Theory." *Financial Analysts Journal*, vol. 40, no. 1 (January/February):22–33.

Mehra, Rajnish. 2003. "The Equity Premium: Why Is It a Puzzle?" *Financial Analysts Journal*, vol. 59, no. 1 (January/February):54–69.

Mehra, Rajnish, and Edward Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, vol. 15, no. 2 (March):145–161.

Miller, Merton, and Franco Modigliani. 1961. "Dividend Policy, Growth and the Valuation of Shares." *Journal of Business*, vol. 34, no. 4 (October):411–433.

Shiller, Robert J. 2000. *Irrational Exuberance*. Princeton, NJ: Princeton University Press.

Siegel, Jeremy J. 1999. "The Shrinking Equity Risk Premium." *Journal of Portfolio Management*, vol. 26, no. 1 (Fall):10–17.

Vuolenteenaho, Tuomo. 2000. "Understanding the Aggregate Book-to-Market Ratio and Its Implications to Current Equity-Premium Expectations." Working paper, Harvard University.

Welch, Ivo. 2000. "Views of Financial Economists on the Equity Premium and Other Issues." *Journal of Business*, vol. 73, no. 4 (October):501–537.

Wilson, Jack W., and Charles P. Jones. 2002. "An Analysis of the S&P 500 Index and Cowles' Extensions: Price Indexes and Stock Returns, 1870–1999." *Journal of Business*, vol. 75, no. 3 (July):505–535.

Stocks, Bonds, Bills, and Inflation®

2015 Ibbotson® SBBI® Market Report

Data as of December 2014

MORNINGSTAR®

Published by:

Morningstar, Inc.
22 W. Washington Street
Chicago, Illinois 60602

Main +1 312 696-6000
Product Sales +1 888 298-3647
Fax +1 312 696-6010
global.morningstar.com/DataPublications

Table of Contents

Results for 2014 Capital Markets

Large-Cap Stocks
The market for U.S. large-capitalization stocks is represented here by the total return on the S&P 500 Index (the total return includes reinvestment of dividends). Large-cap stocks for the year posted a total return of 13.69%, down from 32.39% in 2013. Eight months of 2014 produced positive returns; February delivered the highest return at 4.57%, while January's -3.46% was the lowest.

An index of large-cap stock total returns, started at $1.00 on Dec. 31, 1925, increased to $5,316.85 by the end of 2014. That was up from $4,676.88 a year earlier.

Small-Cap Stocks
Small-cap stocks delivered a total return of 2.92% in 2014, down from 45.07% the prior year. Seven months of 2014 produced positive returns; October posted the highest return at 6.52%, while September and July saw losses of 5.69% and 5.84%, respectively.

The cumulative wealth index grew to $27,419.32 from $1.00 at the end of 1925 and $26,641.17 at the end of 2013.

Long-Term Corporate Bonds
Long-term corporate bonds (with maturity near 20 years) returned 17.28% in 2014, well ahead of the 7.07% loss the previous year. Total returns were positive in 11 months of 2014, with August having the highest return of 3.56%, and September, at -2.71%, the lowest.

The bond default premium, or net return from investing in long-term corporate bonds rather than long-term government bonds of equal maturity, was negative at -5.32% in 2014, compared with 4.84% in 2013.

One dollar invested in long-term corporate bonds at year-end 1925 grew to $189.76 at the end of 2014, up from $161.80 a year earlier.

Long-Term Government Bonds
Long-term government bonds (with maturity near 20 years) returned 23.87% in 2014. This return was significantly higher than the -11.36% return in 2013 and more than four times the long-term average return (1926–2014) of 5.7%. Ten months produced positive returns, with January's the highest at 4.99%, and the -1.72% in September the lowest.

A wealth index of long-term government bonds grew to $135.18 at year-end 2014 from $1.00 at year-end 1925. The capital appreciation index of long-term government bond returns closed at $1.44 at year's end, up from $1.19 in 2013. December's close hit an all-time high, finally eclipsing the previous high set in February 1946.

Results for 2014 Capital Markets

Intermediate-Term Government Bonds
The total return on intermediate-term government bonds (with maturity near five years) in 2014 was 3.12%, above the -1.07% in 2013, but below the long-term (1926–2014) average return of 5.3%. Five months had positive returns, with October posting the highest return of 2.26% while June had the lowest return at -1.03%.

The wealth index of intermediate-term government bonds grew to $95.88 as of year-end 2014 after starting at $1.00 at year-end 1925. The index dipped in 2013 to $92.98.

Treasury Bills
An investment in bills with approximately 30 days to maturity returned 0.02% in 2014, repeating the return of 2013 and trailing the long-term average (1926–2014) of 3.5%. The cumulative index of Treasury bill total returns ended the year at $20.58, unchanged from a year earlier. Because monthly Treasury bill returns are nearly always positive, each monthly index value typically sets a new all-time high.

Inflation
Inflation decreased to 0.76% in 2014, compared to 1.50% in 2013. The result is lower than the long-term historical average (1926–2014) of 2.9%. Inflation has remained below 5% for 32 of the last 33 years (the exception was the 6.11% rate in 1990).

A cumulative inflation index, beginning at $1.00 at year-end 1925, finished 2014 at $13.10, up from $13.00 at year-end 2013. That is, a “basket” of consumer goods and services that cost $1.00 in 1925 would cost $13.10 today. The two baskets are not identical, but are intended to be comparable.

Graph 1
Wealth Indexes of Investments in the U.S. Capital Markets
Index (Dec. 31, 1925 = $1.00)

From December 1925 to December 2014

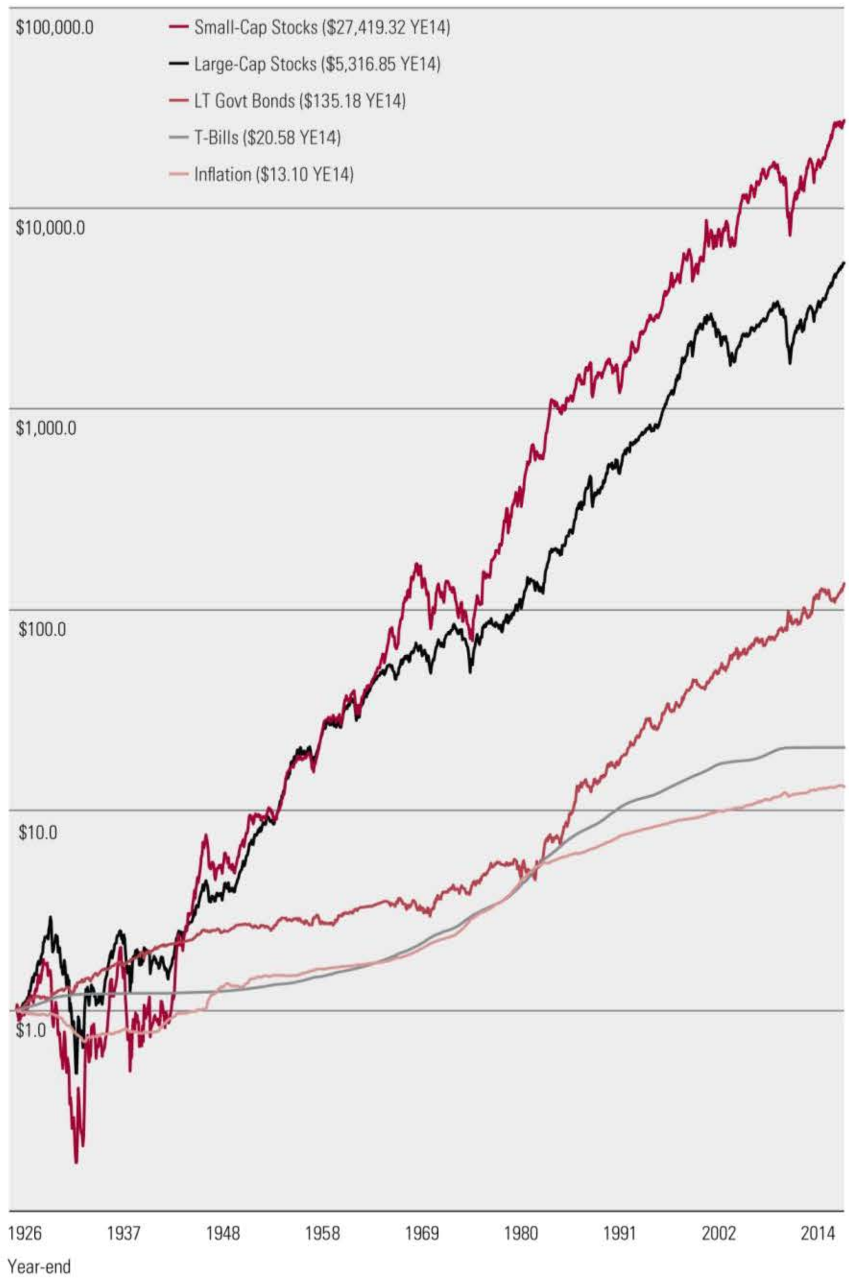


Table 1
Basic Series: Annual Total Returns in Percent

Year	Large-Cap Stocks	Small-Cap Stocks	Long-Term Corporate Bonds	Long-Term Government Bonds	Intermediate-Term Government Bonds	U.S. Treasury Bills	Inflation
2005	4.91	5.69	5.87	7.81	1.36	2.98	3.42
2006	15.79	16.17	3.24	1.19	3.14	4.80	2.54
2007	5.49	-5.22	2.60	9.88	10.05	4.66	4.08
2008	-37.00	-36.72	8.78	25.87	13.11	1.60	0.09
2009	26.46	28.09	3.02	-14.90	-2.40	0.10	2.72
2010	15.06	31.26	12.44	10.14	7.12	0.12	1.50
2011	2.11	-3.26	17.95	28.23	9.46	0.04	2.96
2012	16.00	18.24	10.68	3.31	2.07	0.06	1.74
2013	32.39	45.07	-7.07	-11.36	-1.07	0.02	1.50
2014	13.69	2.92	17.28	23.87	3.12	0.02	0.76

Table 2
Portfolios: Annual Total Returns in Percent

Year	100% Large-Cap Stocks	90% Stocks 10% Bonds	70% Stocks 30% Bonds	50% Stocks 50% Bonds	30% Stocks 70% Bonds	10% Stocks 90% Bonds	100% Long-Term Govt. Bonds
2004	10.88	10.70	10.29	9.84	9.34	8.80	8.51
2005	4.91	5.28	5.96	6.58	7.12	7.60	7.81
2006	15.79	14.30	11.33	8.40	5.49	2.61	1.19
2007	5.49	6.03	7.03	7.95	8.79	9.54	9.88
2008	-37.00	-32.14	-21.55	-9.72	3.43	18.02	25.87
2009	26.46	21.86	12.97	4.49	-3.58	-11.23	-14.90
2010	15.06	14.97	14.52	13.70	12.53	11.02	10.14
2011	2.11	4.77	10.07	15.34	20.56	25.70	28.23
2012	16.00	14.83	12.42	9.91	7.32	4.66	3.31
2013	32.39	27.33	17.70	8.69	0.26	-7.62	-11.36
2014	13.69	14.73	16.80	18.84	20.87	22.87	23.87

Table 3
Basic Series: Monthly and Quarterly Returns in Percent

Month	Large-Cap Stocks Total Return	Large-Cap Stocks Income Return	Large-Cap Stocks Cap Appr	Small-Cap Stocks Total Return	Long-Term Corporate Bonds Total Return	Long-Term Government Bonds Total Return	Long-Term Government Bonds Income Return	Long-Term Government Bonds Cap Appr	Long-Term Government Bonds Yield	Intermediate-Term Government Bonds Total Return	Intermediate-Term Government Bonds Income Return	Intermediate-Term Government Bonds Cap Appr	Intermediate-Term Government Bonds Yield	Treasury Bills Total Return	Inflation Rate
12/13	2.53	0.18	2.36	1.89	0.02	-1.64	0.31	-2.81	3.67	-0.19	0.09	-0.29	1.13	0.00	-0.01
1/14	-3.46	0.10	-3.56	-4.43	3.31	4.99	0.31	4.68	3.35	1.56	0.14	1.42	1.33	0.00	0.37
2/14	4.57	0.26	4.31	4.21	1.68	-0.10	0.25	-0.35	3.38	0.26	0.10	0.16	1.29	0.00	0.37
3/14	0.84	0.15	0.69	0.97	0.62	1.26	0.29	0.98	3.31	-0.69	0.11	-0.80	1.46	0.00	0.64
4/14	0.74	0.12	0.62	-3.41	1.60	1.38	0.27	1.11	3.24	0.97	0.12	0.85	1.28	0.00	0.33
5/14	2.35	0.24	2.10	0.10	1.88	2.71	0.28	2.43	3.07	0.46	0.11	0.35	1.20	0.00	0.35
6/14	2.07	0.16	1.91	4.34	0.20	0.28	0.25	0.03	3.07	-1.03	0.10	-1.13	1.45	0.00	0.19
7/14	-1.38	0.13	-1.51	-5.84	0.24	1.10	0.26	0.84	3.01	-0.02	0.12	-0.14	1.48	0.00	-0.04
8/14	4.00	0.23	3.77	4.54	3.56	3.11	0.25	2.86	2.81	-0.13	0.12	-0.25	1.54	0.00	-0.17
9/14	-1.40	0.15	-1.55	-5.69	-2.71	-1.72	0.23	-1.94	2.95	-0.09	0.13	-0.22	1.59	0.00	0.08
10/14	2.44	0.12	2.32	6.52	2.25	3.54	0.25	3.30	2.73	2.26	0.13	2.12	1.09	0.00	-0.25
11/14	2.69	0.24	2.45	-0.75	1.73	1.64	0.22	1.42	2.63	-0.01	0.09	-0.10	1.12	0.00	-0.54
12/14	-0.25	0.17	-0.42	3.37	1.83	3.59	0.22	3.37	2.40	-0.43	0.10	-0.52	1.24	0.00	-0.57
2014	13.69	2.16	11.39	2.92	17.28	23.87	3.33	20.17	2.40	3.12	1.39	1.72	1.24	0.02	0.76

Quarter	Large-Cap Stocks Total Return	Large-Cap Stocks Income Return	Large-Cap Stocks Cap Appr	Small-Cap Stocks Total Return	Long-Term Corporate Bonds Total Return	Long-Term Government Bonds Total Return	Long-Term Government Bonds Income Return	Long-Term Government Bonds Cap Appr	Long-Term Government Bonds Yield	Intermediate-Term Government Bonds Total Return	Intermediate-Term Government Bonds Income Return	Intermediate-Term Government Bonds Cap Appr	Intermediate-Term Government Bonds Yield	Treasury Bills Total Return	Inflation Rate
I-12	12.59	0.57	12.00	12.25	-0.59	-4.75	0.62	-5.36	2.90	-0.75	0.20	-0.94	0.97	0.01	1.65
II-12	-2.75	0.53	-3.29	-3.60	6.72	9.52	0.67	8.84	2.25	2.74	0.19	2.55	0.44	0.01	0.04
III-12	6.35	0.58	5.76	6.46	3.81	0.46	0.54	-0.08	2.26	-0.23	0.11	-0.34	0.51	0.02	0.84
IV-12	-0.38	0.62	-1.01	2.63	0.49	-1.42	0.58	-1.99	2.41	0.32	0.09	0.24	0.46	0.02	-0.78
I-13	10.61	0.56	10.03	12.05	-2.40	-2.28	0.64	-2.92	2.84	-0.17	0.15	-0.33	0.77	0.01	1.38
II-13	2.91	0.55	2.36	4.83	-5.69	-5.82	0.71	-6.50	3.29	-0.91	0.17	-1.08	1.01	0.01	0.31
III-13	5.24	0.55	4.69	10.98	-0.29	-0.97	0.85	-1.83	3.42	0.03	0.26	-0.23	1.06	0.00	0.28
IV-13	10.52	0.58	9.92	11.29	1.25	-2.74	0.85	-4.43	3.67	-0.02	0.26	-0.29	1.13	0.01	-0.47
I-14	1.81	0.50	1.30	0.56	5.71	6.21	0.87	5.33	3.31	1.12	0.35	0.77	1.46	0.01	1.39
II-14	5.23	0.53	4.69	0.88	3.71	4.42	0.81	3.60	3.07	0.40	0.33	0.07	1.45	0.00	0.87
III-14	1.13	0.51	0.62	-7.17	1.01	2.45	0.75	1.70	2.95	-0.24	0.37	-0.61	1.59	0.00	-0.13
IV-14	4.93	0.54	4.39	9.28	5.91	9.02	0.71	8.29	2.40	1.81	0.32	1.49	1.24	0.00	-1.35

Table 4
Portfolios: Monthly and Quarterly Returns in Percent

Month	100% Large-Cap Stocks	90% Stocks 10% Bonds	70% Stocks 30% Bonds	50% Stocks 50% Bonds	30% Stocks 70% Bonds	10% Stocks 90% Bonds	100% Long-Term Govt. Bonds
12/13	2.53	2.11	1.28	0.45	-0.39	-1.22	-1.64
1/14	-3.46	-2.61	-0.92	0.77	2.46	4.15	4.99
2/14	4.57	4.11	3.17	2.24	1.30	0.37	-0.10
3/14	0.84	0.88	0.97	1.05	1.14	1.22	1.26
4/14	0.74	0.80	0.93	1.06	1.19	1.32	1.38
5/14	2.35	2.38	2.46	2.53	2.60	2.67	2.71
6/14	2.07	1.89	1.53	1.17	0.82	0.46	0.28
7/14	-1.38	-1.13	-0.64	-0.14	0.35	0.85	1.10
8/14	4.00	3.91	3.73	3.56	3.38	3.20	3.11
9/14	-1.40	-1.43	-1.50	-1.56	-1.62	-1.68	-1.72
10/14	2.44	2.55	2.77	2.99	3.21	3.43	3.54
11/14	2.69	2.58	2.37	2.16	1.95	1.74	1.64
12/14	-0.25	0.13	0.90	1.67	2.44	3.20	3.59
2014	13.69	14.73	16.80	18.84	20.87	22.87	23.87

Quarter							
I-12	12.59	10.77	7.19	3.68	0.25	-3.10	-4.75
II-12	-2.75	-1.51	0.97	3.44	5.88	8.31	9.52
III-12	6.35	5.76	4.58	3.40	2.22	1.04	0.46
IV-12	-0.38	-0.47	-0.67	-0.88	-1.09	-1.31	-1.42
I-13	10.61	9.28	6.66	4.07	1.51	-1.02	-2.28
II-13	2.91	2.04	0.30	-1.45	-3.20	-4.95	-5.82
III-13	5.24	4.62	3.38	2.14	0.89	-0.35	-0.97
IV-13	10.51	9.14	6.42	3.75	1.12	-1.46	-2.74
I-14	1.81	2.28	3.21	4.10	4.97	5.80	6.21
II-14	5.23	5.15	4.99	4.83	4.67	4.50	4.42
III-14	1.13	1.26	1.53	1.80	2.06	2.32	2.45
IV-14	4.93	5.34	6.16	6.98	7.79	8.61	9.02

Table 5
Basic Series: Monthly Index Values
Dec. 31, 1925 = $1.00

Month	Large-Cap Stocks Total Return	Large-Cap Stocks Capital Appreciation	Small-Cap Stocks Total Return	Long-Term Corporate Bonds Total Return	Long-Term Government Bonds Total Return	Long-Term Government Bonds Capital Appreciation	Intermediate-Term Government Bonds Total Return	Intermediate-Term Government Bonds Capital Appreciation	Treasury Bills Total Return	Inflation
12/13	4,676.682	144.853	26,641.173	161.802	109.138	1.195	92.977	1.714	20.579	13.001
1/14	4,514.990	139.699	25,460.969	167.162	114.584	1.251	94.424	1.738	20.580	13.049
2/14	4,721.522	145.722	26,532.876	169.977	114.470	1.247	94.674	1.741	20.581	13.097
3/14	4,761.210	146.732	26,790.245	171.035	115.916	1.259	94.023	1.727	20.581	13.182
4/14	4,796.405	147.642	25,876.698	173.771	117.515	1.273	94.939	1.742	20.581	13.225
5/14	4,908.996	150.747	25,902.574	177.034	120.695	1.304	95.376	1.748	20.582	13.271
6/14	5,010.404	153.620	27,026.746	177.389	121.036	1.305	94.395	1.729	20.582	13.296
7/14	4,941.306	151.303	25,448.384	177.820	122.363	1.315	94.381	1.726	20.582	13.291
8/14	5,138.983	157.001	26,603.741	184.154	126.169	1.353	94.254	1.722	20.582	13.269
9/14	5,066.916	154.565	25,089.988	179.172	124.005	1.327	94.172	1.718	20.582	13.279
10/14	5,190.676	158.151	26,725.855	183.199	128.398	1.370	96.297	1.755	20.583	13.245
11/14	5,330.277	162.031	26,525.411	186.359	130.501	1.390	96.286	1.753	20.582	13.174
12/14	5,316.850	161.353	27,419.317	189.762	135.185	1.437	95.875	1.744	20.583	13.099

Table 6
Portfolios: Monthly Index Values
Dec. 31, 1925 = $1.00

Month	100% Large-Cap Stocks	90% Stocks 10% Bonds	70% Stocks 30% Bonds	50% Stocks 50% Bonds	30% Stocks 70% Bonds	10% Stocks 90% Bonds	100% Long-Term Govt. Bonds
12/13	4,676.682	3,765.535	2,192.132	1,107.822	486.743	186.031	109.138
1/14	4,514.990	3,667.155	2,171.897	1,116.313	498.698	193.743	114.584
2/14	4,721.522	3,817.763	2,240.791	1,141.287	505.192	194.455	114.470
3/14	4,761.210	3,851.469	2,262.470	1,153.294	510.935	196.830	115.916
4/14	4,796.405	3,882.404	2,283.538	1,165.510	517.001	199.419	117.515
5/14	4,908.996	3,974.933	2,339.601	1,194.961	530.435	204.744	120.695
6/14	5,010.404	4,049.956	2,375.413	1,208.989	534.771	205.687	121.036
7/14	4,941.306	4,004.130	2,360.296	1,207.282	536.663	207.433	122.363
8/14	5,138.983	4,160.752	2,448.419	1,250.208	554.790	214.070	126.169
9/14	5,066.916	4,101.102	2,411.785	1,230.720	545.794	210.466	124.005
10/14	5,190.676	4,205.782	2,478.650	1,267.547	563.327	217.689	128.398
11/14	5,330.277	4,314.474	2,537.497	1,294.976	574.333	221.485	130.501
12/14	5,316.850	4,320.176	2,560.341	1,316.581	588.326	228.582	135.185

Table 7
Basic Series and Portfolios: Summary Statistics of Annual Total Returns in Percent

	From 1926 to 2014		
Asset Class	Geometric Mean	Arithmetic Mean	Standard Deviation
Large-Cap Stocks	10.1	12.1	20.1
Small-Cap Stocks	12.2	16.7	32.1
Long-Term Corporate Bonds	6.1	6.4	8.4
Long-Term Government Bonds	5.7	6.1	10.0
Intermediate-Term Government Bonds	5.3	5.4	5.6
U.S. Treasury Bills	3.5	3.5	3.1
Inflation	2.9	3.0	4.1
90% Stocks/10% Bonds	9.9	11.4	18.1
70% Stocks/30% Bonds	9.2	10.2	14.3
50% Stocks/50% Bonds	8.4	9.0	11.2
30% Stocks/70% Bonds	7.4	7.8	9.3
10% Stocks/90% Bonds	6.3	6.7	9.2

Table 8
Derived Series: Monthly and Quarterly Returns in Percent

Month	Equity Risk Premium*	Small-Cap Premium	Bond Default Premium	Bond Horizon Premium	Inflation Adjusted Total Returns (%) Large-Cap Stocks	Small-Cap Stocks	LT-Corp Bonds	LT-Govt Bonds	IT-Govt Bonds	T-Bill
12/13	2.53	-0.63	1.69	-1.64	2.54	1.90	0.03	-1.63	-0.18	0.01
1/14	-3.46	-1.01	-1.60	4.99	-3.82	-4.78	2.93	4.60	1.18	-0.37
2/14	4.57	-0.35	1.79	-0.10	4.19	3.83	1.31	-0.47	-0.10	-0.36
3/14	0.84	0.13	-0.63	1.26	0.20	0.32	-0.02	0.62	-1.32	-0.64
4/14	0.74	-4.12	0.22	1.38	0.41	-3.73	1.27	1.05	0.64	-0.33
5/14	2.35	-2.20	-0.81	2.71	1.99	-0.25	1.52	2.35	0.11	-0.35
6/14	2.06	2.23	-0.08	0.28	1.88	4.15	0.01	0.10	-1.21	-0.18
7/14	-1.38	-4.52	-0.84	1.10	-1.34	-5.80	0.28	1.14	0.02	0.04
8/14	4.00	0.52	0.44	3.11	4.17	4.71	3.74	3.28	0.03	0.17
9/14	-1.40	-4.35	-1.01	-1.72	-1.48	-5.76	-2.78	-1.79	-0.16	-0.07
10/14	2.44	3.98	-1.25	3.54	2.70	6.79	2.51	3.80	2.51	0.25
11/14	2.69	-3.35	0.09	1.64	3.25	-0.21	2.28	2.19	0.53	0.54
12/14	-0.25	3.63	-1.70	3.59	0.32	3.96	2.41	4.18	0.14	0.57
2014	13.67	-9.47	-5.32	23.85	12.83	2.15	16.40	22.94	2.34	-0.73

Quarter	Equity Risk Premium*	Small-Cap Premium	Bond Default Premium	Bond Horizon Premium	Large-Cap Stocks	Small-Cap Stocks	LT-Corp Bonds	LT-Govt Bonds	IT-Govt Bonds	T-Bill
I-12	12.58	-0.30	4.37	-4.75	10.76	10.43	-2.20	-6.29	-2.36	-1.61
II-12	-2.76	-0.87	-2.56	9.51	-2.79	-3.63	6.68	9.48	2.71	-0.03
III-12	6.34	0.10	3.34	0.44	5.47	5.58	2.94	-0.38	-1.06	-0.82
IV-12	-0.40	3.02	1.94	-1.44	0.41	3.44	1.28	-0.64	1.11	0.81
I-13	10.60	1.30	-0.13	-2.28	9.10	10.52	-3.73	-3.61	-1.53	-1.36
II-13	2.90	1.86	0.14	-5.83	2.59	4.50	-5.98	-6.12	-1.22	-0.31
III-13	5.25	5.45	0.69	-0.98	4.96	10.68	-0.56	-1.25	-0.24	-0.27
IV-13	10.51	0.70	4.11	-2.75	11.04	11.81	1.73	-2.28	0.45	0.48
I-14	1.80	-1.23	-0.48	6.20	0.41	-0.82	4.25	4.75	-0.26	-1.36
II-14	5.23	-4.13	-0.67	4.41	4.33	0.02	2.82	3.52	-0.47	-0.86
III-14	1.13	-8.20	-1.41	2.45	1.26	-7.04	1.14	2.59	-0.11	0.13
IV-14	4.93	4.15	-2.85	9.01	6.37	10.78	7.36	10.51	3.20	1.37

* In this table, equity risk premium is calculated as the geometric difference between large-cap stock total returns and U.S. Treasury bill total returns.

Table 9
Derived Series: Monthly Index Values
Dec. 31, 1925 = $1.00

	Inflation Adjusted Total Return ($)					
Month	Large Stocks	Small Stocks	LT-Corp Bonds	LT-Govt Bonds	IT-Govt Bonds	T-Bill
12/13	359.728	2,049.227	12.446	8.395	7.152	1.583
1/14	346.004	1,951.187	12.810	8.781	7.236	1.577
2/14	360.498	2,025.841	12.978	8.740	7.229	1.571
3/14	361.202	2,032.403	12.975	8.794	7.133	1.561
4/14	362.677	1,956.647	13.140	8.886	7.179	1.556
5/14	369.898	1,951.787	13.340	9.095	7.187	1.551
6/14	376.838	2,032.709	13.342	9.103	7.100	1.548
7/14	371.786	1,914.746	13.379	9.207	7.101	1.549
8/14	387.306	2,005.025	13.879	9.509	7.104	1.551
9/14	381.587	1,889.517	13.493	9.339	7.092	1.550
10/14	391.892	2,017.783	13.831	9.694	7.270	1.554
11/14	404.617	2,013.521	14.146	9.906	7.309	1.562
12/14	405.899	2,093.246	14.487	10.320	7.319	1.571

Table 10
Long-Horizon Expected Equity Risk Premium and Size Premium
As of Dec. 31, 2014

Equity Risk Premium

Long-horizon expected equity risk premium (historical): large-cap stock total returns minus long-term government bond income returns [1]	7.00%
Long-horizon expected equity risk premium (supply-side): historical equity risk premium minus price-to-earnings ratio calculated using three-year average earnings	6.19%

Size Premiums (market capitalization in millions) [2]

Decile	Smallest Company		Largest Company	Size Premium (Return in Excess of CAPM)
Mid-Cap (3-5)	2,552,441	–	10,105,622	1.10%
Low-Cap (6-8)	549,056	–	2,542,913	1.77
Micro-Cap (9-10)	3,037	–	548,839	3.69

Breakdown of Deciles 1-10

1 – Largest	24,428,848	–	591,015,721	-0.32%
2	10,170,746	–	24,272,837	0.65
3	5,864,266	–	10,105,622	0.94
4	3,724,624	–	5,844,592	1.05
5	2,552,441	–	3,724,186	1.65
6	1,688,895	–	2,542,913	1.63
7	1,011,278	–	1,686,860	1.77
8	549,056	–	1,010,634	2.18
9	300,752	–	548,839	2.64
10 – Smallest	3,037	–	300,725	5.72

[1] Expected equity risk premium is based on the difference of historical arithmetic mean returns for 1926-2014. Large-cap stocks are represented by the S&P 500 Index.

[2] Return in excess of CAPM estimation. Mid-Cap stocks are defined here as the aggregate of size-deciles 3–5 of the NYSE/AMEX/NASDAQ; Low-Cap stocks are defined here as the aggregate of size-deciles 6–8 of the NYSE/AMEX/NASDAQ; Micro-Cap stocks are defined here as the aggregate of size-deciles 9–10 of the NYSE/AMEX/NASDAQ. The betas used in CAPM estimation were estimated from CRSP NYSE/AMEX/NASDAQ decile portfolio monthly total returns in excess of the 30-day U.S. Treasury bill total return versus the S&P 500 total returns in excess of the 30-day U.S. Treasury bill, January 1926–December 2014. Calculated (or derived) based on data from CRSP US Stock Database and CRSP US Indices Database ©2015 Center for Research in Security Prices (CRSP®), The University of Chicago Booth School of Business. Used with permission.

Glossary

Bond Default Premium
Calculated as the geometric difference between long-term corporate bond total returns and long-term government bond total returns.

Bond Horizon Premium
Calculated as the geometric difference between long-term government bond total returns and Treasury bill total returns.

Equity Risk Premium
Calculated as the geometric difference between large-capitalization stock total returns and U.S. Treasury bill total returns.

Inflation
Represented by Consumer Price Index for All Urban Consumer (CPI–U), not seasonally adjusted.

Intermediate–Term Government Bonds
Measured using a one-bond portfolio with a maturity near five years.

Large Capitalization Stocks
Represented by the Standard and Poor's 500 Stock Composite Index® (S&P 500) 1957–present; and the S&P 90, 1926–1956.

Long–Term Corporate Bonds
Represented by the Citigroup long-term, high-grade corporate bond total return index.

Long–Term Government Bonds
Measured using a one-bond portfolio with a maturity near 20 years.

Small-Capitalization Stocks
A portfolio of stocks represented by the fifth capitalization quintile of stocks on the NYSE for 1926–1981. For January 1982 to March 2001, the series is represented by the DFA U.S. 9–10 Small Company Portfolio and the DFA U.S. Micro Cap Portfolio thereafter.

Small Stock Premium
Calculated as the geometric difference between small-cap stock total returns and large-cap stock total returns.

U.S. Treasury Bills
Measured by rolling over each month a one-bill portfolio containing, at the beginning of each month, the bill having the shortest maturity not less than one month.

METHODOLOGY FOR ESTIMATES

A GUIDE TO UNDERSTANDING THOMSON REUTERS METHODOLOGIES, TERMS AND POLICIES FOR THE FIRST CALL AND I/B/E/S ESTIMATES DATABASES

OCTOBER 2009

Please visit Thomson Reuters Customer Zone, the single online entry point for Thomson Reuters support and service functions, including on-line support capabilities via the "Contact Us" link located at the top left of the page.

If you choose to contact Thomson Reuters by phone, please click here for global product contact information.

THOMSON REUTERS

NOTICE

THOMSON REUTERS

TABLE OF CONTENTS

THOMSON REUTERS

INTRODUCTION

About Thomson Reuters

Thomson Reuters is the most complete source for integrated information and technology applications in the global financial services industry. Working in partnership with our clients, we develop individual workflow solutions that answer their specific data and analysis needs. Among those needs, clients would like insight on future earning prospects of publicly traded companies. As a result, Thomson Reuters tracks the reported and forecast earnings of these firms globally. *Earnings Per Share* is a key metric, and one most commonly utilized in two ways: to measure performance gains and to gauge companies' results versus expectations.

About This Document

This document provides an in depth look at the methodologies Thomson Reuters uses for estimates. The purpose of this document is to outline, describe and provide reference for the different policies that affect Thomson Reuters estimates data.

ACCOUNTING REGULATIONS

International Financial Reporting Standards (IFRS)

The European Union has passed a regulation that requires listed European companies to comply with International Financial Reporting Standards (IFRS) in 2005 for their consolidated financial statements. There is a limited exception for certain companies to delay implementation until 2007. Generally, the regulation applies to consolidated financial statements for accounting periods starting on or after January 1, 2005. Thus for those companies with 12-month accounting periods covering the calendar year, IFRS will first apply to periods ending on December 31, 2005. As a result, companies will first publish IFRS financial information as at March 31, 2005 (if they report quarterly) or as at June 30, 2005 (if they report semi-annually).

Estimates collected by Thomson Reuters will reflect the adoption of this ruling on a majority basis. The transition period to IFRS is visible for companies in Europe effective April 25, 2005. In addition to countries in Europe, IFRS will be adopted by parts of Asia, including Australia and New Zealand. The transition period to IFRS is visible for companies in Australia and New Zealand effective September 12, 2005.

Dedicated company level footnotes are used to label the majority accounting basis for the company, as well as estimate level footnotes to label and exclude minority accounting basis estimates.

Instrument Level Footnote Code (Majority)	Footnote Text
3	Earnings on a fully adjusted basis
4	Accounting differences exist: Estimate on a Fully-Reported/GAAP basis
W	Estimates based on IFRS

Estimate Level Footnote Code (Minority)	Footnote Text
3	Earnings on a fully adjusted basis
4	Earnings on a fully reported basis
W	Estimates based on IFRS

FAS123(R)

On December 16, 2004, The Financial Accounting Standards Board (FASB) issued FAS123(R). This ruling requires companies to calculate the fair value of stock options granted to employees, and amortize that amount over the vesting period as an expense through the income statement. FAS123(R) is currently effective for fiscal years beginning after June 15, 2005, with company transition choices of: modified prospective, modified retrospective or early adoption. The effective date of the ruling was then extended from quarterly to annual periods beginning after June 15, 2005.

Thomson Reuters will treat the expensing of stock options on a company-by-company basis. Stock option expenses will only be included in the primary EPS mean when the majority of the contributing analysts have included the expenses in their estimates. Estimates will be footnoted describing whether estimates include or exclude the options expense. Once the majority of the analysts are including stock option expenses in their estimates, the remaining estimates that do not include the expenses will be footnoted, filtered, and excluded from the primary EPS mean calculation. In the event that a contributing analyst provides two sets of EPS estimates for a given company (one including options expenses and one excluding), the majority basis estimate will appear under the EPS field and the alternative estimate will appear under the EPX field.

The GAAP EPS measure (GPS) will however, include option expenses per FAS123(R) for periods where GAAP requires the inclusion of option expenses in reported results, and when the impact is known. When available, estimates from contributing analysts on a GAAP basis appear under the GPS measure.

For periods where GAAP requires the inclusion of stock options expense, estimates excluding stock options expense will be filtered and footnoted once the impact of stock options expense is known for that period, as determined by any of the following:

- company issued guidance,
- a quarterly report,
- the presence of a GAAP estimate including options expense from a single contributor.

For example, if 10 brokers provide a GPS estimate that excludes stock options expense, but 1 broker provides an estimate that includes stock options expense for a period where GAAP requires inclusion, the 10 brokers excluding options will be filtered and footnoted and the 1 broker will remain unfiltered and comprise the GPS mean.

Dedicated company level footnotes are used to label the majority accounting basis for the company, as well as estimate level footnotes to label and exclude minority accounting basis estimates.

Company Level Footnote Code (Majority)	Footnote Text
E	Estimates reflect adoption of FAS123(R)
F	Estimates do not reflect adoption of FAS123(R)
I	Estimates have always reflected adoption of FAS123(R)
N	No known impact from FAS123(R) on estimates

Estimate Level Footnote Code (Minority)	Footnote Text
5	Estimate includes stock option expenses
6	Estimate excludes stock option expenses

FASB APB 14-1

On May 9, 2008 The Financial Accounting Standards Board (FASB) issued FASB APB 14-1. This ruling requires companies to change how they account for convertible debt in their financial statements - specifically, debt that can be converted into cash. Companies will be required to amortize the excess of the principal amount of the liability component over its carrying amount. This will result in higher interest costs. The effective date of the change will be the first fiscal year that begins after December 15, 2008, and will impact 2009 fiscal year estimates for most companies. For US traded companies carrying this type of debt, GAAP earnings will be negatively affected starting with 2009.

Thomson Reuters will treat estimates impacted by FASB Staff Position APB 14-1 on a company-by-company basis. Post-FASB APB 14-1 estimates will only be included in the EPS mean when the majority of the contributing analysts have adopted this accounting change in their estimates. Estimates will be footnoted describing whether estimates reflect or do not reflect the accounting change. Once the majority of analysts reflect FASB APB 14-1 in their estimates, the remaining estimates that do not include the expenses will be footnoted, filtered, and excluded from the EPS mean calculation.

The GAAP EPS (Fully Reported) measure will be post FASB APB 14-1 for periods where GAAP requires the amortization of cash-convertible debt in reported results and when the impact is known. When available, estimates from contributing analysts on a GAAP basis appear under the GAAP EPS measure on Thomson Reuters products.

Dedicated company level footnotes are used to label the majority accounting basis for the company, as well as estimate level footnotes to label and exclude minority accounting basis estimates.

Company Level Footnote Code (Majority)	Footnote Text
8	Estimates reflect FASB APB 14-1
9	Estimates do not reflect FASB APB 14-1

Estimate Level Footnote Code (Minority)	Footnote Text
8	Estimate reflects FASB APB 14-1
9	Estimate does not reflect FASB APB 14-1

ACTUALS

Evaluation

Thomson Reuters Market Specialists enter both quarterly period and annual actuals where analyst estimates exist on a real-time global basis - as sourced from multiple newswire feeds, press releases, company websites and public filings. When a company reports their earnings, the data is evaluated by a Market Specialist to determine if any Extraordinary or Non-Extraordinary Items (charges or gains) have been recorded by the company during the period. If no items have been recorded during the period the reported value is entered. If one or more items have been recorded during the period, actuals will be entered based upon the estimates majority basis at the time of reporting. The Market Specialist will still review each item in relation to the estimate submissions and how similar items have been treated in past periods. If after review it is determined that majority basis is to be changed, Thomson Reuters will update the actual and corresponding surprise values accordingly.

Certain differences exist across regions pertaining to prioritization, coverage, and timeliness. Companies in Asia-Pacific, North America and Latin America are updated the same day of reporting. In the EMEA region, Tier1 companies (445 companies including FTSE 100 and other major indices) are also updated the same-day of reporting, with the Tier 2 companies updated within 15 days.

Please note that Thomson Reuters collects actuals only for periods and measures where current analyst estimates exist.

Majority Basis

Thomson Reuters goal is to present actuals on an operating basis, whereby a corporation's reported earnings are adjusted to reflect the basis that the majority of contributors use to value the stock. In many cases, the reported figure contains unusual or one-time items that the majority of analysts exclude from their actuals. The majority accounting basis is determined on a quarter-by-quarter basis. Typical adjustments are for the effects of extraordinary and non-extraordinary items.

Thomson Reuters examines each reported item, and includes or excludes the item from the actual based on how the majority of contributing analysts treat the item for that period. Once the Thomson Reuters Market Specialist determines whether the item is being included or excluded by the majority of contributors, they will enter the actual and a footnote detailing the type of the item, whether it is included or excluded, the size of the item, and the period affected.

If after the comparable actual for the period is saved for a company and a go-forward majority is established on a different accounting basis, that actual will be replaced to reflect the change and footnoted to indicate the majority basis change. The announce and activation dates of the original comparable actual will remain.

Any submission of an estimate by a contributing analyst using a non majority actual or on a non majority basis results in a call from a Thomson Reuters Market Specialist requiring the contributing analyst to adjust to the majority basis or have their estimates footnoted for an accounting difference and excluded from the mean calculation for the fiscal years in question. In all cases, appropriate footnotes are added to the estimate to denote what items are included or excluded. In some cases, a company's actuals number will be temporarily withheld so that analysts may be contacted and additional research conducted.

Elimination of Held-Out Actuals Practice (September 2009)

Thomson Reuters made changes to the collection of actuals to provide increased data timeliness. As companies report, values will be adjusted to the estimates majority basis for the period, then entered into the database without a "hold out" period.

- Previously, when a company reported results, actuals were collected according to the estimates majority basis for the period at the time of report. If however, unexpected charges or gains were reported, actuals would temporarily be "held out" from products to see if the majority basis would change going forward.
  - This process introduced possible timeliness issues whilst the sell-side analyst community reacted to the company news and issued reports, and subsequently Thomson Reuters re-evaluated the majority basis.
- Going forward, this "hold out" period will be eliminated in cases where unexpected charges or gains are reported. Actuals will be entered strictly based upon the estimates majority basis at the time of report – significantly increasing timeliness of actuals under these scenarios.
  - The review of analyst reaction will still be done by Thomson Reuters, however only after the actual was already saved to the database and available on products.
  - If the analyst majority basis changes after the fact, Thomson Reuters will update the actual and corresponding surprise values accordingly, and footnote the reason.

BASIC VS. DILUTED ESTIMATES

Dilution occurs when a company issues securities that are convertible into common equity. Such issues can take the form of convertible bonds, rights, warrants or other instruments. When Thomson Reuters refers to "fully diluted" earnings estimates it means that the forecasts assume that all eligible shares are converted. Fully diluted earnings per share are, by definition, less than basic EPS (which is based solely on common shares outstanding).

- To be an eligible convertible security, the contributing analyst must predict that the share price will be greater than the strike price.

- If the contributing analyst predicts that the convertible security will be eligible, the convertible shares are included in the analyst's share count, and the interest expense associated with the conversion is included in their EPS estimate. If the contributing analyst does not predict the convertible security will be eligible, the share count does not include the convertible shares, and there is no interest expense associated with the convertible. (Interest expense is associated with the conversion and this scenario has no conversion.)

Thomson Reuters determines whether a company is followed on basic versus diluted shares based on the majority rule. If a contributor is on the minority basis, the estimate is filtered, footnoted and excluded from the mean calculation using the estimate level footnotes listed below.

Estimate Level Footnote Code (Minority)	Footnote Text
B	Accounting differences exist: Estimate on a basic share count basis
E	Accounting differences exist: Estimate on a diluted share count basis

North America

Thomson Reuters defaults to using diluted shares in North America, as this is the most widely used valuation method. Estimates are displayed on a diluted basis taking into account all eligible convertible securities. The only circumstances where basic shares would be the default for a company would be when a company reports a loss, as basic is the more conservative valuation method.

International

For international companies, Thomson Reuters determines whether a company is followed on basic vs. diluted shares based on the majority rule, due to the high amount of variance in which companies are followed. In cases where an analyst follows a company on a basis that is different from the mean, filters/footnotes are applied to their estimates, which are then excluded from the mean calculation.

CORPORATE ACTIONS

Corporate actions are defined as any event which can bring material change to a stock, which include the following:

- Mergers
- Acquisitions
- Spin-offs
- Stock splits

Thomson Reuters obtains information on corporate actions via real-time news feeds as well as information received directly from companies. Thomson Reuters Market Specialists then process corporate actions on a real-time basis. Thomson Reuters Market Specialists verify the corporate action announcement by using original press releases from companies. Corporate action announcements are then footnoted in the appropriate tables (see examples below):

Estimate Level Footnote Code (Minority)	Footnote Text
L	Accounting differences exist: Estimate reflecting corporate action
V	Contributor update pending: Estimate not reflecting corporate action
A	Accounting Differences Exist

Example:

St. Paul Travelers Cos Inc. (ticker STA)

Corporate Action Announcement: 17-Nov-03 announced merger with Travelers Property Casualty Corp.

Mergers, Acquisitions and Spin Off's

Thomson Reuters will reflect estimates on the post-event basis, reflecting the completion of a merger/acquisition/spin-off, when the first of two events occur:

- The majority of analysts covering the company submit estimates on a post-event basis or;
- The event itself actually closes/completes (usually signified by a press release on or around the closure date).

When a corporate action occurs, before Thomson Reuters makes any data changes, all of the following action details are thoroughly researched:

- **All information must be confirmed,** including the action, the date, and how current and historical estimates will be treated going forward. For example, to which company estimates will be attached.
- Great importance is also placed on how the company will be treating its financial statements going forward. This research is done by using Datastream, the company's website, or by contacting the company's IR group directly.
- The corporate action is always treated in the database in accordance with the company's guidelines (who will be the surviving entity, etc.).

Policies involved with introducing the Merger/Acquisition include:

- Footnotes will be added describing the announced merger/acquisition to all publicly traded companies involved that we have established in our database.
- All Thomson Reuters mean estimates will reflect a merger/acquisition according to how the majority of analysts covering the company treat the action. The mean will follow this majority policy up until the date the merger/acquisition closes. An additional footnote will be added to the database detailing how the mean is treating the action that will remain present until the action closes. Once the merger/acquisition is closed and finalized, the estimates must reflect the full affects of the action.
- Upon the date of closing several actions may need to be taken on the part of Thomson Reuters depending on the type of merger/acquisition that has occurred. All of the possible actions performed are to update the Thomson Reuters estimates database to reflect all effects of the closed corporate action. Below are some broader steps taken but more specific instructions are listed with each possible scenario below:
- The closing of the merger/acquisition is footnoted. All records and consensus data for surviving or newly formed companies affected by the merger/acquisition must now fully reflect the effects of the completed corporate action. This may involve company name or identifier changes of the acquiring company or the creation of a completely new entity in our database formed through a merger. It will involve making sure all estimate data included in consensus for these companies reflects the completed action. Historical estimates for the surviving company, normally the company doing the acquiring, will remain.
- If a company has been acquired or merges with another and no longer exists as a separate entity, the estimates/recommendations/price targets associated with that ticker must be stopped and the ticker end-dated upon closing of the action. Since the company will no longer exist, there will be no visible outstanding or active records on our products or database. Please note that when estimates are stopped, the user will not have a link between the former company and the newly created one. Thomson Reuters does, however, keep a record of the movement of companies in the central estimates database.

The policies Thomson Reuters follows in the case of Spin-Off/De-Merger include:

- Footnotes are added describing the announced spin-off/demerger to all publicly traded companies involved that are established in the Thomson Reuters database.
- All mean estimates will reflect a spin-off/demerger according to how the majority of analysts covering the company treat the action. The mean will follow this majority policy up until the date the spin-off/demerger closes. An additional footnote will be added to the database detailing how consensus is treating the action that will remain present until the action closes. Once the spin-off/demerger is closed and finalized, the estimates must reflect the full effects of the action.
- Upon the date of closing several actions may need to be taken on the part of Thomson Reuters depending on the type of spin-off/demerger that has occurred. All of the possible actions performed are to update the estimates database to reflect all effects of the closed corporate action. Below are some broader steps taken but more specific instructions are listed with each possible scenario below:
- The closing of the spin-off/demerger is footnoted. All records and consensus data for surviving or newly formed companies affected by the spin-off/demerger must now fully reflect the effects of the completed corporate action. This may involve the creation of a completely new entity in the estimates database formed through the spin-off/demerger. This will involve making sure that all estimate data included in consensus for these companies reflect the completed action.
- If a previously existing company will no longer exist or no longer trades publicly, all estimates, recommendations and price targets must be stopped and the ticker end-dated upon closing of the transaction.

Stock Splits & Stock Dividends

A security begins trading on a post-split or post-stock dividend basis the day after the payment date (date the declared split or dividend is paid). Thomson Reuters enters a footnote that indicates the size of the stock split or stock dividend and the effective date (the day after the payment date).

After the market closes on the day before the stock begins trading on the new basis, all estimates data in Thomson Reuters – both current and historical - will be adjusted for the new shares. If a contributing analyst submits estimates on an adjusted basis prior to the effective date or unadjusted basis after the effective date, Thomson Reuters will contact that analyst to request properly adjusted estimates.

Please note that Thomson Reuters does not make adjustment factors for corporate actions *which do not affect the number of shares.* This document describes the actions taken when a company's share count changes. This could include, but is not limited to, spin offs, mergers or cash payments / special payments.

Example of Stock Split:

Meritage Homes Corp [MTH]

Footnote: 20-Dec-04 2 for 1 Split Effective 10-Jan-05

Thomson Reuters does not adjust estimates for cash payments. The effect of cash payments on estimates is treated as a revision by the contributing analyst. On the effective date of the cash payment, a Thomson Reuters market specialist will contact all contributing analysts to request updated figures that include the cash payment. Estimates that are not updated to reflect the cash payment are footnoted as update pending, and will be filtered from the mean until they are updated by the contributing analyst.

Example of Stock Split with Cash Payment:

United Business Media PLC [UBM]
14 for 17 share consolidation
Special cash dividend of 89p per share

Thomson Reuters will apply a split factor of 1.214 reflecting the share consolidation. It is expected that contributors will revise their models to reflect the 89p cash dividend. Contributors that do not revise their estimates to reflect the cash dividend will be footnoted as update pending and filtered from the mean estimate.

Rights Issues

Rights Issues are treated in the following manner:

- When rights issues becomes effective, like stock splits, the ex date triggers all current and historical adjustments for price, shares and earnings.
- Even before the majority of analysts switch to post rights issue estimates, estimates will be collected and displayed on products prior to the ex-date, but will be excluded from the mean with a new estimate level footnote type:

Estimate Level Footnote Code (Minority	Footnote Text
7	Accounting differences exist: Estimate reflecting rights issue prior to ex-date

- Once the ex-date occurs, footnotes of excluded estimates will be automatically end-dated and will be then added back into the mean calculation where appropriate.

CONTRIBUTOR REQUIREMENTS

In order to maintain a quality, professional standard for all contributing analysts, Thomson Reuters Contributor Relations requires a candidate to pass a strict set of guidelines before being enlisted as a contributor. A potential contributor must provide information to establish that they are a reputable firm. This process includes providing example research reports, three references from institutional clients, three references from company investor relations, detail on the number of companies covered per analyst in the firm, and background information on the director of research. Thomson Reuters currently collects and analyzes the research, ratings and forecasts from many different sell-side or independent contributors.

Please reference the Thomson Reuters Contributor Approval Policy document for further details.

CURRENCY

The default currency displayed on Thomson Reuters is generally the currency in which the company reports*. Thomson Reuters will however, accept estimates in any currency.

The following describes the treatment of non-default currency conversions on Thomson Reuters products: (Please note that product update schedules vary for currency conversions.)

- All estimates revisions received in a non-default currency are updated using the prior day's currency conversion rate.
- All non-default estimates have the currency conversion recalculated on Friday night using Friday's end of the day conversion rate.
- When a contributing analyst confirms a default currency estimate, there is no change in the raw value estimate stored in the database.
- Thomson Reuters provides normalized Summary and Detail history offerings which provide a smooth historical view for companies that have had a currency change over time and it is intended to simplify clients' workflow.

A confirmation of a non-default currency estimate however, does result in a reconverted estimate being sent to products. This estimate will represent the conversion rate as of the day prior to the confirmation.

Please note one exception: the per-share data measures of United Kingdom companies are always covered in BPN (pence) and the values for non-per share data measures are displayed in GBP (pounds). The label for all estimates, regardless of per share or non-per share measure type however are BPN.

Treatment of Currency Changes

Thomson Reuters follows companies based on their reporting currency. In some cases however, where the reporting currency does not reflect the clear majority of estimate submissions, Thomson Reuters may exercise the option to set the default based on the currency of the majority of estimate submissions. In cases where companies report in multiple currencies, Thomson Reuters will set the default currency based on the majority of estimate submissions.

Occasionally, companies will change the currency in which they report and/or the majority of analysts covering a company will change the currency of their estimates. As a result, Thomson Reuters will change the default currency of a company in order to align with the reporting company or majority of contributing analysts as part of the operational process.

Normalized Summary & Detail History (Currency)

Thomson Reuters provides normalized summary and detail history in addition to regular summary and detail history, providing a smooth historical view for companies that have had a currency change over time and it is intended to simplify clients' workflow. Whereas the regular summary and detail history offering provides a clear time series of when a company changes reporting currencies, the normalized offering will provide all historical estimates for a company in the current reporting currency of that company.

ENTITLEMENTS INFORMATION

Thomson Reuters is recognized for providing the most timely and accurate estimates data available to investment professionals. This is made possible in part by an agreement with our contributing analysts which restricts the distribution of individual analyst's estimates to certain parties.

The following policy is strictly adhered to:

- Individual estimates with the associated contributor names are provided exclusively to institutional 'buy-side' investors and the research departments of the contributing analysts.
- Institutional investors are defined as users who are involved in executing trades through multiple brokerage firms.
- Investment banking, corporate finance and trading firms are not considered institutional investors as they do not have a trading relationship with any of the contributing firms and in effect, are competitors of those contributing analysts. Therefore, these firms are not privy to seeing individual analyst's earnings estimates.
- Analyst's research is considered proprietary information, unlike news articles or SEC filings. Detailed earnings estimates are also considered a part of an analyst's research and therefore proprietary in nature.

Examples of disentitlement views by product would be:

- Thomson ONE — Broker and analyst names are displayed while displaying estimate value as "PERMISSION DENIED"
- First Call — Blank records for entire entry are sent with the detail record – no broker or analyst name or estimate value are displayed.

THOMSON REUTERS

- I/B/E/S — Estimator and Analyst Name will be replaced by a numeric code, effectively meaning “Permission Denied” while displaying estimate value.

In order to gain access to the research reports of a broker with ‘Prior Approval’ status, a client need only speak with their Thomson Reuters ***Relationship Manager*** or ***Sales Representative*** directly. Thomson Reuters will contact those brokers in question and seek approval to access their reports on behalf of the client. If approved, the client will have access to view the research reports within 24-48 hours.

ESTIMATES COLLECTION

Process

Thomson Reuters gathers earnings forecasts and other data from hundreds of brokerage and independent analysts who track companies as part of their investment research work. Thomson Reuters calculates a mean consisting of estimates utilizing the same accounting standards (basis).

Majority Policy

Most institutional clients prefer to view estimates on an “operating” basis, reflecting the majority of the analysts covering a security. Consequently, Thomson Reuters follows a ‘majority’ policy, where the accounting basis of each company estimate is determined by the basis used by the majority of contributing analysts.

Once the majority basis has been established, contributing analysts in the minority may keep their original estimates, or are also given the opportunity to adjust to the majority basis. On rare occasions, the majority basis may be revised as additional analysts are heard from or as some change their opinion. In all cases, appropriate footnotes are added to the Thomson Reuters database stating the appropriate basis of each estimate, and if the item has been included or excluded from the mean estimate.

Adoption of Post-Event Mean (as of September 2009)

As of September 21, 2009, Thomson Reuters adopted more stringent updating rules for analyst’s estimates which are not reflecting current company events, such as:

- **Issuance of Company Guidance**
  Detail estimates which have not been updated or confirmed following the issuance of guidance and do not fall within the guidance range (e.g. “$1.00 - $1.10”) will be filtered / excluded from the mean at the time of guidance. In those cases where single-point guidance is issued (e.g. “about $1.00”), estimates not within 5% of the guidance will be footnoted and excluded from the mean. The aforementioned guidance filter will only apply to the specific measure and period.

  Those estimates that are excluded will be labeled with a (N) estimate level footnote. Then, excluded estimates that are updated or confirmed will have the footnote end-dated and added back into the mean calculation.

- **Actual(s) Reporting**
  Detail estimates for unreported periods which are not updated or confirmed within 10 business days of a prior-period reported actual will be excluded from the mean, based on the reporting of the EPS actual for that/their specified period(s).

  Those estimates that are excluded from the mean will be labeled with a type (P) estimate level footnote. The reported actual(s) filter will be applied to all measures and subsequent periods for that fiscal year. Then, excluded estimates that are updated or confirmed will have the footnote end-dated and added back into the mean calculation.

Estimate Level Footnote Code (Minority)	Footnote Text
N	Contributor update pending: Estimate not reflecting recent company guidance
P	Contributor update pending: Estimate not reflecting recent reported actual

THOMSON REUTERS

Please note that all other scenarios, including corporate actions, will continue with the original policy of waiting for the full majority of analyst treatment however they will be enhanced with new descriptive footnotes, illustrated below in the Footnotes section of this document.

Extraordinary Items

Extraordinary items are defined by the accounting conventions of the Financial Accounting Standards Board. Companies are required to present extraordinary items as a separate item in their financial statements. Thomson Reuters will always exclude them from the reported figures, since the majority of contributing analysts always choose to exclude extraordinary items. Thomson Reuters uses the word "extraordinary" in the most limited sense as defined by accounting convention (some analysts have the habit of applying the word "extraordinary" to any unusual charges or gains).

The most common extraordinary items are:

- Cumulative Effect of FASB Accounting Changes
- Tax Loss Carry forwards
- Discontinued Operations
- Early Retirement of Debt

Please note that as each quarter is treated independently of each year, any exclusion from a given quarter would result in an exclusion from the annual estimate

Example:	*Q1*	*Included*
	Q2	*Excluded, minority basis*
	Q3	*Included*
	Q4	*Included*
	FY	*Excluded, due to Q2 exclusion*

Non-Extraordinary Items

Non-extraordinary and non-operating items are charges or gains that may or may not be seen as pertinent to ongoing operations, depending on the industry and the opinion of the majority of contributing analysts. In contrast to the uniform recognition of extraordinary items, there is a great deal more variance within the analyst community concerning the treatment of non-extraordinary/non-operating items.

When submitting estimates, contributors are encouraged to include or exclude any non-extraordinary items they deem non-recurring and/or non-operating. Once a non-extraordinary or non-operating item is recognized, a Thomson Reuters Market Specialist will poll all contributor's estimates covering a particular company, to establish if the majority of them are including or excluding the event. If there is no clear majority, then the charge or gain is included in the mean. If at any point the majority basis cannot be determined, the Thomson Reuters Market Specialist will further research the affected estimates, including potentially contacting the contributing analysts, to determine the majority basis.

Examples of Non-Extraordinary items include:

- Restructuring charges - larger ones are usually excluded
- Asset sale gains or losses - larger ones are usually excluded
- Inventory adjustments - included in the majority of cases
- Currency adjustments - included in the majority of cases; always included in the Oil industry
- Realized securities gains or losses - always excluded in the Insurance industry; always included in the Banking industry
- Acquisition expenses or gains from acquisition - larger ones are usually excluded
- Litigation charges or gains from litigation
- Tax settlements or adjustments
- Write-offs

Majority Basis Footnotes

A new series of valuable company and estimate level footnotes is now available for enhanced transparency of estimate accounting basis and rationale for exclusions.

COMPANY LEVEL FOOTNOTE

Footnote Code	Footnote Text
M	Majority Basis includes/excludes… (freeform criteria utilized to define specific accounting scenario of the mean calculation)

This new company level footnote is designed for flexibility, and as such it will be edited to reflect any specific company scenario. Just a few possible examples of what this new freeform footnote will label include, but are not limited to, the following:

- Majority Basis excludes restructuring charge
- Majority Basis includes tax adjustment gain
- Majority Basis includes currency adjustment gain
- Majority Basis excludes litigation charge

ESTIMATE LEVEL FOOTNOTES

In addition to labeling a company's majority accounting basis, Thomson Reuters also introduced new estimate level footnotes to clarify the specific reasoning of why an estimate was excluded from the mean. Both the company and estimate level footnotes work in tandem in the event of a change in basis (e.g. if a company's basis changes, both sets of footnotes will be 'flipped' to account for the new majority basis).

New / Modified footnotes to be used are as follows:

Footnote Code	Footnote Text
4	Accounting differences exist: Estimate on a Fully-Reported/GAAP basis
7	Accounting differences exist: Estimate reflecting rights issue prior to ex-date
B	Accounting differences exist: Estimate on a basic share count basis
E	Accounting differences exist: Estimate on a diluted share count basis
G	Accounting differences exist: Excludes charge(s)
H	Accounting differences exist: Includes charge(s)
I	Accounting differences exist: Excludes gain(s)
J	Accounting differences exist: Includes gain(s)
L	Accounting differences exist: Estimate reflecting corporate action
M	Accounting differences exist: Estimate on a non-GAAP basis
X	Accounting differences exist: Estimate on a Cash EPS basis
N	Contributor update pending: Estimate not reflecting recent company guidance
O	Contributor update pending: Estimate failed freshness policy
P	Contributor update pending: Estimate not reflecting recent reported actual
V	Contributor update pending: Estimate not reflecting corporate action

Existing footnotes which will continue to be used where appropriate are as follows:

Footnote Code	Footnote Text
3	Earnings on a fully adjusted basis
5	Estimates Include Stock Options Expense
6	Estimates Exclude Stock Options Expense
8	Estimate reflects FASB APB 14-1
9	Estimate does not reflect FASB APB 14-1
A*	Accounting Differences Exist
C	Estimate Received directly from Analyst

THOMSON REUTERS

D	Est rec'd in currency other than default
F	Freeform Footnote
K	Forecast estimate not a 12-month figure
S	Estimate Confirmed in analysts notes.
T	Accounting basis unknown - contributor contacted
U*	Contributor Update Pending
W	Estimates based on IFRS

Please note that whenever possible, the newly created granular footnotes above will be used, but the existing "A" and "U" footnotes will still continue to be utilized when multiple minority basis scenarios exist.

ESTIMATES TO RESEARCH LINKING (JUMP-TO)

Through use of the Thomson ONE platform, clients subscribing to both Detail-Estimates and Real-Time Research reports have the capability to click from a sell-side analyst's estimate to the exact research document from which it was sourced. This will provide greater transparency to identify the details around estimate movements and pinpoint the exact reasons why a contributor is revising or confirming an estimate.

Estimates sourced directly from a research report contain a link to the exact report from where the estimate was first received (identified on the platform as any underlined estimate value in blue). If the estimate was confirmed more recently, an additional link will display to take the user to the most recent confirmation document.

These links are offered for current or previous estimates available on the detail estimates, full year, all measures and revision analysis pages of Thomson ONE.

Note that a user must be entitled to Real-Time Research to be able to see the Estimates to Research (Jump-To) functionality. Additionally the page will only contain links to contributor's documents the user is entitled to view.

Please note: If Estimates were received through automated feeds or files, the value will display without a link.

FISCAL YEAR

The fiscal year displayed on Thomson Reuters products is determined by the calendar year the last month of the fiscal year falls in. For example, if a company reports fiscal year results ending in January 2007, they are reporting Fiscal Year 2007. If a company reports fiscal year results ending in October 2006, they are reporting Fiscal Year 2006.
Thomson ONE platforms contain estimate data for up to five annual fiscal periods, four quarterly fiscal periods and long-term growth. (Analysts typically do not make forecasts for periods beyond the third fiscal year and fourth quarter.) Since not all companies have the same fiscal year end, Thomson Reuters uses the familiar FY1, FY2... convention to identify estimates for each unique period.

The following is a description of how this labeling technique works:

- The most recently reported earnings number is denoted as time slot **0 (** can be FY, Q, or SAN).
- A company's last reported annual earnings is referred to as FY0, the most recently reported quarter is Q0 and the most recent semiannual reported earnings is SAN0.
- Using these periods as a base, the period end dates for all estimated periods are easily found.
- If FY0 corresponds to the December 2006 year-end, the FY1 mean estimate is for December 2007 and the FY2 mean estimate is for the period ended December 2008. The same holds true for the interim periods.
- If Q0 refers to the period ended March 2007 (the last reported quarter), then the Q1 estimate is for the June quarter. A frequent misunderstanding is that Q1 refers to the first fiscal quarter instead of the first estimated quarter.

Fiscal Year-End Changes:

- If a company decides to change their fiscal period end, stops will be inserted in the database for all existing estimates on the company with the previous fiscal period end.
- New estimates data will then be collected under the new fiscal period end going forward.

- For example if a company changed from an October year end to December year end, all 10-2007Y estimates would be stopped, then only 12-2007Y estimates would collected on the effective date of the change.

FOOTNOTES

Footnotes are attached to estimates to alert clients as well as Thomson Reuters Market Specialists of special actions or situations affecting estimates. There are three distinct types of footnotes that can be entered: Company, Instrument and Estimate Level Footnotes.

Company-Level Footnotes

Company-level footnotes are footnotes that apply to estimates received from all contributors in a specific measure for a specific period. All company level footnotes apply to the majority EPS accounting basis, which translates down to all related data measures as well. Thomson Reuters Market Specialists use company-level footnotes to relay the majority basis of a table to clients. For example, if the analysts covering a company are including/excluding a specific charge or gain, a Company-level footnote would be attached to clearly identify this.

The footnotes below show the types of Company-level footnotes available:

Footnote Code	Purpose	Footnote Text
8	Accounting	Estimate reflects FASB APB 14-1
9	Accounting	Estimate does not reflect FASB APB 14-1
A	Accounting	Quarters may not add to annual due to changes in shares outstanding
B	Accounting	Estimates reflect adoption of SFAS 142
C	Accounting	Stock Carries Goodwill Amortization
D	Accounting	No Goodwill Amortization Present In Stock
E	Accounting	Estimates reflect adoption of FAS123(R)
F	Accounting	Estimates do not reflect adoption of FAS123(R)
G*	Accounting	Free Form Extraordinary Event Footnote
I	Accounting	Estimates have always reflected adoption of FAS123(R)
M*	Accounting	Majority basis Includes / Excludes <text>
N	Accounting	No Known impact from FAS123(R) on estimates

Footnote utilizes free-form criteria to define specific accounting scenarios of the mean calculation.

Instrument-Level Footnotes

Instrument-level footnotes are footnotes without a time frame or specific measure. These footnotes apply to all estimates entered on a particular ticker across every year and every measure.

For example, if the company tracks FFO instead of EPS, an Instrument-level footnote would be attached to clearly identify this.

Footnote Code	Purpose	Footnote Text
3	Accounting	Earnings on a fully adjusted basis
4	Accounting	Earnings on a fully reported basis
8	Accounting	Estimate reflects FASB APB 14-1
9	Accounting	Estimate does not reflect FASB APB 14-1
A*	Accounting	Accounting Alert. Free Form
C	Accounting	Accounting Alert, Company followed on a Cash Earnings basis
E	Accounting	Estimates reflect adoption of FAS123(R)
F	Accounting	Estimates do not reflect adoption of FAS123(R)
G	Accounting	Accounting Alert, Company earnings before goodwill amortization
I	Accounting	Estimates have always reflected adoption of FAS123(R)
M*	Accounting	Majority basis Includes / Excludes <text>

THOMSON REUTERS

N	Accounting	No known impact from FAS123(R) on estimates
W	Accounting	Estimates based on IFRS

Footnote utilizes free-form criteria to define specific accounting scenarios of the mean calculation

Estimate-Level Footnotes

Estimate-level footnotes are attached to a specific contributor, ticker, year, measure, and/or period estimate.

The footnotes below show the types of Estimate-level footnotes available. The purpose of Estimate-level footnotes is to exclude estimates from the mean calculation, and give a label as to the reason why it is excluded. Footnotes in *italics* however do not automatically exclude estimates from being part of the mean (C, D, F and S).

Footnote Code	Purpose	Footnote Text
3	Accounting	Earnings on a fully adjusted basis
4	Accounting	Accounting differences exist: Estimate on a Fully-Reported/GAAP basis
5	Accounting	Estimate includes stock option expenses
6	Accounting	Estimate excludes stock option expenses
7	Accounting	Accounting differences exist: Estimate reflecting rights issue prior to ex-date
8	Accounting	Estimate reflects FASB APB 14-1
9	Accounting	Estimate does not reflect FASB APB 14-1
A	Accounting	Accounting differences exist
B	Accounting	Accounting differences exist: Estimate on a basic share count basis
E	Accounting	Accounting differences exist: Estimate on a diluted share count basis
G	Accounting	Accounting differences exist: Excludes charge(s)
H	Accounting	Accounting differences exist: Includes charge(s)
I	Accounting	Accounting differences exist: Excludes gain(s)
J	Accounting	Accounting differences exist: Includes gain(s)
K	Accounting	Forecast estimate not a 12-month figure.
L	Accounting	Accounting differences exist: Estimate reflecting corporate action
M	Accounting	Accounting differences exist: Estimate on a non-GAAP basis
T	Accounting	Accounting basis unknown - contributor contacted
W	Accounting	Estimates based on IFRS
X	Accounting	Accounting differences exist: Estimate on a Cash EPS basis
N	Freshness	Contributor update pending: Estimate not reflecting recent company guidance
O	Freshness	Contributor update pending: Estimate failed freshness policy
P	Freshness	Contributor update pending: Estimate not reflecting recent reported actual
U	Freshness	Contributor update pending.
V	Freshness	Contributor update pending: Estimate not reflecting corporate action
C	*Supplemental*	*Estimate received directly from analyst*
D	*Supplemental*	*Est rec'd in currency other than default*
F	*Supplemental*	*Freeform Footnote*
S	*Supplemental*	*Estimate confirmed in analysts notes.*

GLOBAL ESTIMATES FRESHNESS POLICIES

Thomson Reuters strives to provide the freshest estimates content possible to clients and consequently, contributors are asked to regularly send confirmations of their existing estimates. Thomson Reuters maintains active policies on the 'freshness' of estimates provided by contributing analysts. All forecasted data measures are accompanied by original announce and confirmation dates (in Eastern Time) and are subject to policies designed to prevent stale data:

Estimates

If an estimate has not been updated for 105 days, the estimate is filtered, footnoted with the following estimate level footnote and excluded from the mean. (Estimates are updated by a contributing analyst sending a confirmation, revision or drop in coverage.)

Footnote Code	Footnote Text
O	Contributor update pending: Estimate failed freshness policy

- When Q4 is the current reporting period, Q4 and FY1 estimates are an exception to this rule: Q4 and FY1 estimates will be filtered when they have not been updated for 120 days. (This allows extra time for companies to report year-end results.)

If an estimate is not updated for a total of 180 days, the estimate is stopped.

Note:
- All non-updated estimates are auto-filtered at 105 days. If an estimate is later confirmed as current, the filter/footnote/exclusion will be end-dated and the estimate will be confirmed.
- All non-updated estimates are auto-stopped at 180 days. If an estimate is later re-sent by a contributor, it will be treated as a new estimate initiation.

Recommendations

If a recommendation is not updated for a total of 180 days, the recommendation is stopped. (Recommendations are updated by a contributing analyst sending a confirmation, revision or drop in coverage.)

Price Targets

Price target data is stopped at the expiration of it's time horizon (For example, a 12-month price target would be stopped 12 months after it was last revised by a contributing analyst).

GUIDANCE

Guidance is any forward-looking expectation issued directly by a company regarding its future financial performance. Most importantly, guidance is used by company management to manage investor expectations and by investors to evaluate the company and predict future performance. Under current full disclosure regulations, guidance is the only legal method a company can utilize to communicate its expectations to investors.

Thomson Reuters StreetEvents obtains guidance information via real-time news feeds as well as information received directly from companies. Thomson Reuters Market Specialists analyze estimates and guidance together on a real-time basis. Thomson Reuters Market Specialists verify the guidance by using original press releases from companies; comments made by analysts are not used as guidance. Guidance will be evaluated and compared with the earnings estimates mean before reflecting on product.

Issuance of Company Guidance

Detail estimates which are not updated in a timely fashion after the issuance of guidance will be excluded in order to create a post-event mean value. Detail estimates which have not been updated or confirmed following the issuance of guidance and do not fall within the guidance range (e.g. "$1.00 - $1.10") will be excluded from the mean at the time of guidance. If a single-point guidance is issued (e.g. "about $1.00"), estimate(s) not within 5% of the guidance would be excluded from the mean with appropriate addition of footnotes (see below). Once excluded estimates are updated or confirmed, they will have the footnote end-dated and added back into the mean calculation.

Footnote Code	Footnote Text
N	Contributor update pending: Estimate not reflecting recent company guidance

Product Views

In Q307, Thomson Reuters began offering a "Mean/Guidance Comparison" page on **Thomson ONE**, which is separate from the standard StreetEvents guidance offering. This enhancement allows clients to view mean estimates, actuals and guidance on the same accounting basis side-by-side to ensure a consistent analysis. Additionally, guidance and estimates not on the same accounting basis are indicated with a footnote. This comparable guidance data is fielded and adjusted for corporate actions. Most importantly it is normalized and adjusted to match the accounting basis of estimates; percentages are translated into values, extraordinary items are included/excluded to adhere to estimates majority.

Thomson Reuters offers estimates-comparable guidance on 14 data measures for over 2,350 companies globally, with history for the S&P500 back to January 2006.

Thomson Reuters also offers **Thomson Reuters Guidance Datafeed**, bringing I/B/E/S Estimates and Guidance together into one consistent format allowing clients to perform true comparisons. Thomson Reuters Guidance is a unique, intra-day datafeed that offers quantitative (numeric) company expectations from press releases and transcripts of corporate events and plots them alongside the I/B/E/S mean estimate at the time of the release. This offering enables investment professionals to access company expectations alongside earnings forecasts in a single feed, and most importantly, direct from the market-leading source including the benefits of:

- Global coverage
- Historical content dating back to 1994
- Available for fiscal quarters and years
- Announcement dates and timestamps

Estimates Comparable Guidance is available for the following 14 data measures:

Code	Data Measure
CPX	Capital Expenditure
DPX	Dividends Per Share
EBS	EBITDA Per Share
EBT	EBITDA
EPS	Earnings Per Share
FFO	Funds From Operations Per Share
GPS	Fully Reported Earnings Per Share
GRM	Gross Margin
NET	Net Income
OPR	Operating Profit
PRE	Pre-Tax Income
ROA	Return On Assets (%)
ROE	Return On Equity (%)
SAL	Sales

HISTORY

Thomson Reuters I/B/E/S historical earnings database is revision-based. Therefore, a new 'record' is not written into history unless the current estimate changes (referred to as "revised"). In the event that a contributing analyst is confident in the current estimate and does not wish to revise the estimate, a confirmation is requested. Confirmations add integrity to the estimates (a 30-day old estimate, although in-line with all other estimates, is not regarded as confidently as a day-old estimate). Confirmations are easily identifiable in the database in that the announce (effective) date remains unchanged while the confirmation date is updated to the date of the confirmation.

Error-Corrected History

Thomson Reuters has traditionally made error corrections to historical data if it can be substantiated through published research documentation. While there are certain types of estimate data that contain "As published" information (e.g., Surprise values), the majority of the data is error corrected. Policies on historical corrections are defined by data item. In general, historical corrections are made upon request/review and are granted based on: corresponding documentation and if necessary, after the basis is verified.

There are two main types of data items:

- Earnings forecasts and other period-specific data items
- Recommendations or Target Prices

For each of the types, the following factors are taken into consideration when making historical changes:

How long ago did the error occur?

- Within the last six months: Changes are made to the database. History is captured in the recalculated mean figures.
- Prior to the past six months: These changes are made but do not automatically result in recalculated mean figures. This is due to the need to adjust history products and tables, or else detail data will not match mean data. As a result, summary history may not match detail history due to such error corrections.

How was the data received?

- Data can be received via: Notes, PDF Research, or Universe Files.

Types of changes made to historical data:

- Value, Effective Date (and Activation Date for Actuals), Analyst Coverage, Deletion, Addition of Missed Revision

  Historical corrections are made to ensure the highest quality data. Errors are minimized; however it is possible that discrepancies exist due to contributing analysts never sending Thomson Reuters the data originally, or that it was sent incorrectly. As a general rule, corrections are only made, if the contributing analyst can support the value through published research. This policy has been in effect for the treatment of both recent and older history - regardless of whether or not the company reported.

As-Was Summary History

In addition to the traditional 'error-corrected' history offering, Thomson Reuters has recently made a new historical summary-level dataset available, which is unaltered in any way. The As-Was historical daily mean estimates dataset provides daily mean values as they appeared on a particular day; regardless if the underlying detail estimates have since been corrected or not.

Daily Historical Mean is a collection of detail estimates from analysts calculated on a daily basis. The mean is the average of the detail estimates as reported by the analyst at that particular point in time, without making any revisions or corrections to the data once it's published. Quantitative researchers utilize "as was" data to analyze the market impact on the actual day the official record was released. Subscribers of this data set will have the ability to view over 20 financial measures, including 5 types of per share data for US and International companies.

- This powerful data set is *extremely* important to quantitative portfolio managers wishing to see historic data free from modifications due to error corrections.
- As-was history enables clients to see a true snapshot of the *exact* information available to the market at a given point in time - to see the effect that the company's estimates had on market events.

****Note that Thomson Reuters presently only offers summary-level daily as-was history. As-was detail-level estimates history will be a future enhancement to this offering.**

Differences between 'Error-Corrected' and 'As-Was' History

There are certain circumstances when Thomson Reuters needs to adjust or correct a historical detail estimate that has been stored in the database. This happens when brokers go back to Thomson Reuters to correct a previously provided estimate, or when an estimate was missed from an update. In these cases, Thomson Reuters will change the detailed estimate which may or may not cause the mean to change. If the mean changes, it is no longer an "as-was" figure. Instead, the mean becomes "error-corrected" because it is recalculated based on a corrected detail.

Example:

Company ABC has 10 estimates from 10 different brokers. As of 11-01-2006, the mean for the 12-06 quarter is $2.15. One of the brokers covering Company ABC is Broker XYZ who provided Thomson Reuters with an estimate of $2.20 for the same time period.

On November 30, 2006, Broker XYZ told Thomson Reuters that their $2.20 should have been $2.26. Broker XYZ provides documented proof that the estimate that was sent to Thomson Reuters via a feed was incorrect, and that their research reports support that the estimate is actually $2.26. Thomson Reuters will apply the correct value to the detail estimate for the applicable quarter, on the date that the estimate was effective. Because of the change, the mean will change to $2.17. In this scenario, the "as-was" mean is $2.15 and the "error-corrected" mean is $2.17.

In summary, all traditional estimates history products offer 'error-corrected' history in which any time an incorrect value is found, it is then corrected – on either a summary or detail estimate level. Thomson Reuters new 'as-was' history offers historical mean estimates, free of any modification, and shows any given mean estimate value as it appeared in that particular day.

History is also available for Normalized Summary & Detail History (Currency) and is detailed in the Currency section above.

INDUSTRY CLASSIFICATIONS SOURCE / SCHEMA

The sector/industry classification schema for I/B/E/S and Thomson ONE products presently are based upon:

- For U.S. companies follow the S&P scale for sector/industries/groups
- For international companies the MSCI schema is used.

Future products will adopt the new proprietary Thomson Reuters Business Classification schema.

KEY PERFORMANCE INDICATORS

Thomson Reuters offers Key Performance Indicators (KPIs) to quickly identify and retrieve analyst forecast information on key drivers within the retail, restaurant and pharmaceutical industries. These key performance indicators are industry-specific measures that facilitate comparisons among similar peer groups. Consensus and detail forecasts are available for Same Store Sales and Pharmaceutical Sales, including business segment and product breakdowns, enabling efficient comparisons between analysts' expectations on these indicators and your own.

Thomson Reuters collects and displays forecasted and reported industry-specific Key Performance Indicators on products including Thomson ONE Analytics and Thomson ONE Investment Management (under Security -> Estimates -> Detail – Single Period). Estimates data is available on both a detail analyst as well as summary mean level.

Thomson Reuters also offers a Key Performance Indicators (KPI) datafeed collection of current detail and summary level estimates as well as actuals information.

See "Glossary of Estimates Data Measures" section under "Product-Level Measures" for all KPIs collected.

MULTI LISTED SECURITIES

Companies may enlist to trade on multiple exchanges or may have more than one share type trade on a common exchange. The Thomson Reuters estimates database will store forecast information for all listings covered by analysts. The primary listing is referred to as an "S" type Security (Instrument Type: S). This type of security's I/B/E/S ticker will usually reflect the ticker used for trading on the local exchange, such as MSFT for Microsoft Corporation based in the US and traded on the NASDAQ exchange. It is usually the most liquid share class with the highest trading volume.

In addition to the primary listing, companies may also have other listings including:

- Multiple Shares (Instrument Type M)
- Multiple Listings/Inter-listed Securities (Canada Only) (Instrument Type D)
- American Depository Receipts - ADR's (Instrument Type A)
- Combination of all Security Types
- Dual Listed Companies

Multiple Share Classes (Instrument Type M)

Please note: Presently, multiple share listings - indicated by Instrument Type M and having I/B/E/S Tickers with a slash "/" - are not displayed on Thomson Reuters platforms nor included in datafeeds such as I/B/E/S QFS & History.

THOMSON REUTERS

Multiple share classes of a company occur when more than one share class is traded for that company on the same exchange within the same country. The additional shares are referred to as multiple shares of the same equity.
Multiple shares for companies are usually issued because:

- Different levels of voting rights are attached to each share class
- There is a restriction within the market on foreign ownership and a secondary class is created for foreigners
- The company wishes to increase the liquidity of its shares by adding share classes with small nominations
- Other reasons as determined by the company

A multiple share of a company is added to the estimates database as a Multi Share listing (I/B/E/S Type: M). This type of security's I/B/E/S ticker will always be the I/B/E/S ticker of the S type listing, with a slash "/" and a numeric digit suffix. For example, if the ticker for the S type listing of a company is @ALZ, the ticker for the M type listing will be @ALZ/1. If the numeric digit is greater than 9, then a letter is used in place of a numeric, for example: @ALZ/A.

Company Name	Market Symbol	I/B/E/S Ticker	I/B/E/S Type	Exchange Country	Exchange
Royal Dutch Shell	RDSA.NL	@RDN	S	NETHERLANDS	Euronext Amsterdam
Royal Dutch Shell	RDSB.NL	@RDN/1	M	NETHERLANDS	Euronext Amsterdam

Royal Dutch Shell plc has two classes of shares, "A" and "B" shares. "A" shares and "B" shares have identical rights except in relation to the source of dividend income where "A" shares have a Dutch source and "B" shares are intended to have a UK source.

Source: www.unification.shell.com

- Unique tickers are created in the database for each share class – the primary share as type S and the additional share classes as type M (with a slash "/" in the ticker).
- All estimates forecasts (with the exception of price targets, DPS, and recommendations) are stored and displayed under the type S listing regardless of the listing sent by the contributor. Minority data are stored under the share class for which it was received and then copied over to the primary listing with the exception of Price targets, DPS, and recommendations.

Multi-listed Securities/Inter-listed Securities/Dual Listed Securities (Instrument Type D)

A multi-listed/inter-listed security has the same class of shares listed on two different exchanges. Multi-listed securities are an additional listing of any security of the company, but are typically related to the primary listing. In this case, the company's shares are listed on more than one stock exchange in two different geographic locations. Inter-listed securities are those listed on both Toronto Stock Exchange (TSX) and a US exchange, including the NASDAQ, AMEX or NYSE. Each inter-listed security has one CUSIP, is fungible, and can therefore be traded and cleared in either Canada or the US.

A multi-listed/inter-listed security is added to the database as a D Type security under the same issuer name as the primary S type listing. The primary ticker is setup as an S type security and the secondary listing as a D type security.

Example:

Company Name	Market Symbol	I/B/E/S Ticker	I/B/E/S Type	Exchange Country	Exchange	Share Class
Royal Dutch Shell	RDSA.NL	@RDN	S	NETHERLANDS	Euronext Amsterdam	A Shares
Royal Dutch Shell	RDSA.GB	@SHE	D	UNITED KINGDOM	London Stock Exchange	A Shares

Company Name	Market Symbol	I/B/E/S Ticker	I/B/E/S Type	Exchange Country	Exchange
Barrick Gold	RDSA.NL	@RDN	S	NETHERLANDS	Euronext Amsterdam
Barrick Gold	RDSA.GB	@SHE	D	UNITED KINGDOM	London Stock Exchange

Company Name	Market Symbol	I/B/E/S Ticker	I/B/E/S Type	Exchange Country	Exchange
Barrick Gold	ABX.US	ABXF	S	Canada	TSX
Barrick Gold	ABX.CN	ABX3	D	USA	NYSE

THOMSON REUTERS

- Unique tickers are created for each listing -- the listing on the local exchange as type S and the multi-listed/inter-listed as type D.
- Estimates are stored and displayed under the listing provided by the contributing broker.
- Thomson Reuters platforms display both types of securities and feed files include data on both types of securities.

A dual-listed security is a Canadian company that trades on both the US and Canadian stock exchanges. In order to increase granularity of its data, Thomson Reuters uses the following method to capture estimate, recommendation and price target data for Canadian dual-listed companies.

- Thomson Reuters adds a secondary instrument or ticker for Canadian dual-listed companies when estimate data is received for both listings. In order to link the tickers, there are two types of securities: The primary security is denoted as type 'S' and the dual-listed security is denoted as type 'D'.
- Duplicate identifiers (CUSIPS) exist since Canadian companies that trade both in Canada and the US share the same CUSIP, but carry a separate SEDOL for each exchange on which they trade. A CUSIP is a number identifying all stocks and registered bonds – *C*ommittee on *U*niform *S*ecurities *I*dentification *P*rocedures. A SEDOL is a code which identifies a foreign stock that has a CUSIP number but does not trade in the U.S. – *S*tock *E*xchange *D*aily *O*fficial *L*ist.
- Thomson Reuters implements this process in a two-step approach in order to accommodate clients who currently use CUSIP as the identifier to load data. A second dual listed instrument is added and data is captured as received from contributing analysts. An artificial CUSIP is attached, which is the first seven digits of the primary listing and "X" as the last digit eg. 3748593X. The unique SEDOL for each listing is captured in the database in order to maintain correct pricing information.
- The second step requires that data file products be amended in order to adequately support duplicate CUSIPS. Once implemented, Thomson Reuters will continue to maintain the dual listed instruments by properly capturing data and attaching the correct CUSIP for both instruments. The correct digit will replace the artificial "X" once the long-term approach is implemented. At least three months notification will be provided to clients preceding any changes to the ID files.
- Thomson Reuters publishes estimates on whichever security a contributor provides estimates. If an analyst supplies forecasts under both securities then estimates/coverage will be made viewable on both securities. If the analyst supplies forecasts for one security, estimates will be displayed under that particular security and no other.
- Target Price will be the basis for determining which security is covered. For example, if an analyst sends their Target Price under the CAD listing yet supplies US estimates, Thomson Reuters will display coverage under the CAD security. Analyst's have the ability to cover both listings as long as both target prices are supplied. The currency of estimates will have no determining factor on which listing an analyst covers. Dual-listed securities are shown in the exchange opposite of the primary security. For example, if the primary security is listed on the Canadian Exchange, the newly created security would be listed under the US Exchange.

Example of Dual-Listed Company:

Canadian National Railway

Local Tickers:	U.S. – CNI
	Canada – CNR
I/B/E/S Tickers:	U.S. – CNI
	Canada – CN2

Thomson Reuters uses this policy on dual-listed companies due to the request of analysts. Analysts wish to show coverage with specific security. These methods allow analysts to forecast price targets for one or both securities. Having two separate securities increase granularity of data and allow for correct pricing information. It also allows for proper analyst ranking for each security.

American Depository Receipts – ADR's (I/B/E/S Type A)

American Depository Receipts are listings for a foreign traded company on an American exchange. An ADR is a negotiable certificate issued by a U.S. bank representing a specified number of shares (or one share) in a foreign stock that is traded on a U.S. exchange. ADR's are denominated in U.S. dollars, with the underlying security held by a U.S financial institution overseas, and help to reduce administration and duty costs on each transaction that would otherwise be levied. ADR's make it easier for Americans to invest in foreign companies, due to the widespread availability of dollar-denominated price information, lower transaction costs, and timely dividend distributions.

ADR's are treated the same as US companies. If an ADR is covered by one of the Thomson Reuters contributing analysts, estimates are collected as well as actuals, and mean data is created based off the number of analysts included in the mean calculation. ADR's are grouped, however, with US companies, and not by the countries of their local security.

An ADR security is added to the I/B/E/S database as an A type security under the same issuer name as the primary S type listing. The primary ticker is setup as a type S and the secondary listing as a type A security.

Example:

Company Name	Market Symbol	I/B/E/S Ticker	I/B/ES Type	Exchange Country	Exchange	Share Class
Royal Dutch Shell	RDSA.NL	@RDN	S	NETHERLANDS	Euronext Amsterdam	A Shares
Royal Dutch Shell	RDS/A.US	RD	A	USA	NYSE	A Shares

- Unique I/B/ES tickers are created for each listing - the listing on the local exchange as type S and the ADR as type A.
- Estimates are stored and displayed under the listing provided by the contributing broker.
- All platforms display both types of securities and feed files include data on both types of securities.

Combination of All Security Types

Some companies have a combination of different listing types including dual listings, multiple share classes and ADR's, as is the case for Royal Dutch Shell PLC.

Example:

Company Name	Market Symbol	I/B/E/S Tickers	I/B/E/S Type	Exchange Country	Exchange	Share Class
Royal Dutch Shell	RDSA.NL	@RDN	S	NETHERLANDS	Euronext Amsterdam	A Shares
Royal Dutch Shell	RDSB.NL	@RDN/1	M	NETHERLANDS	Euronext Amsterdam	B Shares
Royal Dutch Shell	RDSA.GB	@SHE	D	UNITED KINGDOM	London Stock Exchange	A Shares
Royal Dutch Shell	RDSB.GB	@SHE/1	M	UNITED KINGDOM	London Stock Exchange	B Shares
Royal Dutch Shell	RDS/A.US	RD	A	USA	NYSE	A Shares
Royal Dutch Shell	RDS/B.US	RD/1	M	USA	NYSE	B Shares

Thomson Reuters publishes estimates on whichever security a contributor provides estimates. If an analyst supplies forecasts under both securities then estimates/coverage will be made viewable on both securities. If the analyst supplies forecasts for one security, estimates will be displayed under that particular security and no other.

- Target Price will be the basis for determining which security is covered. For example, if an analyst sends their Target Price under the CAD listing yet supplies US estimates, Thomson Reuters will display coverage under the CAD security. Analyst's have the ability to cover both listings as long as both target prices are supplied. The currency of estimates will have no determining factor on which listing an analyst covers. Dual-listed securities are shown in the exchange opposite of the primary security. For example, if the primary security is listed on the Canadian Exchange, the newly created security would be listed under the US Exchange.

PARENT / CONSOLIDATED INDICATOR

Indicates whether the estimates of a company are carried (by Thomson Reuters) on a parent or consolidated basis. The way a company appears on the database is based on the majority of the earnings estimates received. Contributors are free to provide either parent or consolidated estimates for any given company. Using sales estimates as an example, consolidated sales estimates would be under SAL, whereas sales for parent company would be under SALPAR. The primary basis (either P or C) is determined by whichever is the majority basis.

THOMSON REUTERS

Consolidated Companies

Companies are classified as consolidated when the earnings of the investee companies where the parent holds a 20% voting stake or more are combined with the earnings of the parent company, after elimination of inter-company transactions.

Parent Companies

Companies are classified as parent when only the earnings of the reporting entity, including dividends, interest, royalties, etc. received from its investee companies, are presented as net income.

Companies Without Subsidiaries

Companies without subsidiaries are classified as consolidated by default since a great majority of the markets adhere to the consolidated basis.

Consolidated / Parent Companies

If companies are carried in two-basis (Consolidated and Parent) and use a different calculation, a review and shifting of the affected measures are necessary to ensure that the majority and minority of broker submissions are stored in the right primary measures (Primary Parent/ Primary Consolidated) and secondary measures (Secondary Parent/ Secondary Consolidated). Switching the primary basis from secondary and vice versa is imperative when there is a significant drop or increase in either broker submission.

Shifting Company Indicators

The reason for the need to shift is that there are two main data products that are dependent on current collection:

- History- The detail history product only includes primary basis. Due to constraints it is imperative that the primary basis includes the majority of contribution.
- Global Aggregates- This product also offers history. If EPS history for primary basis is deleted/ removed/ relabeled calculations that includes these companies will be affected.

The switch from consolidated primary to parent primary or vice versa should be based on two main factors:

- Change in reporting standards/ actual availability - Availability of actual data for the basis identified as primary. When company does not have subsidiaries and no earnings to consolidate.
- Change in broker submission- when there is a shift in majority of basis brokers is sending their data.

When a significant number of brokers are shifted to a different basis, the primary measure is shifted to the basis where the majority of the brokers are sending. The basis where the minority of the brokers are sending will be the new secondary measure. All measures for the same basis will be shifted all together.

When equal contribution is submitted for both bases, the deciding factor should be the availability of the actuals for that company/market based on proposed/ reviewed and approved by the accounting board.

When equal contributions are submitted for both bases and there is an actuals available for both bases as well, the company basis should remain as of the day of the review. When companies have minimal (1 or 2 contributor in the P/C status) difference in contribution and majority have shifted to a different basis, the current measures remain until a significant number of contributors have shifted. Significant number is considered as 60% if company has fewer than 8 estimates & 40% if it is has 9 estimates up.

PERIODICITY

Periodicity is the frequency for which a company reports their full financial results. A company will have either a quarterly (QTR) periodicity, a semi annual (SAN) periodicity, or an annual (ANN) periodicity once it is established with the database and data is collected.

Quarterly (QTR) periodicity is used when:

- Company reports full financial results quarterly;

- Company reports full financial results semi annually, and contributors are making quarterly EPS or FFO estimates; and;
- Company reports full financial results annually and there are no contributors making interim estimates.

Semi-Annual (SAN) periodicity is used when:

- Company reports full financial results semi annually, and contributors are not making quarterly EPS or FFO estimates. There are cases where contributors will supply quarterly sales estimates for companies that only report full financials semi annually. These sales estimates should not be used to determine the periodicity since it is not a shifting measure; and
- Company reports full financial results semi annually, and there are no contributors making interim estimates.

Annual (ANN) periodicity is used when:

- Company reports full financial results every 12 months, and a period year consists of one annual.
- A company's periodicity should be set to the most frequent time interval based on one of the following:
- The company report; or
- EPS or FFO estimates periodicity supplied by contributors

Please note that quarterly periodicity is the most frequent interval used as the default periodicity when setting up new companies.

PRELIMINARY ESTIMATES

When Thomson Reuters receives a contributor's estimate, it goes through an extensive and thorough verification process prior to delivery to all estimates products to ensure accuracy and consistency. This value-added quality control process ensures estimates are of the highest quality and estimates are delivered to products in the quickest time possible, however there are times where this added level of process may affect the timeliness of estimates.

As a solution for the most time-sensitive clients, Preliminary Estimates are available which combine real-time estimate availability, with an automated quality screening process. A Preliminary Estimate bypasses the manual portion of Thomson Reuters value-added quality control checks and verification tests – and is only subjected to limited automated verification tests. This data is then available in true real-time, enabling clients to view a contributor's updated forecasts prior to the Thomson Reuters full verification, filtering and footnoting process. The majority of Preliminary Estimates will be followed by a 'fully-verified' estimate, which are subjected to all of Thomson Reuters quality control checks.

- Preliminary Estimates enable true real-time delivery to clients.
- Preliminary Estimates are useful to any customers making investment decisions based on estimate revisions and related time sensitive activity.
- Preliminary estimates are currently being offered via the First Call Datalink feed, as well as Thomson ONE Analytics and Thomson ONE Investment Management platforms.
- First Call Datalink offers Preliminary Estimates for the following data measures: EPS, Sales, Cash Flow per Share, Recommendations and Price Target.
- Thomson ONE Analytics and Thomson ONE Investment Management offer Preliminary Estimates for all 26 data measures.

Please note that Preliminary Estimates are available in real-time after fielded receipt of estimate values from analysts (either once automated feeds/files are received from brokers, or once Thomson Reuters Market Specialists extract estimate values from PDF research documents.

PRICE FORECASTS

In addition to publically traded companies, Thomson Reuters also collects forecasts on the price levels of commodities, as well as both bottom-up and top-down price forecasts on select indices.

Commodity Price Forecasts

Commodities are something that are relatively easily traded, that can be physically delivered, and that can be stored for a reasonable period of time. A common characteristic of commodities is that their prices are determined on the basis of an active market. Examples of commodities include metals, minerals, and energy sources such as crude oil, natural gas,

THOMSON REUTERS

aluminum, gold, diamonds, or silver. Sales and purchases of commodities are usually carried out under future contracts on exchanges, which standardize both the quantity and minimum quality of the commodity being traded.

Commodity price forecasts are collected by Thomson Reuters if available from contributing analysts. Unique I/B/E/S tickers are created for each commodity with sell-side analyst estimates coverage and are set up as a Type "O" Instrument type. For a complete listing of all available commodity price forecasts, please reference the document *"Thomson Reuters Top-Down Index & Commodity Price Forecasts"*.

Actuals

Commodity price actuals are entered within 15 days of the end of the period by using the calculated average price of the preceding three (3) months period. Please note that this method is also used by the contributing analysts, who take the average closing price of the quarter to determine actuals, not the closing price at the end of the quarter.

Estimates

Commodity price forecasts are based off spot prices and are entered using the same majority basis policy as estimates on companies. These estimates are sourced from the same sell-side analysts covering companies and related industries.

Index Price Forecasts

Thomson Reuters collects and calculates price forecasts for a handful of US stock indices, most notably including the S&P500 and Dow Jones Industrial Average (DJIA). Unique I/B/E/S tickers are created for each index with sell-side analyst estimates coverage and are set up as a Type "I" Instrument type. For a complete listing of all available index price forecasts, please reference the document *"Thomson Reuters Top-Down Index & Commodity Price Forecasts"*.

Two types of index price forecasts are available on Thomson Reuters; top-down, which are an average of market strategists' forecasts, and bottom-up, which are aggregations of all analyst mean forecasts for each individual company in an index.

Top-Down Estimates

Index price forecasts are based off index prices and are entered using the same majority basis policy as estimates on individual companies. These detail estimates are sourced from sell side industry analysts, as well as market strategists who forecast based upon macroeconomic conditions, rather than individual company performance. All of these individual estimates are then averaged to create a mean (consensus) top down forecast.

Bottom-Up Estimates

In addition to Thomson Reuters collecting top-down forecasts from sell-side contributors, bottom-up forecasts are calculated as well. These forecasts are sourced from aggregating all of the individual mean estimates for each individual company in an index, and then weighted by market cap. The explicit bottom-up index forecasts calculation used by Thomson Reuters is as follows:

Avg_eps = spi * total_cons_shares / total_price_shares

Where:

Avg_eps =	bottom-up index estimate displayed on products
spi =	price index value
total_cons_shares =	consensus eps * shares of each company of the Index
total_price_share =	price * shares of each company of the index

Actuals

The current policy for updating actuals for index estimates is to enter the bottom up calculated figure two quarters after the end of the period. Bottom-up estimates and actuals are calculated on a calendarized basis, in order to account for different fiscal year ends for companies and allow for comparison of companies regardless of fiscal period. The calendar quarter end is taken along with the month before and the month after to create a quarter number that allows companies with different fiscal periods to be compared against each other.

Actuals Entry Schedule:

Quarter	Period Ending	Enter Actual Value on
Q1	March 31	July 1
Q2	June 30	October 1
Q3	September 30	January 1
Q4	December 31	April 1

Calendarization Methodology:

Quarter	Period Ending
Q1	February, March, April
Q2	May, June July
Q3	August, September, October
Q4	November, December, January (of next calendar year)

PRIORITIZATION

Estimates and recommendations are researched and reviewed by Thomson Reuters Market Specialists to insure accuracy – prior to becoming available on products. Every revision is subject to a stringent quality control process – both before and after the data is available on products. If the accuracy or accounting basis cannot be verified by the data source alone, Thomson Reuters Market Specialists will further research the affected estimates/recommendations, by contacting the contributing analysts directly for clarification. It is however Thomson Reuters goal to deliver accurate and reliable estimate revisions as timely as possible.

During peak times such as earning seasons, the added revision volume can sometimes cause slight delays. Thomson Reuters uses a rolling 'priority scheme' which gives higher priority to market movers, index constituents, higher market cap companies, companies in the news/reporting etc. – to ensure that estimate revisions for these types of companies are the first to be updated.

All of the following would be considered as higher priorities when updating estimates; surprising earnings news, pre-announcements, reported earnings, S&P companies, market capitalization, major merger announcements/ completions and post-market prior day events (e.g., companies in the news to which the market has yet to react). Index Constituents tend to be considered market movers and therefore given priority over lesser-followed companies. For that reason, the mechanism is in place to highlight an index as a priority grouping.

Please note that Preliminary Estimates are available in real-time after fielded receipt of estimate values from analysts – prior to the manual verification process. See Preliminary Estimates section for more details.

REASONS FOR CONTACT WITH CONTRIBUTING ANALYSTS

All phone calls between Thomson Reuters Market Specialists and Contributors/IR Representatives are logged in a phone call database.

Cases that would typically trigger Thomson Reuters to contact a contributor include but are not limited to:

- Quarterly estimates within the published research document do not add to the annual provided (indicating use of non-majority prior period actual).
- Quarterly or annual estimates received from a contributor (either via research or feed) which fail quality control tests and validations for accuracy, such as standard deviations, decimalization errors, etc.
- An accounting basis issue is identified within a contributor's estimate or reported actual – contributor contacted and communicated what the 'majority' basis is using.
- A company issues guidance, and the contributor either does not update/confirm their estimate or it is outside of the guidance range.
- An estimate fails the Thomson Reuters Freshness Policy and a contributor is contacted to confirm/revise their estimates.
- A company announces a merger/acquisition/spinoff – a contributor is contacted for their post-event estimate.

THOMSON REUTERS

- A contributor's estimates are not updated after a company reports their quarterly/annual results.
- Pre-split estimates are provided in research, after a company has gone through a stock dividend or split of their stock.
- A company goes through a FYE change and the contributor sends numbers on the old FYE.

RECOMMENDATIONS

Recommendation Mapping: Thomson Reuters I/B/E/S 1-5 Scale

The Thomson Reuters I/B/E/S recommendation scale is as follows:

1 - Strong Buy
2 – Buy
3 – Hold
4 – Underperform
5 – Sell

Each contributor determines how their individual recommendation scale maps to the Thomson Reuters I/B/E/S 5-point scale. Every firm, no matter if they have a 3-point scale or a dual-tiered system, must map their scale to the normalized 1-5 scale utilized by Thomson Reuters. The only stipulation being that the mapping requested must allow for negative to negative ratings, positive to positive ratings and neutral to neutral ratings when mapping to Thomson Reuters I/B/E/S 1-5 scale. A contributor using a 3-point scale of BUY, HOLD, SELL would not be allowed to have a mapping of 1,2,3 on the 1-5 Thomson Reuters Scale. Contributors are made aware that the 1-5 value will be calculated to create a mean and displayed across Thomson Reuters products.

Please note that while contributors may have elaborate multi-tier recommendation scales, including both company and industry/sector ratings, all points in their scale must map back to the standardized Thomson Reuters I/B/E/S scale is 1-5. In cases of broker scales being greater than 5 points, multiple points in a broker's scale may map back to a single point in the Thomson Reuters I/B/E/S scale.

Recommendation Mapping: Impact on Products

Clients viewing the Recommendations data measure, depending upon the product, can view analyst recommendations in multiple versions:

- Contributor Text format – the actual text provided by the contributor
- Normalized Text format – the corresponding text on Thomson Reuters normalized scale
- Normalized Code format – the corresponding code on Thomson Reuters normalized scale

Contributor Text format is the exact recommendation language used by that specific contributing firm. Normalized Text and Code make the Contributor Text more consistent, by mapping the Contributor Text to Thomson Reuters standard 1-5 recommendation scale. It is the Normalized Codes which are used to calculate the Thomson Reuters Mean Recommendation.

Recommendation Scale Changes

If a contributor changes their recommendation scale, stops must be applied to the database to prevent false revisions, followed directly by new recommendations applied on the same day. When recommendation scale changes occur, Thomson Reuters Market Specialists work closely with the contributor to outline the implications, and make decisions on how the change should be represented, based on the guidelines Thomson Reuters uses in mapping contributor scales to the normalized scale.

Note: Recommendation scale change requests received from contributors will be processed on a go-forward basis

Recommendation Drops

If a contributor drops coverage of a company, a stop is applied to the recommendation field. Additionally, if a contributor is "restricted" on the stock or has suspended their recommendation, a stop would be applied to the recommendation field.

RESTATEMENT POLICY (ACTUALS)

Thomson Reuters actuals restatement policy addresses the needs of two distinct sets of end users: those who prefer the actual data as it was initially reported and those who wish to view the company as it is constituted today.

- Thomson Reuters can restate actuals for any available measures; however the ones most commonly restated are EPS, Sales and FFO.
- Thomson Reuters will restate the quarterly figures for the current fiscal year, as well as the prior year's actuals data to provide comparability. Thomson Reuters will not restate actual data for more than one year back.
- All other actuals data will be left as originally entered, to allow historical examination.
- In all cases, footnotes will be entered to explain the basis of the modified figures.
- Once a restatement has taken place, any existing estimates or new estimate submissions must use the restated actual data: this ensures a proper apples-to-apples comparison among contributing analysts. If a contributor is not using the restated figure, a Thomson Reuters Market Specialist will contact the analyst to adjust to the restated basis, or will have their estimates footnoted and excluded from the mean for the fiscal year in question.

Examples of events that would require restatement include:

- Changes in the accounting basis
- Classification of certain operations as discontinued
- Sales and acquisitions of business lines

Example of company with restated actuals:

Integrated Circuit Systems (ticker ICST)

Restated EPS Actual:	Q105 = 0.24R
Accompanying Footnote:	11-Nov-04 SEP04Q Restated from 0.23 upward for accounting change

Thomson Reuters will only restate actuals after a company has officially made the restatement, and can be documented via a press release, or by confirmation of all the contributing analysts.

SHARE CLASS

Default share class is determined by the majority of estimates submitted. Policies differ slightly for the US and International companies.

U.S.

1. Determined by majority of coverage.
2. If there is not a majority of coverage, then defer to liquidity.
3. If liquidity is comparable then defer to the share class with the most voting rights.

International

1. Determined by majority of coverage.
2. If there is not a majority of coverage, then defer to the share class with voting rights.

Only recommendations and target prices are affected by share class; all other estimates are generally available under the primary share class.

Shares Outstanding Data

Number of Shares Outstanding (NOSH)
Current number of shares outstanding (NOSH) data is provided as a supplemental data item in I/B/E/S datafeeds as well as on Thomson ONE (Security->Overview->Snapshot). This data provided is based on the NOSH for the specific security (SEDOL-specific), and not on the consolidated/company level.

Shares Outstanding Used in Per-Share Estimates
The shares outstanding data, for per-share data measures, which is utilized in individual analyst's detail estimates, and subsequently the summary level mean data, are all consolidated/company-specific data (it is not share class specific, like the NOSH data displayed on products is).

- *The above is only for per-share measures. Exclusions would be Dividend Per Share and Price Targets, which would be based upon NOSH for the particular share class.*

Example
To illustrate, here is an example using Viacom:

- NOSH data would display 549.503m for VIAB, and VIAB/1 has 57.364m number of shares outstanding; each security showing security-specific shares outstanding.
- Analyst research reports, and subsequently estimates data, would show 607m number of shares outstanding; showing consolidated/company level shares outstanding.

STOP, FILTER AND DELETION SCENARIOS

***Stop* - Results in a contributing analyst's estimates no longer being displayed on products.**

- The contributing analyst has dropped coverage.
- The contributing analyst is "restricted" on the stock.
- Estimate/recommendation has not been updated (confirmed or revised) for 180 days or more.
- Recommendation / Target Price under review

***Filter* - Contributing analyst's estimates are still displayed on products but are footnoted and excluded from the mean calculation.**

- Estimate is on a different accounting basis than the majority of contributing analysts.
- Estimate has not been confirmed or revised at the issuance of a company's earnings guidance and it is either outside of the guidance range or >5% of a single-point guidance value; applying only to the specific measure and period issued.
- Estimate is not on the majority basis pertaining to a corporate action or the estimate has not been updated to reflect a corporate action after the effective date.
- Quarterly estimates revised without a corresponding adjustment to the annual estimate (all other period estimates for the same year are filtered).
- Annual estimate revised without a corresponding adjustment to the quarterly estimates (all quarterly estimates for the same year are filtered).
- A Thomson Reuters Market Specialist has requested data verification and no response was received for more than 48 hours.
- Estimate is under review by the contributing analyst.
- Estimate has not been updated (confirmed or revised) for 105 days or more.
- After an actual is reported, an estimate is excluded from the mean if it is not or confirmed within 10 business days of a prior-period reported actual.
- Estimate is updated for post-Rights Issue prior to the ex-date.

***Deletion* - Estimate is removed from the database and history. The previous estimate becomes the current estimate.**

- Incorrect estimate was entered into the database (only if verified by published research).

TAX RATES

A quarterly estimate is only considered to be on a different basis with respect to taxes if some analysts are taxing the estimates and others are not. For example, if an analyst is not taxing their estimates and the other analyst is using a tax rate of 30%, those two estimates are on a different basis and one of them needs to be excluded from the mean calculation. On the other hand, if one analyst is using a tax rate of 20% and the other is using a tax rate of 33%, and there are no other basis issues, those estimates are on the same basis and should both be included in the mean.

This holds true for an annual estimate as long as the analyst is using the same tax rate for the actuals that we are using. If the analyst is using a different tax rate for a reported period (different actual), then the annual estimate should be filtered. Any future quarters should remain unfiltered if they do not violate the quarterly rule above.

TREATMENT OF SMALL ESTIMATES REVISIONS

Thomson Reuters accepts data from contributors to varying degrees of precision. Most contributors provide estimates to 2 or 3 decimal places. The following are scenarios under which small estimates revisions would be treated:

Second Decimal Place

- An estimate revision that is less than 0.01, which *does not result* in a new value after rounding to the **second decimal place**, is treated as a confirmation of the existing estimate (i.e., it is not recorded in the Thomson Reuters I/B/E/S collection database as a revision and is not fed to products as a revision).
- An estimate revision that is less than 0.01 which *does result* in a new value after rounding to the **second decimal place** is treated as a revision and is fed to products as a revision.

Third Decimal Place *(in effect since June 15, 2009)*

- All estimates revisions that impact the **third decimal place** after rounding will now be recorded and fed to products as a revision, for select currencies, in order to provide additional estimates granularity for markets that are regularly impacted by very small revisions:
  - Australian Dollar (AUD)
  - Japanese Yen (JPY)
  - Malaysian Ringgit (MYR)
  - New Zealand Dollar (NZD)
  - Singapore Dollar (SGD)
  - South African Rand (ZAR)
  - South Korean Won (KRW)

Scenario 1: New estimate differs from the current estimate by less than 0.01, but *does not impact* the second decimal place after rounding.

Example 1 – Not Impacting Second Decimal Place

	ANALYST'S ESTIMATE		3+ DECIMAL PLACE PRODUCTS (I/B/E/S QFS, I/B/E/S HISTORY, REUTERS KNOWLEDGE, 3000 XTRA)			2 DECIMAL PLACE PRODUCTS (THOMSON ONE, FIRST CALL DATALINK)		
	Estimate	Revision Date	Estimate	Revision Date	Confirmation Date	Estimate	Revision Date	Confirmation Date
Existing	0.241	05-May-2009	0.241	05-May-2009	05-May-2009	0.24	05-May-2009	05-May-2009
New	0.244	03-Jun-2009	0.241	05-May-2009	03-Jun-2009	0.24	05-May-2009	03-Jun-2009

In Example 1, the new estimate is treated as a confirmation on all products since the change does not impact the second decimal place after rounding. No subsequent revision dates change, but confirmation date is updated.

Example 2 – Impacting Third Decimal Place - Select Currencies

	ANALYST'S ESTIMATE		3+ DECIMAL PLACE PRODUCTS (I/B/E/S QFS, I/B/E/S HISTORY, REUTERS KNOWLEDGE, 3000 XTRA)			2 DECIMAL PLACE PRODUCTS (THOMSON ONE, FIRST CALL DATALINK)		
	Estimate	Revision Date	Estimate	Revision Date	Confirmation Date	Estimate	Revision Date	Confirmation Date
Existing	0.241	05-May-2009	0.241	05-May-2009	05-May-2009	0.24	05-May-2009	05-May-2009
New	0.244	03-Jun-2009	0.244	03-Jun-2009	03-Jun-2009	0.24	03-Jun-2009	03-Jun-2009

In Example 2, the new estimate is treated as a revision on products displaying 3 decimal places since it is for one of the select currencies and it impacts the third decimal place after rounding. On products with 2 decimal places it appears as the same value since the second decimal place is not impacted, however the revision and confirmation dates are updated.

Scenario 2: new estimate differs from the current estimate by less than 0.01, but does impact the second decimal place after rounding.

Example 3 – Impacting Second Decimal Place

	ANALYST'S ESTIMATE		3+ DECIMAL PLACE PRODUCTS (I/B/E/S QFS, I/B/E/S HISTORY, REUTERS KNOWLEDGE, 3000 XTRA)			2 DECIMAL PLACE PRODUCTS (THOMSON ONE, FIRST CALL DATALINK)		
	Estimate	Revision Date	Estimate	Revision Date	Confirmation Date	Estimate	Revision Date	Confirmation Date
Existing	0.244	05-May-2009	0.244	05-May-2009	05-May-2009	0.24	05-May-2009	05-May-2009
New	0.246	03-Jun-2009	0.246	03-Jun-2009	03-Jun-2009	0.25	03-Jun-2009	03-Jun-2009

In Example 3, the new estimate is treated as a revision on all products since it impacts the second decimal place after rounding.

GLOSSARY OF ESTIMATES DATA MEASURES

Product-Level Measures

Key Performance Indicator Description	Relevant Industries	Measure Code	Measure Abbreviation
Pharmaceutical Sales	Drug Manufacturers	SAL	PS
Same Store Sales	Retailers, Restaurants, Lodging	SSS	SS

Pharmaceutical Sales

Pharmaceutical Sales represents the revenue associated with individual pharmaceutical drug unit products.

- Thomson Reuters collects reported company results and forecasted sales estimates on a quarterly and annual basis for pharmaceutical companies globally.
- Estimates data available on both a detail analyst as well as summary mean level.

- Thomson Reuters links these drugs on multiple levels depending on the business relationship, chemical ingredients and purpose associated with each - allowing not only specific forecast data for each separate drug but also aggregate sales of generic ingredients and instances where global revenues are shared as a joint venture between companies.

Same Store Sales

Same Store Sales represents a percentage sales growth for retail stores and restaurants that have been open for more than one year. Same Store Sales allows investors to decipher what portion of sales growth is due to true retail growth and what portion is due to new store openings.

- Thomson Reuters collects reported company results and sales growth forecasts on a monthly, quarterly and annual basis for North American companies.
- Estimates available on a store line as well as consolidated basis, where available.
- Estimates data available on both a detail analyst as well as summary mean level.
- Companies followed include discount retailers, department stores, specialty retailers, casual dining, quick serve restaurants and more.

Company-Level Measures

Data Measure Description	Primary Consolidated Code	Secondary Consolidated Code	Primary Parent Code	Secondary Parent Code
Book Value Per Share	BPS	SBP	BPSPAR	SBPPAR
Capital Expenditure	CPX	SPX	CPXPAR	SPXPAR
Cash Flow Per Share	CPS	SCP	CPSPAR	SCPPAR
Dividend Per Share	DPS			
Earnings Before Interest & Taxes (EBIT)	EBI	SBI	EBIPAR	SBIPAR
Earnings Before Interest, Taxes, Depreciation & Amortization (EBITDA)	EBT	SBT	EBTPAR	SBTPAR
Earnings Per Share	EPS	SEP	EPSPAR	SEPPAR
Earnings per Share - Alternate	EPX			
Earnings per Share - Before Goodwill	EBG	SBG	EBGPAR	SBGPAR
Earnings per Share - Cash	CSH	SCS	CSHPAR	SCSPAR
Earnings per Share - Fully Reported / GAAP	GPS	SGP	GPSPAR	SGPPAR
EBITDA Per Share	EBS	SEB	EBSPAR	SEBPAR
Enterprise Value	ENT	SNT	ENTPAR	SNTPAR
Funds From Operations Per Share	FFO	SFO	FFOPAR	SFOPAR
Gross Profit Margin	GRM	SGM	GRMPAR	SGMPAR
Long Term Growth Rate (%)	LTG			
Net Asset Value	NAV	SAV	NAVPAR	SAVPAR
Net Debt	NDT	SND	NDTPAR	SNDPAR
Net Income	NET	SNI	NETPAR	SNIPAR
Operating Profit	OPR	SOP	OPRPAR	SOPPAR
Pre-tax Profit	PRE	SPR	PREPAR	SPRPAR
Price Target	PTG			
Recommendation	REC			
Return on Assets (%)	ROA	SOA	ROAPAR	SOAPAR
Return on Equity (%)	ROE	SOE	ROEPAR	SOEPAR
Revenue	SAL	SSA	SALPAR	SSAPAR

**While EPS, Revenue, Price Target and Recommendations are the most popular measures contributed, analysts are free to contribute forecasts for any or all of the collected data metrics specified above. Thomson Reuters doesn't require any minimums in terms of collected data measures, and is willing to accept all metrics a broker provides.*

**For companies followed on both a parent and consolidated basis (see the Parent/Consolidated Indicator section), both Primary and Secondary data measures are available. The markets where two-basis measures are usually available include India, Japan, South Korea, Taiwan, and Thailand.*

Book Value per Share (BPS)

A company's common stock equity as it appears on a balance sheet equal to total assets minus liabilities, preferred stock, and intangible assets such as goodwill, divided by the weighted average number of total shares outstanding for the year. This is how much the company would have left over in assets per share after all debts are paid, if it went out of business immediately. Thomson Reuters provides both expected and actual BPS data (where available).

Capital Expenditure (CPX)

Funds used by a company to acquire or upgrade physical assets such as property, industrial buildings, or equipment or the amount used during a particular period to acquire or improve long term assets such as property, plant, or equipment. Thomson Reuters provides both expected and actual CPX data (where available).

Cash Flow per Share (CPS)

Cash Flow per Share is a corporation's cash flow from operations, before investing and financing activities, divided by the weighted average number of common shares outstanding for the year. Investing includes the sale or purchase of land, factories, buildings etc.

- Financing includes dividend payments, loan proceeds and sale of stock. Thomson Reuters provides both expected and actual CPS data (where available).
- Interest payments are an operating activity.
- Thomson Reuters CPS is a company's Operating Cash Flow. The basic formula is Operating Cash flow less maintenance capital = Distributable Cash flow per unit.
- CPS is generally calculated after-tax.
- Thomson Reuters does not have DCFPU (Distributable Cash Flow per Unit) as a measure. This is something to consider as an industry specific measure as well as payout ratio. If the company does not provide operating cash flow, Thomson Reuters will collect the DCFPU estimate and place it in the CPS filtered with "A" for accounting difference.

Dividend per Share (DPS)

DPS are a corporation's common stock dividends on an annualized basis, divided by the weighted average number of common shares outstanding for the year. In the US dividend per share is calculated before withholding taxes (though for some non-US companies DPS is calculated after withholding taxes). Thomson Reuters provides both expected and actual DPS data (where available).

- Thomson Reuters DPS is equivalent to Cash Distribution (not the same as Distributable Cash Flow per Unit.)
- For DPS estimates a "0" is a valid estimate, indicating no expected dividend payment for a company. The absence of any estimate or a "stopped" estimate indicates that a contributor does not have any DPS estimate.

Earnings per Share (EPS)

Valuation earnings per share, defined as the EPS that the contributing analyst considers to be that with which to value a security. This figure may include or exclude certain items depending on the contributing analyst's specific model. Estimates that are not on the majority basis for a given security are displayed on certain Thomson Reuters products but filtered from the mean calculation. Thomson Reuters provides both expected and actual EPS data where available.

Earnings per Share - Alternate (EPX)

Alternate EPS is a corporation's net income from continuing operations, divided by the weighted average number of shares outstanding. This measure tracks the estimates of contributing analysts who wish to forecast EPS on the non-majority basis. This alternate basis is not included in the mean calculation; it is filtered from the main EPS data measure. This data measure therefore, will not have corresponding Summary-Level (mean), nor actuals data.

Earnings per Share - Before Goodwill (EBG)

EBG measures a company's per share earnings before the amortization of goodwill. In some countries (France, for example) goodwill is treated as a part of ordinary income for companies and the amortized component of goodwill is added back to yield earnings before goodwill amortization. EBG is a corporation's net income from continuing operations before goodwill amortization divided by the weighted average number of shares outstanding. Thomson Reuters provides both expected and actual EBG data (where available).

- Due to the implementation of International Financial Reporting Standards (IFRS) in various European countries, goodwill will no longer be amortized but instead written off as an impairment charge and will be treated as an exceptional item. This change eliminates the necessity for a separate EBG measure for companies residing in those countries. In such markets, Thomson Reuters will only collect and display EPS and GPS (valuation EPS and fully-reported EPS).

Earnings per Share - Cash (CSH)

Cash Earnings Per Share is a company's net income, plus depreciation, amortization of goodwill, intangibles, and prepaid assets (non-cash items); divided by weighted average number of shares outstanding. Thomson Reuters provides both expected and actual CSH data (where available).

Earnings per Share – Fully Reported / GAAP (GPS)

Statutory or reported earnings per share, defined as net profit (on continuous activities) divided by the weighted average number of shares outstanding during the period. Where a company carries exceptional items or goodwill amortization, this measure is post-exceptional, post-goodwill. Thomson Reuters provides both expected and actual GPS data (where available).
In North America this figure is referred to as GAAP Earnings per Share and is calculated according to Generally Accepted Accounting Principles (GAAP), which is reported in SEC filings. The mean estimate for the GPS data measure will only reflect the strict adaptation of GAAP basis estimates. Estimates from contributors on an adjusted GAAP basis will be displayed but footnoted and filtered from the mean, even if the adjusted basis is the majority. A-type footnotes will include as much information as possible regarding the difference in accounting basis from the strict GAAP basis. This policy may result in the majority of estimates being filtered under GPS if the majority basis is an adjusted GAAP basis.

In countries that have adopted International Financial Reporting Standards (IFRS) this figure will include all items according to IFRS rules.

EBIT / Earnings Before Interest & Taxes (EBI)

EBIT represents the earnings of a company before interest expense and income taxes paid. As such, EBIT is a gauge of corporate earnings before any debt servicing to creditors (including bondholders) and the payment of corporate taxes. It is calculated in general form by taking the pretax corporate income of a company, adding back interest expense on debt, and subtracting any interest capitalized. Thomson Reuters provides both expected and actual EBIT data (where available).

- Displayed in whole number terms (millions).
- In certain European and Asian markets, EBIT is calculated as total sales and subtracting total costs and operating expenses. In these cases EBIT will be similar to Operating Profit.

EBITDA / Earnings Before Interest, Taxes, Depreciation & Amortization (EBT)

EBITDA gauges the raw earnings power of a company before debt servicing, corporate taxes, and any allowances made for depreciation and amortization costs the company faces. It is calculated in general form by taking the pretax corporate income of a company, adding back any depreciation and amortization costs charged, plus any interest expense on debt (subtracting any capitalized interest). Thomson Reuters provides both expected and actual EBITDA data (where available).

- Displayed in whole number terms (millions).
- In the United Kingdom, the general market standard is to include royalties as part of gross revenue, net of royalty tax. This tax portion would be included as part of the royalties, and would therefore be deducted before EBITDA, rather than as part of the income taxes lower down the income statement.

EBITDA per Share (EBS)

EBITDA per share represents EBITDA divided by the weighted average number of shares outstanding. Thomson Reuters provides both expected and actual EBS data (where available).

Enterprise Value (ENT)

Enterprise Value is calculated as market capitalization plus debt, minority interest and preferred shares, minus total cash and cash equivalents. Cash equivalents are defined as an item on the balance sheet that reports the value of a company's assets that can be converted into cash immediately. Examples of cash and equivalents are bank accounts, marketable securities and Treasury bills. An Enterprise Value actual is calculated using the closing price at the end of the fiscal period. Thomson Reuters provides both expected and actual ENT data (where available).

Funds from Operations per Share (FFO)

A measure used by real estate and other investment trusts to define the cash flow from trust operations. It is earnings with depreciation and amortization added back. A similar term increasingly used is Funds Available for Distribution (FAD), which is FFO less capital investments in trust property and the amortization of mortgages. Thomson Reuters provides both expected and actual FFO data (where available).

Gross Margin (Gross Profit Margin) (GRM)

A company's total sales revenue minus cost of goods sold, divided by the total sales revenue, expressed as a percentage. Thomson Reuters provides both expected and actual GRM data (where available).

Long Term Growth Rate (%) (LTG)

The long term growth rate represents an expected annual increase in operating earnings over the company's next full business cycle. These forecasts refer to a period of between three and five years, and are expressed as a percentage.

Long term growth rate forecasts are received directly from contributing analysts; they are not calculated by Thomson Reuters. While different analysts apply different methodologies, the Long Term Growth Forecast generally represents an expected annual increase in operating earnings over the company's next full business cycle. In general, these forecasts refer to a period of between three to five years. Due to the variance in methodologies for Long Term Growth calculations, Thomson Reuters recommends (and uses as its default display) the median value for Long Term Growth Forecast as opposed to the mean value. The median value (defined as the middle value in a defined set of values) is less affected by outlier forecasts.

Net Asset Value (NAV)

Net Asset Value is the total book value of a company's securities. It is calculated in general form by taking the total assets of a company and subtracting the value of the company's intangible assets (goodwill, patents, etc.) minus current and long-term liabilities. NAV is helpful in determining under-priced equities by indicating the ultimate value of a company's securities in the event of their liquidation. Thomson Reuters provides both expected and actual NAV data (where available).

- Displayed in whole number terms (millions).
- As NAV is not a measure companies generally report in filings or press releases, Thomson Reuters calculates NAV actual data as total shareholders equity including minority share or total assets minus total liabilities.

Net Debt (NDT)

Net Debt is calculated as short and long term interest bearing debt minus cash (and equivalents). Thomson Reuters provides both expected and actual NDT data (where available).

Please note the examples below:

Rule: If debt is greater than cash, the value collected will be a positive number in the database.
From the balance sheet.

Cash and Equivalents	$175
Short and Long Term Debt	$400
Net Debt =	$400 – 175
NDT =	$225

Rule: If debt is less than cash then the value collected will be a negative number in the database.
From the balance sheet.

Cash and Equivalents	$300
Short and Long Term Debt	$250
Net Debt =	$250 – 300
NDT =	($50)

Net Income (NET)

Net income is defined as a corporation's after-tax income. This item varies significantly from market to market as regards the inclusion or exclusion of non-recurring items. In most markets, non-recurring items are backed out of net income and this measure is restricted to income from continuing operations only (also referred to as normalized income). Some markets (Japan, for example) apply reported net income, including any and all extraordinary items. Recent accounting changes in still other markets (particularly Southeast Asia) have resulted in a reclassification of extraordinary versus exceptional items, bringing many formerly extraneous items above the net income line. Thomson Reuters provides both expected and actual NET data (where available).

Operating Profit (OPR)

Operating Profit is the difference between a company's revenues and its costs and expenditures arising directly out of a company's regular operations. Operating Profit is calculated before any deductions in income owing to non-operating activities (generally such items as interest expense, corporate tax payments, material gains or losses arising from changes in accounting policy, and the like) and excludes any income derived from outside the firm's regular activities. Thomson Reuters provides both expected and actual OPR data (where available).

- Displayed in whole number terms (millions).
- In certain European and Asian markets, EBIT is calculated as total sales and subtracting total costs and operating expenses. In these cases EBIT will be similar to Operating Profit.

Pre-Tax Profit (PRE)

Pre-tax profit is a company's net income before tax expense. Where applicable, extraordinary items and non-recurring charges are subtracted from net income. Thomson Reuters provides both expected and actual PRE data (where available).

- In Japan, companies compliant with Japan Accounting Standards use Recurring Profit.

Price Target (PTG)

Price target is the projected price level forecasted by the analyst within a specific time horizon. Note that while detail-level data can be collected for various time horizons, Thomson Reuters summary-level mean data is only calculated for targets with 12-month time horizons.

Recommendation (REC)

The recommendation value reflects the contributing analyst's rating for a particular company.

Return on Assets (ROA)

Return on Assets is a profitability ratio and as such gauges the return on investment of a company. Specifically, ROA measures a company's operating efficiency regardless of its financial structure (in particular, without regard to the degree of leverage a company uses) and is calculated by dividing a company's net income prior to financing costs by total assets. Thomson Reuters provides both expected and actual ROA data (where available).

- **Displayed as a percentage.**

Return on Assets is calculated as follows:

$$\text{ROA (Return on Assets)} = \frac{\text{Net Income}}{\text{AverageTotalAssets}}$$

Return on Equity (ROE)

Return on Equity is another profitability ratio, which gauges return on investment by measuring how effectually the company is employing stockholder money. ROE is calculated by dividing a company's net income by total equity of common shares. Unlike ROA, ROE does consider the degree to which a company uses leveraging, as interest expense paid to creditors is generally deducted from earnings to arrive at Net Income. Thomson Reuters provides both expected and actual ROE data (where available).

THOMSON REUTERS

- **Displayed as a percentage.**

Return on Equity is calculated as follows:

$$\text{ROE(Return on Total Equity)} = \frac{\text{Net Income}}{\text{Average Total Equity}}$$

Revenue (Sales) (SAL)

The Revenue measure is a corporation's net revenue, generally derived from core business activities. For non-financial companies, the calculation of net revenue (or net turnover) in most markets generally involves subtracting transportation and related operational costs from gross revenue/sales. Revenue recognition practices vary significantly from market to market, though generally the recording of revenue is based upon sales invoices issued (or anticipated for forecast purposes) during the accounting period.

For banks, revenue is generally defined as net interest income plus net non-interest income. Net interest income is defined as interest income minus interest expenses. Net interest income components generally include net interest earned on loans, reserve deposits and deposits with other banks, and net interest earned from inter-bank money market operations (IMMO) and marketable securities. Net non-interest income components generally include net income from fees and commissions, net gains from capital market and foreign exchange operations, and net income earned from participations.

For insurance companies, revenue is generally defined as net technical income plus net financial income. Net technical income is generally defined as technical income minus technical expenses. Technical income components generally include income from premiums and commissions received, re-insurer's share of claims paid, transferred net technical reserves, and re-insurer's share of technical reserves. Net financial income is generally defined as financial income minus financial expenses. Net financial income components generally include net interest income, net dividend income, and net foreign exchange gains. Thomson Reuters provides both expected and actual SAL data (where available).

Expected Returns on Stocks and Bonds

Investors must moderate their expectations.

Antti Ilmanen

ANTTI ILMANEN
is a managing director in European Fixed Income Strategy at Citigroup, London, U.K.
antti.ilmanen@citigroup.com

The equity-bond risk premium—the long-run expected return advantage of stocks over government bonds—is one of the biggest questions in financial markets. The extent of the premium is widely debated, but it is reasonably clear that it declined in the last quarter of the 20th century, to partly rebound in the first years of the 21st century.

Our review provides a road map to the complex literature on the topic. We explain the key drivers of the risk premium and varying assumptions about them, letting investors themselves assess the long-run prospects for stocks versus bonds. Long-term government bond yields are known, while prospective equity returns are inherently less transparent and thus more open to question.

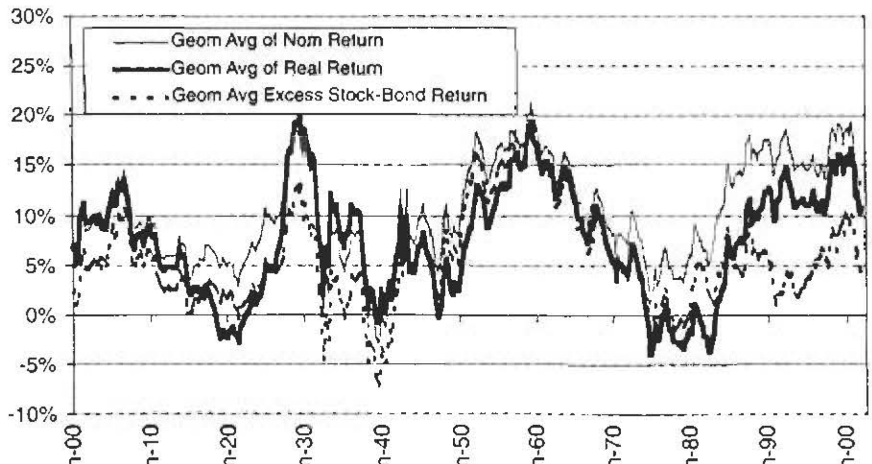
There is an ongoing shift in opinion about expected returns. Long-term equity premiums have traditionally been predicted from historical average asset performance assuming a constant risk premium, but today they are increasingly predicted with the help of dividend discount models, assuming time-varying expected returns.

We first review the historical average returns of major asset classes and explain why these are misleading guides for the future. Essentially, the double-digit returns of the 20th century were due to equities starting cheap and getting richer over time. Many investors extrapolated this past performance and expected (at least) as high future returns. Investors thus missed, first, the fact that a part of realized returns was unexpected windfalls from rising equity valuation multiples, and, second, that when starting from high valuation levels it is not reasonable to

EXHIBIT 1
Road Map to Equity Risk Premiums—Alternative Means for Assessing Levels

	Historical Ex Post Excess Returns	Surveys	Ex Ante Models and Market Data
Means of Assessing the Equity-Bond Risk Premium	Historical average is a popular proxy for the ex ante premium – but likely to be misleading.	Investor and expert surveys can provide direct estimates of prevailing expected returns/premiums.	Current financial market prices (simple valuation ratios or DDM-based measures) can give most objective estimates of feasible ex ante equity-bond risk premiums.
Problems/ Debated Issues	Time-variation in required returns and systematic selection and other biases have boosted valuations over time, and have exaggerated realized excess equity returns compared with ex ante expected premiums.	Limited survey histories and questions of survey representativeness. Surveys may tell more about hoped-for expected returns than about objective required premiums due to irrational biases such as extrapolation.	Assumptions needed for DDM inputs, notably the trend earnings growth rate, make even these models' outputs subjective. Range of views on this growth rate (plus debates on relevant stock and bond yields) => range of premium estimates.

EXHIBIT 2
Moving Average of 10-Year Stock Market Performance 1900–2001



Sources: Ibbotson Associates, Arnott (private correspondence), Shiller website, and Schroder Salomon Smith Barney.

expect as high returns as in the past.

The painful lessons of the recent bear market have made investors more aware of forward-looking expected return measures; the starting price matters. Since market yields give good proxies for the expected returns of long-term bonds, the question of the ex ante equity-bond premium boils down to the ex ante equity return. The dividend discount model (DDM) shows that in the absence of predictable valuation changes (often a good base case), feasible long-run equity return is the sum of dividend yield and a long-run earnings growth rate.[1]

We stress the distinction between two types of expected returns—objectively feasible long-run returns, and subjective return expectations—as well as the balance between them. Objectively high feasible returns are bullish for equities, while excessive subjective investor expectations are bearish, because high hopes make future disappointment more likely.

Neither expected return can be directly observed, but we attempt to estimate them by analyzing historical returns, investor surveys, and market valuation indicators (*see Exhibit 1*). Surveys provide direct estimates of changing return expectations, but they may reflect hoped-for returns as much as required returns.[2]

As of the time of writing in mid-2002, long-term bond yields are 4%–5%, and the DDM suggests feasible long-run equity returns between 5% and 8% (depending on input assumptions). There may still be an imbalance between the objective return prospects and subjective expectations that we put between 8% and 10%. The gap has narrowed significantly from the year 2000 when feasible returns were even lower (due to higher valuation multiples), while subjective return expectations were well into double-digits.[3]

PITFALLS OF BACKWARD-LOOKING RETURNS

The 20th century was the century of equities. Dimson, Marsh, and Staunton [2002] review the 1900-2000 asset returns in 16 countries, and conclude that in all markets stocks handily outperformed bonds and cash. We extend the data to include the 2001 experience, and discuss primarily the U.S. market history.

Even after large losses in the last two years, U.S. equities' average real returns over the 1900-2001 period are 6.5%, with excess return over long-term government bonds of 4.8 percentage points.[4] Looking at just the 1950-1999 period, stocks did even better, outperforming bonds by 7.7 percentage points per year. For comparison, the excess return of equities over bonds was much slimmer (0.5 percentage point) in the 19th century (1802-1899), while the realized average real equity return was similar (6.2%) (see Siegel [1998] and Arnott and Bernstein [2002]).[5]

Exhibit 2 plots the ten-year average compound returns of stocks since 1900—comparing nominal returns, real returns, and excess returns over bonds. In some studies, equity performance is expressed in raw returns, while in others the inflation rate or long-term bond return (or short-term bill return) is subtracted from it. Another distinction is between compound (geometric) average returns and simple (arithmetic) average returns.

Given that the United States has been the world's most successful economy of the past two centuries, it is not surprising that real equity returns have been somewhat lower in most other markets. For example, the average real equity returns for the other G-5 markets over the 1900-2001 period range between 3.4% (Germany) and 5.6% (the United Kingdom). Hyperinflation experiences make excess stock returns versus government bonds harder to gauge.

Did Realized Returns Exaggerate Expected Returns?

A consensus is emerging that the high long-term returns on equities, relative to bonds, are unlikely to persist. The 20th century was favorable to stocks and unfavorable to bonds. Improved valuations boosted ex post equity returns, while rising inflation expectations and real yields hurt bonds. Thus, the realized return gap almost surely exaggerates the expected return gap investors actually required (in the past, let alone after the decline in required returns).

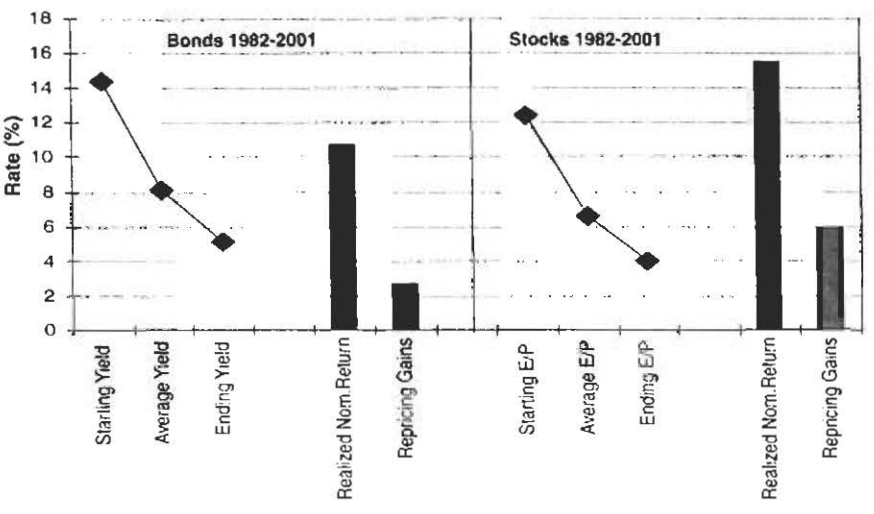
- Various systematic biases make it likely that the publicized realized equity market returns from historical studies exceed the returns that were anticipated—notably survivorship bias, easy data bias, and the so-called peso problem (see Dimson, Marsh, and Staunton [2002] and Fama and French [2002], among others).
- Survivorship bias raises the odds that we examine countries that have had good capital market performance (say, the current G-5 as opposed to Russia, Austria-Hungary, India, Turkey, or Argentina).
- Easy data bias makes it likely that we start samples after unusual events (war, hyperinflation, market closure), which often means that assets are cheap at the start of the period and that no comparable turmoil occurs again during the period.
- The peso problem literature recognizes that past U.S. market pricing was influenced by what could have happened but did not.[6] With hindsight we know that the United States and its market economy survived two world wars, the Cold War, and the Great Depression, and did not suffer the hyperinflation, invasion, or other calamities of many other countries. This was not a forgone conclusion at the time, so it is little wonder that realized equity returns have been boosted by a repricing effect.

Despite these arguments, it is common to use historical excess returns as a proxy for the ex ante risk premium; indeed this is the approach taken in most investment textbooks. Historical average returns equal expected returns, however, only if expected returns are constant, and if unexpected returns from trendwise valuation changes do not distort the within-sample results. Such valuation changes can materially impact average realized returns even over long sample periods—and indeed they have done so in the 20th century. Thus the crucial distinction between realized (ex post) average excess returns and expected (ex ante) risk premiums.

Bond investors understand better than equity investors the folly of extrapolating expected returns from past average returns drawn from a time when valuation levels have trended up or down. A rally—high realized returns—caused by falling discount rates will reduce future yields (feasible expected returns), rather than raise them.

The example in Exhibit 3 shows that between 1982 and 2001 ten-year Treasury yields averaged 8.1%, but the realized annual return was 10.7% because the downtrend in yields (from 14.4% to 5.1%) added almost 3 percentage points of annual capital gains to the yield income. Using the 10.7% realized annual return or even the 8.1% average yield as an expected return proxy makes little sense

EXHIBIT 3
Bond and Stock Market Repricing Gains Due to Falling Discount Rates Between 1982 and 2001



Source: Schroder Salomon Smith Barney.

now that the yield is 5%. The transparency of market yields prevents bondholders from harboring excessive return expectations after a long bull market.

Exhibit 3 shows that the revaluation effect was even greater for equities. The earnings-to-price (E/P) ratio fell from 12.4% to 4.0% in 20 years; that is, the market paid 3.1 times more for a given amount of dollar earnings at the end of 2001 than at the end of 1981. This repricing explains almost 6 percentage points of the S&P 500's 15.5% realized annual return (11.8% real). Again the realized average return clearly exceeds the forward-looking return that was feasible in the 1980s, let alone now. Unfortunately, most equity investors may have focused more on historical returns than on forward-looking returns.

Repricing: Valuation-Neutral Sample or Adjusted Realized Returns

If required returns vary over time, past average returns may be poor predictors of future returns. We try to recover the past average expected returns using two approaches—by selecting a sample period when valuation changes were minimal, and by adjusting realized returns for the estimated repricing impact.

We first focus on a relatively valuation-neutral subperiod—1960-2001. Realized average returns can be dominated by unexpected capital gains/losses even over long sample periods if markets undergo significant valuation changes. Indeed, starting from 1900 or 1950, D/P and E/P ratios have fallen dramatically, while bond yields have risen. These within-sample changes are much smaller between 1960 and 2001, which means that future expected return extrapolations from this subperiod should be less distorted.

The 3.3 percentage point excess return in the United States falls short of the 4.8 percentage points for the 1900-2001 period. During the same period, the excess returns in Germany and Japan (1.1 and 0.0 percentage points) are even slimmer as real equity returns have been lower and real bond returns higher than in the U.S.

These average returns conceal significant time variation in market performance. Besides the equity correction of 2000-2002, these numbers show that equities can underperform long bonds over a period as long as a decade (Germany in the 1970s, Japan in the 1990s). In Japan, the realized excess return over the past 30 years is now negative. Because such a sustained underperformance did not take place in the United States in the last century, many investors took the idea of equities' long-run superiority too far, and believed that equities will always beat bonds over a 20- to 30-year horizon.

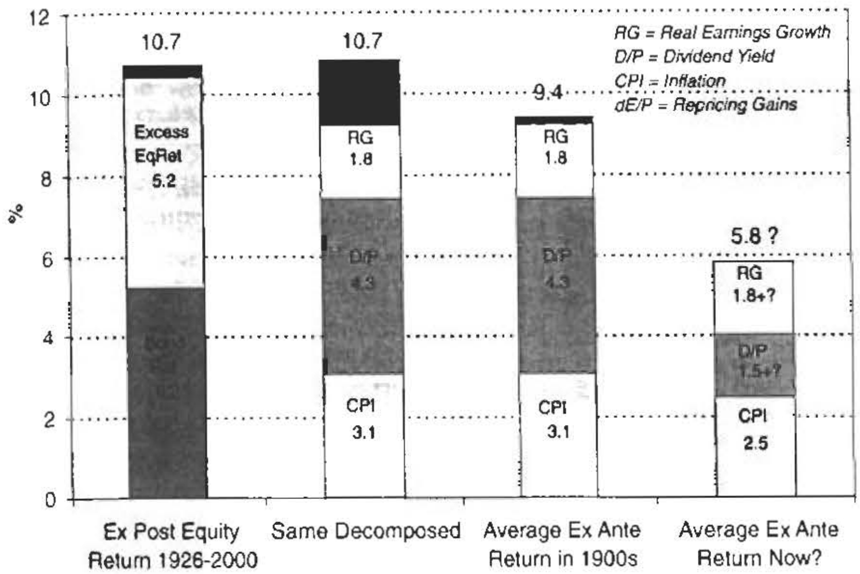
By now it is clear that all statements about the probability of stocks beating bonds were distorted by the favorable sample period, and that the outperformance odds are much slimmer now, given the narrower equity-bond premium.

Alternatively, we can pick any sample period and adjust the returns for unexpected capital gains. Several recent studies take this approach, notably Dimson, Marsh, and Staunton [2002], Fama and French [2002], and Ibbotson and Chen [2002]. Each study uses a slightly different way to remove the impact of unexpected capital gains to recover the typical expected equity risk premium over the sample period. All three studies find (adjusted) expected equity-bond risk premium near 4 percentage points in the United States, averaged over very long histories.

Moving Toward Forward-Looking Expected Returns

Exhibit 4 shows how Ibbotson and Chen [2002] decompose the realized 75-year average compound stock

EXHIBIT 4
Decomposition of 1926–2000 Equity Market Returns



Sources: Ibbotson-Chen [2002], Schroder Salomon Smith Barney.

market return of 10.7% into demanded or supplied parts. The total return is split either into:

- A sum of demanded returns on the assumption that sample averages capture required returns well (5.2% nominal Treasury bond return + 5.2% ex post equity risk premium + small interaction/reinvestment terms), or into:
- A sum of supplied returns (3.1% inflation + 4.3% dividend yield + 1.8% real earnings growth rate + 1.3% repricing effect + small interaction/reinvestment terms).

The third column in Exhibit 4 removes from the supplied returns the unexpected repricing effect (1.3%, the annualized impact of the within-sample change in E/P ratio). The study concludes that investors required a nominal equity market return of 9.4% between 1926 and 2000, on average.

Analysis of *past average* levels can be a misleading guide for the future when current dividend yields and inflation expectations are much lower than the sample average. It misses the point that if expected returns and valuations vary over time, historical averages incorporate limited information about medium-term market prospects. Using strictly the dividend yield and inflation expectations from mid-2002 together with the historical real earnings growth rate, in the spirit of the DDM, the prospective long-term equity market return is below 6%. The implicit equity-bond premium is about 1 percentage point.

The question marks in the last column in Exhibit 4 are related to debates that we review below.

The ongoing shift from constant risk premiums and rational investors to time-varying risk premiums and partly irrational investors means that forward-looking (ex ante) returns are gaining ground over historical (ex post) returns. This change is moderating experts' and investors' perceptions of prospective long-run equity returns and equity-bond premiums, given that the fourth column in Exhibit 4 (ex ante return) is much lower than the first column (ex post return).

Survey Evidence on Subjective Return Expectations

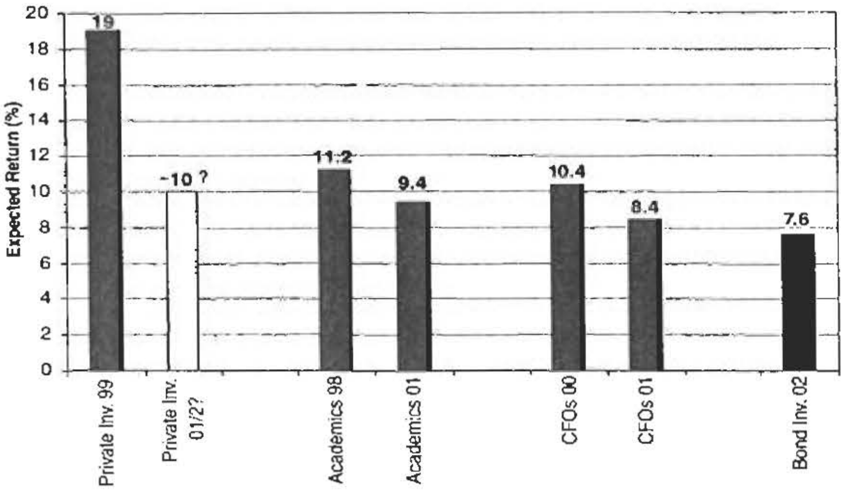
There is a dichotomy between *objectively feasible* return prospects and less rational *subjective* expectations. To provide direct evidence on subjective return expectations, Exhibit 5 summarizes survey views on nominal long-term equity returns from various sources.[7]

Private investors' subjective return expectations were especially high in the late 1990s. Poterba [2001] quotes a broad Gallup poll from 1999 when the consensus of private investors expected 19% annual returns over the long term. Presumably these were deemed moderate expectations after five years of 20%–40% annual returns.

No follow-up surveys tell us how much these excessive expectations have fallen, but we would guess to around 10%. Consensus forecasts in one-year-ahead surveys seem to center around 10% (but dropped in summer 2002 below 8%), while many U.S. pension funds continue to budget well over 10% annual equity returns.

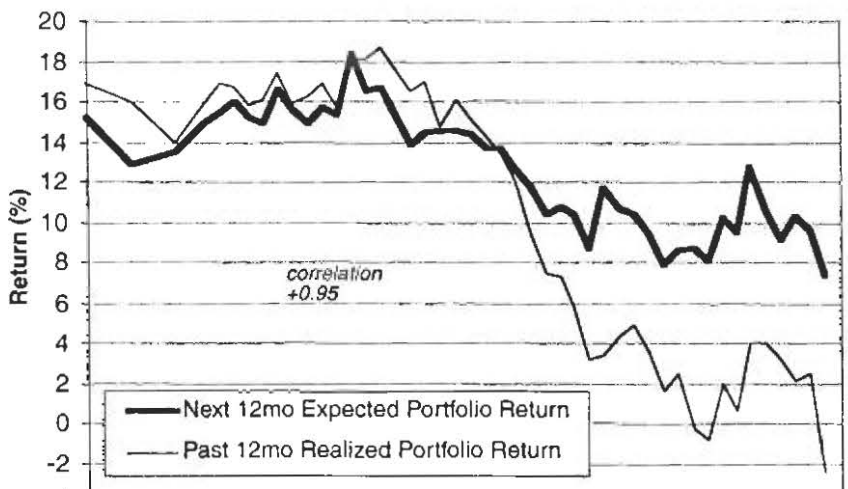
Two surveys of different U.S. experts—finance and economics professors by Welch [2000, 2001] and CFOs and treasurers by Graham and Harvey [2001]—imply long-run equity returns of 8%–9% and stock-bond risk premium estimates of 3.5 to 4.5 percentage points. The equity return forecast in the CFO survey has stabilized at around 8.2% to 8.3% in 2002.

EXHIBIT 5
Survey Forecasts of Long-Term Nominal Expected Returns of U.S. Equities



Sources: Graham and Harvey [2001], Poterba [2001], Welch [2001], and Schroder Salomon Smith Barney.

EXHIBIT 6
Individual Investor Extrapolative Return Expectations—June 1998–August 2002



Sources: Gallup/UBS Paine Webber Survey of Index of Investor Optimism.

Our own survey in April 2002 of global bond investors comes up with the most cautious views on future equity market returns. The mean forecast for next-decade average equity market return is 7.6% for the United States. Compared with bond yields of around 5.2%, these forecasts imply a stock-bond risk premium of 2.4 percentage points.

Are these survey-based risk premium estimates useful proxies for the equity risk premium that the market requires? One can always question how representative any survey is of market views. More important, because of behavioral biases, survey-based expected returns may tell us more about hoped-for returns than about required returns.

Private investor surveys appear especially prone to extrapolation (high hopes after high returns); witness the striking 95% correlation between the past year's returns and next year's expected returns in Exhibit 6. Even the expert surveys are not free from this bias, as consensus views of future risk premiums have edged lower amid poor market performance.[8]

Given the tendency of investors to extrapolate from past returns, the danger of exaggerated expectations and the scope for subsequent disappointment were especially high after two decades of double-digit returns. To quote Dimson, Marsh, and Staunton [2002, p. 4]:

> The most fundamental question of all is: Do investors realize that returns are likely to revert to more normal levels, or do current valuations embody exaggerated expectations based on imperfect understanding of history?

Survey data indicate that investor expectations have corrected lower in the past two years—but it is not possible to say whether the adjustment has gone far enough.

How High Should the Equity-Bond Risk Premium Be?

There is also a normative question about the appropriate size of the equity risk premium, but academic theories provide limited guidance. In the context of the capital asset pricing model, the required market risk premium should reflect the *price* of risk (market risk aversion) and the *amount* of risk (stock market volatility). Other asset pricing models relate the required risk premium to asset return covariances with consumption; intuitively, the risk premium should be high for assets that perform poorly in bad states of the world when losses hurt most (economic downturns with high marginal utility and low consumption).

Given the low observed correlations between equity returns and consumption data, popular utility functions need extremely high risk aversion coefficients to justify the high observed equity risk premium; see Mehra and Prescott [1985]. Academics have proposed various solutions to this equity premium puzzle—alternative utility functions and market imperfections—but there is little agreement on the topic.

While the academic consensus has been shifting from constant risk premiums to time-varying expected returns, opinions vary about the source of the variation: rational time variation in required risk premiums or irrational fluctuations in market sentiment. We believe that both matter.

Because stock prices can be viewed as discounted values of expected future cash flows, it is an accounting identity that higher stock prices and realized returns reflect higher earnings growth expectations or lower required returns. Both factors likely contributed to the run-up in stock prices in the 1990s. The growth optimism was based on a range of factors from real evidence on higher productivity to irrational hopes about the Internet and the new economy (see Asness [2000a] and Shiller [2000]).

Here we focus on a host of possible reasons for the 1990s fall in required equity returns:

- Declines in riskless Treasury yields that contribute to equity discount rates.
- Changing risk—Output volatility and earnings volatility have fallen during past decades; recessions are less frequent (as well as shorter and shallower); monetary and fiscal policies are more stable; improved regulatory and legal infrastructures arguably make transactions safer; and world wars and the Cold War are history.
- Changing risk aversion—Consumer surveys reveal a fall in perceived risk aversion that may be attributed to wealth-dependent risk tolerance or demographic developments. Lower risk and risk aversion are intertwined in many arguments.
  —Higher realized volatility and market losses may remind investors of their risk aversion. Many authors contrast investor caution about equities after the depression of the 1930s with the market-dips-are-buying-opportunities mentality in the 1990s. The optimistic spin is that investors learned in the 1980s-1990s about the consistency of equity long-horizon outperformance, and that this learning enhanced investors' risk tolerance and thereby slimmed equities' required return cushion over less risky assets.
  —Lower trading costs, better market access, greater global diversification opportunities, and negative stock-bond correlations enabled investors to reduce the systematic risk in their portfolios, which in turn raised investors' willingness to take risks.

Some of these factors have reversed since 2000. Although macroeconomic volatility remains low by historical standards, financial market volatility has been extremely high, and perceived risks have risen since September 11, 2001, and various corporate scandals. Sharp falls in share prices certainly have reminded investors of the innate riskiness in equity investing and brought investors closer to their subsistence levels, thereby raising the risk aversion level. If investors perceived, say, a 2 percentage point equity-bond premium sufficient three years ago, we suspect they would now require twice as high compensation for bearing equity risks. Finally, the latest declines in government bond yields appear related to bonds' safe-haven characteristics and should not help reduce the equity discount rates.[9]

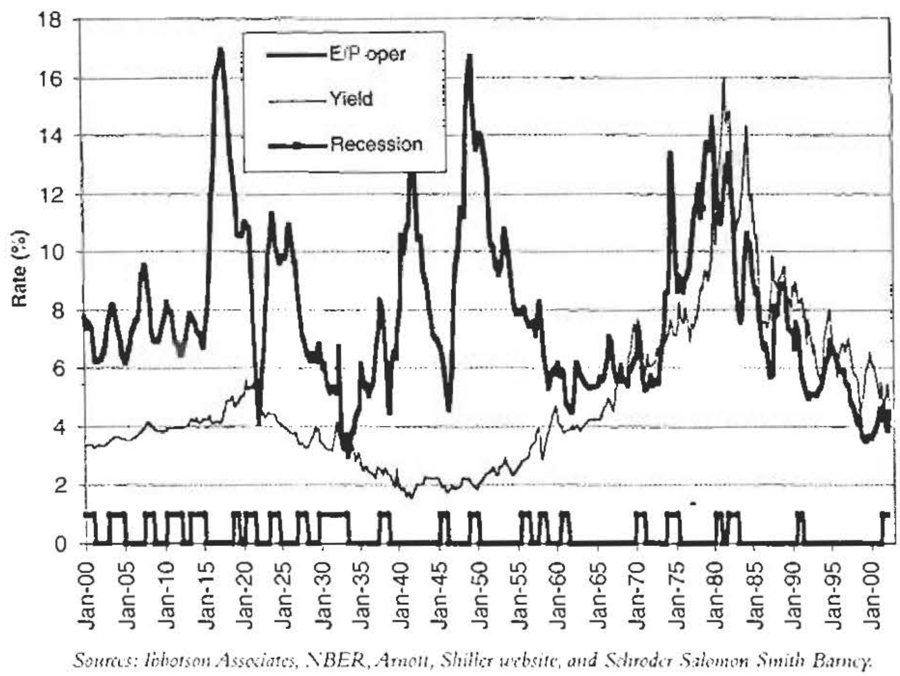
SIMPLE VALUATION RATIOS AS EQUITY-BOND PREMIUM PROXIES

A stock market's price-earnings (P/E) ratio is the most popular pure-equity valuation indicator. Similarly, the ratio of government bond yield (Y) over earnings yield (E/P) is the most popular relative valuation measure for the two major asset classes and thus a shorthand for the equity-bond premium. (Sometimes the earnings yield spread is used instead of the yield ratio, but the broad patterns tend to be similar.)

Lower Bond Yields Explain Lower Earnings Yields

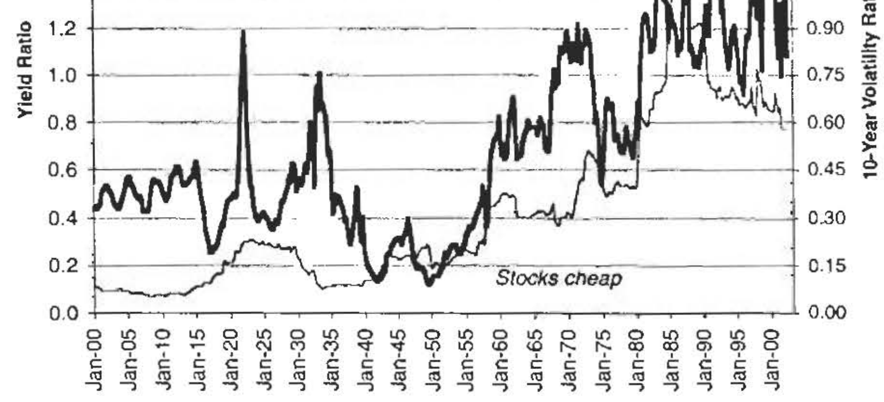
Exhibit 7 shows the history of earnings yield and the ten-year government bond yield for over one century. We focus on the earnings yield rather than its reciprocal

EXHIBIT 7
Earnings Yield of S&P 500 (Operating Earnings) and 10-Year Treasury Yield, 1900–June 2002



Sources: Ibbotson Associates, NBER, Arnott, Shiller website, and Schroder Salomon Smith Barney.

EXHIBIT 8
Bond-Earnings Yield Ratio and Bond-Stock Volatility Ratio 1900–June 2002



Sources: Ibbotson Associates, NBER, Arnott, Shiller website, and Schroder Salomon Smith Barney.

(P/E), because the former is a rate of return measure, akin to a bond yield. Unless otherwise stated, our earnings yield refers to the trailing one-year operating earnings per share of the S&P 500 index and its predecessors.[10]

The broad picture is that the earnings yield has ranged between 4% and 16%, but has been near historical lows for the past few years. Bond yields traded between 2% and 6% for the first 70 years, then hit a 16% peak in the early 1980s, followed by a decline to 4%-5% in 2002. Bond yields traded systematically below earnings yields for most of the century, but traded above them for the last two decades. The measures at the foot of the graph show the timing of the increasingly rare official recessions.

While earnings yields and bond yields were hardly related until 1960, since then they have shared common uptrends and downtrends. Exhibit 8 plots the yield ratio of the Treasury yield over the earnings yield. This ratio is high when stocks are expensive versus bonds, in the sense that bond yields exceed earnings yields.

For the last 20 years, this ratio has been neatly mean-reverting, providing good relative-value signals for asset allocation trades between stock and bond markets. Over this period, we can say that lower bond yields explain lower earnings yields (higher equity market valuations). This is not surprising, because bonds are the main competing asset class for equities, and the bond yield constitutes the riskless part of equities' discount rate.

But what are we to make of the long-run trends in the yield ratio? If we cannot explain them, we may deem the last 40 years' close relation between stock and bond yields as spurious, perhaps related to the broad rises and falls in inflation.

Lower Relative Risk of Stocks versus Bonds Explains the Long-Run Puzzle

The yield ratio series was relatively trendless in the first half of the 20th century but clearly upward-trending in the second, signaling relative richening of stocks versus bonds. Asness [2000b] proposes an appealing explanation for the long upward trend in the yield ratio: The relative risk of bonds versus stocks has grown over time.

The thin line in Exhibit 8 shows the relative return volatility of ten-year government bonds and the stock market index, measured by ten-year moving standard deviations. In the first half of the century, stock market returns were about seven times as volatile as bond returns. By the 1980s, relative volatilities were virtually equal—although subsequent disinflation has reduced bond volatility to about half of stock market volatility.

The trend increase in the volatility ratio reflects an increase in bond volatility, particularly in the 1970s-1980s, and a decline in stock volatility since the 1930s. The related underlying macroeconomic trends are:

- Growing inflation uncertainty associated with the persistent rise in inflation until the early 1980s.
- More stable real growth, as evidenced by lower volatilities in real output and earnings growth rates and by less frequent, shorter, and shallower recessions.[11]

Changing relative risk between asset classes is a structural change that undermines the usefulness of valuation signals like the yield ratio. This ratio will serve well as a mean-reverting signal within any one regime, but it typically gives a wrong value signal when a structural change occurs.

How to watch out for those structural changes? One guidepost is the relative importance of long-run inflation and growth risks.

- If central bank credibility and other arguments, for example, convince people of future inflation stability, and thus of relatively higher real growth risks, relative bond-stock volatility may again shift lower. Such a change should favor bonds and perhaps move the yield ratio back below unity in the medium term. Exhibit 8 shows a reversal in the volatility ratio in the past 15 years but not yet any trend reversal in the yield ratio. (In third quarter 2002, the yield ratio did fall below unity, however.)
- As a more current example, we think that in the world after September 11, 2001, with heightened security concerns and policy uncertainties, both growth and inflation risks have increased. It is less clear which has increased more, making the impact on the yield ratio debatable.
- Deflation would arguably reduce the required bond risk premium and raise the required equity risk premium. Thus, incipient deflation should systematically reduce the yield ratio.

Drivers of Earnings Yields

Since stock prices reflect the discounted values of expected future cash flows, it is an accounting identity that low earnings yields (high P/E ratios) reflect some combination of low discount rates and/or high expected earnings growth rates.

Like many others, we find that various growth indicators are only loosely related to earnings yield fluctuations and that P/E ratios have only a modest ability to predict subsequent earnings growth. Discount rate effects may reflect the riskless yield component or the required equity-bond risk premium. The sensitivity of earnings yields to nominal bond yields can be traced back to expected inflation rates or required real bond yields. Historical analysis suggests that earnings yields have been more closely related to inflation than to any other series, including nominal or real bond yields.

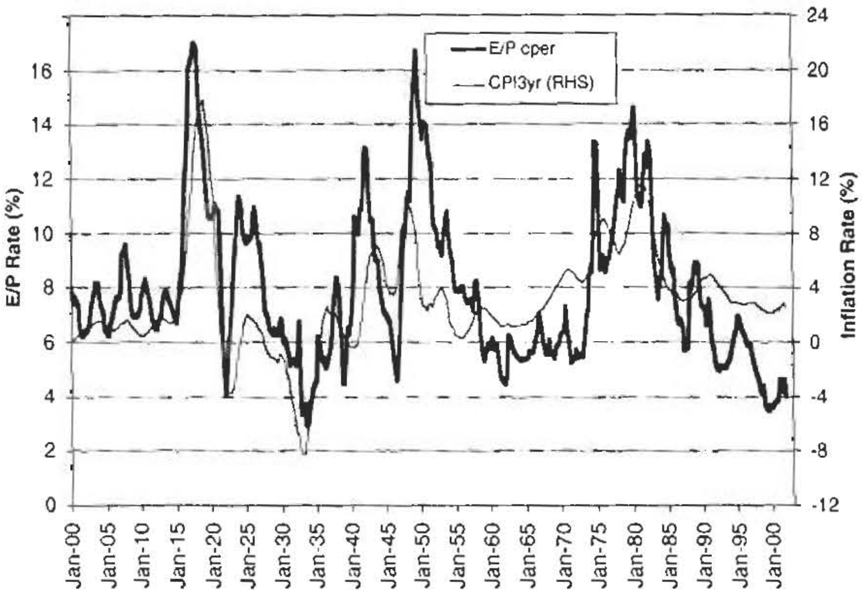
Exhibit 9 depicts the relation between U.S. earnings yields and the previous three years' average inflation. There is a similarly close relationship in other countries, including Japan.[12]

A high correlation between earnings yields and inflation rates may be surprising, because the E/P is supposed to be a real variable. The textbook view is that stocks are real assets since higher inflation should be fully compensated by higher nominal earnings growth rate, with little impact on the stock price or the D/P or E/P ratios.

What explains this anomalous correlation? Here are the main candidates, all of which may contribute:

- Inflation may impact real earning growth prospects —steady low-but-positive inflation appears to be the optimal environment for real growth.
- Inflation may raise prospective real returns because irrational money illusion makes equity markets undervalued (overvalued) when inflation is high (low).[13]

EXHIBIT 9
Dependence of Earnings Yields on Inflation Level—1900–2001



Source: Shiller website, Schroder Salomon Smith Barney.

- Inflation may raise required real returns on bonds *and* equities (rational inflation-related risk premium).

We can explain the bulk of the past 50 years' variation in earnings yields by just two factors: inflation level, and output volatility (see Bernstein [1999], Wieting [2001], and Ilmanen [2002]). The rise and fall in inflation explains the humped shape (20-year rise in earnings yields before 1980 and 20-year fall thereafter), while the trailing volatility of GDP growth rates (or earnings growth rates) explains the general downtrend.

By the end of the century, equity markets benefited from low levels in both factors, in addition to a record-long expansion, productivity optimism, and high risk tolerance after a persistent bull market. No wonder that irrational exuberance and overshooting valuations followed.

The good news is that at least part of the multiple expansion is fundamentally justified. Above-average P/E levels may then be sustainable (as long as inflation stays at the apparently optimal level for equities, near 2%–4%, and macroeconomic stability rather than equity volatility drives equity investors' risk aversion). Yet many observers appear to forget that sustainably high P/E still means low E/P and low long-term equity returns; sustainability would just remove the need for further cheapening in the near term (as the P/E falls to the historical mean).

EXPECTED EQUITY PREMIUMS BASED ON DDM

While the yield ratio is a useful shorthand for the equity-bond premium, the dividend discount model gives us directly what we really want to see: the difference between stocks' and bonds' expected long-run returns.[14] In the basic version of the DDM, equity cash flows (dividends) are assumed to grow at a constant annual rate G. A feasible long-run return on equities is then the sum of the cash flow yield (D/P) and the trend cash flow growth rate (see the appendix). The required return on equities, or the discount rate, can be viewed as a sum of the riskless long-term government yield (Y) and the required equity-bond risk premium (ERP).

Intuitively, markets are in equilibrium when the equity market return that investors require (Y + ERP) equals the rationally feasible expected return (D/P + G). This equality can be reshuffled to express the ex ante equity-bond risk premium in terms of three building blocks:

Equity-Bond Risk Premium ≈
Expected Stock Return –
Expected Bond Return

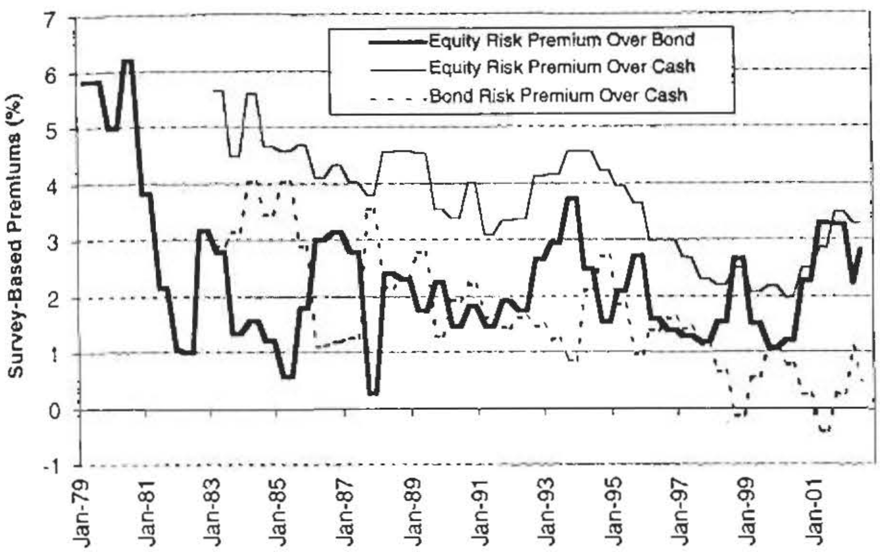
or

$$ERP = D/P + G_{nom} - Y_{nom}$$

The appendix shows how this model can be extended to real (inflation-adjusted) terms or to discounted earnings terms. The DDM framework is simple, but there is a wide disagreement about the inputs to the equity premium calculation. There are two main unobservables, ERP and G. One can either infer ERP for a given G assumption, as we do, or one can reshuffle the equation to infer G (implied growth rate) for a given ERP assumption.

Even the observable inputs—dividend yield and bond yield—are ambiguous. It may be debated whether to include share repurchases in dividend yield and whether to use a ten-year or longer-maturity Treasury yield. The

EXHIBIT 10
Survey-Based Asset Class Premiums—Using Consensus Forecasts of Long-Term GDP Growth, Inflation, and Short-Term Rates—1979–June 2002



Sources: Best and Byrne [2001], Blue Chip Economic Indicators, IBES, and Schroder Salomon Smith Barney.

main source of contention, though, is the assumed trend profit growth rate G.

Instead of assuming a constant profit growth rate, we may allow G to vary over time according to survey forecasts or statistical estimates. Before we explore the various debates, we present equity-bond premium estimates based on survey forecasts of long-term GDP growth rate, motivated by the widely held idea that corporate profit trends are somehow tied to output trends.

Best and Byrne [2001] examine risk premium estimates that use consensus forecasts of next-decade average real GDP growth and inflation as inputs for nominal G. Exhibit 10 shows that the estimated equity-cash risk premium and bond risk premium together trended downward between 1983 and 2000, while the ex ante equity-bond risk premium ranged between 0.5 and 3.5 percentage points.[15]

Debates on Inputs for Statistical Risk Premium Estimates

There will never be full agreement about the equity-bond premium, because there are a wide range of views about DDM inputs. Here we simply summarize the key questions.

Long-Run Growth Rate (G). This is the main debate. Since G is the least-anchored DDM input, differing views on it can shift risk premium estimates by several percentage points, while disagreements about dividend yields and bond yields are worth about 1 percentage point, at most.

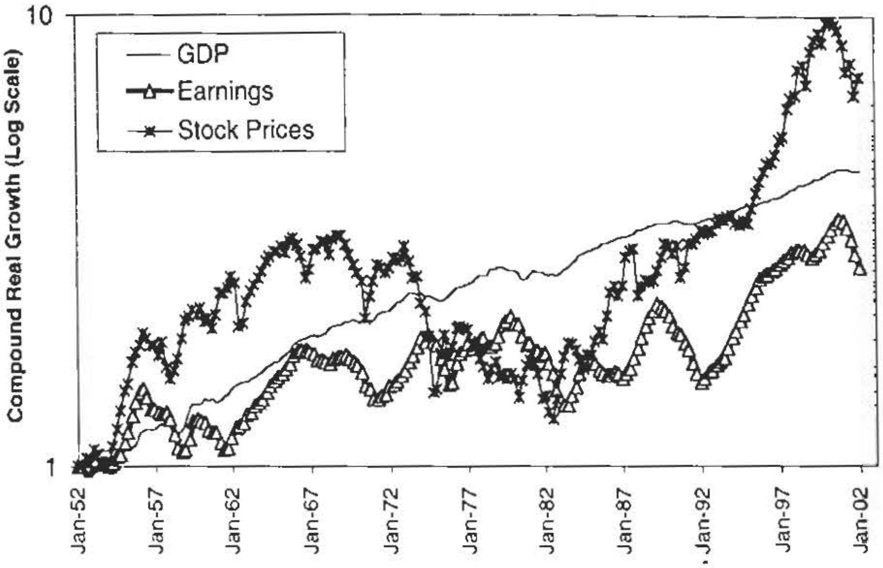
Earnings or dividend data? In historical analyses, some authors use earnings data, others dividend data, and yet others gross domestic product data to proxy for cash flows. While earnings data have their own shortcomings, we use them. Historical dividend growth is arguably understated by the declining trend in dividend payout rate since the late 1970s, partly related to firms' shift from dividend payments toward share repurchases.

Nominal or real G? Many observers refer to historical earnings growth rates in nominal terms (perhaps even using arithmetic averages), thereby overstating future prospects now that inflation rates are quite low. We prefer to assess expected inflation and real earnings growth separately. We do concede that assuming stable nominal earnings growth rates over time could work surprisingly well, because inflation may be inversely related to real earnings growth.

Relation to GDP growth? It is useful to first assess the trend GDP growth rate and then the gap between earnings and GDP growth.

- The long-run productivity growth is important because it determines the potential earnings growth rate, and because persistent changes influence stock prices much more than cyclical changes. If the recent extraordinary productivity growth is sustained, it could be quite bullish for long-run profits and share valuations.
- Historical evidence on the gap between earnings (or dividends) and GDP growth is less encouraging—indeed, recent findings are shocking to many market participants. Several recent studies show that per share earnings and dividends have over long histories lagged the pace of GDP growth and in many cases even per capita GDP growth. Focusing on our past-century sample period (1900-2001), U.S. GDP growth averaged 3.3% in real terms, compared with 1.9% GDP per capita growth, 1.5% earnings growth, and 1.1% dividend growth.

EXHIBIT 11
Cumulative Real Growth of GDP, S&P 500 Operating Earnings, and Stock Prices—1952–2001



Sources: Arnott, Shiller website, and Schroder Salomon Smith Barney.

largely labor-augmenting and wage-enhancing rather than the capital-enhancing type that would spur EPS growth (also see discussion in Nordhaus [2002] and "Proceedings of Equity Risk Premium Forum" [2002]).

Can we do better than using historical averages? Empirical studies find limited predictability in long-term earnings growth rates (see Fama and French [2002]). No predictability implies that the historical sample average may be the best estimate of future earnings growth.

How long a sample? The compound average real earnings growth rate over very long periods is around 1.5%. Others argue that the world has changed, and that the future should be more like the 1990s' experience, with its 4.3% average real earnings growth, and unlike the preceding decades (0.4% in the 1980s and 1.8%-2.9% in the 1950s, 1960s, and 1970s).

Payout rates appear to have some ability to predict future growth, but the results are debatable. Ibbotson and Chen [2002] argue on theoretical grounds that low dividend payout rates are a sign of high growth prospects. Arnott and Asness [2002] show that the empirical experience has been exactly opposite. Low dividend payout rates have preceded low subsequent earnings growth. If this pattern holds, it is a bad omen for the coming years, given the low payout rates of the boom years.[16]

On a positive note, there are some signs that real earnings growth is higher when the trend productivity growth is higher, when the inflation rate is lower (but positive), and when earnings volatility is lower. Lower inflation and volatility drags may have boosted real earnings in the last 15 years and, if sustained, could keep future trend earnings growth more in line with the GDP growth (see Wieting [2001]).

Dividend Yield (D/P). Dividend yields in the United States fell even faster in the 1980s and 1990s than earnings yields. The declining propensity to pay dividends partly reflects a shift toward more tax-efficient share repurchases; by the late 1990s, U.S. firms disbursed cash flows more in share repurchases than in dividends (see Wadhwani [1999], Fama and French [2001], and Jagan-

Exhibit 11 shows that cumulative real growth of earnings has consistently lagged GDP growth in the past 50 years, while stock prices beat GDP only because of the multiple expansion. International evidence in Arnott and Ryan [2001] is hardly more encouraging, and Dimson, Marsh, and Staunton [2002] show that real dividend growth has lagged real GDP per capita growth between 1900-2000 in 15 of the 16 countries they examine.

- What explains these disappointing results? Arnott and Bernstein [2002] attribute them to the dynamic nature of entrepreneurial capitalism. New entrepreneurs and labor (perhaps especially top management) capture a large share of economic growth at the expense of current shareholders. Stock market indexes (made up of listed stocks) do not participate in all growth, and indeed may miss the most dynamic growth of yet-unlisted start-up ventures. Arnott and Bernstein argue that aggregate earnings growth of the corporate sector (listed and unlisted firms) should better keep pace with aggregate GDP growth, and this conjecture seems to hold in the national accounts data.

Siegel [1999] adds that real output growth related to technological progress may have been nathan, McGrattan, and Scherbina [2001]). Adding up dividends and gross buy-backs, however, exaggerates sustainable cash flow yields. One reason is that gross buy-backs should be adjusted for related share issuance (buy-backs are often linked to employee stock options); another is that share repurchase programs are less permanent (easier to discontinue) than dividend payments.

While gross buy-backs added perhaps 2 percentage points and net repurchase payouts 1.5 percentage points to U.S. cash flow yields during the late 1990s peak buy-back years, Liang and Sharpe [1999] argue that adding 0.5 percentage point to dividend yields is a more realistic medium-term estimate. Even this adjustment may be questioned because the 1990s share buy-backs never exceeded new share issuance.

Bond Yield (Y). It is common to use the ten-year government bond yield in equity-bond premium calculations, mainly for data availability reasons. In fact, the "duration" of equities is much longer. Using a longer-maturity yield may thus be appropriate.[17]

Yield curves tend to be upward-sloping, so the use of a longer yield typically reduces the equity-bond premium. But when the yield curve was inverted in the early 1980s, the reverse was true.

Inputs for Ex Ante Asset Returns and Premiums—and Resulting Outputs

Arnott and Bernstein [2002] carefully create a time series of ex ante real long-term stock and bond returns since the early 1800s that would have been realistic to expect, given the information available at the time. Roughly speaking, their inputs include the historical average real dividend growth rate to proxy for the real G (averaging previous 40 years and full-sample experience), a regression-based proxy for expected future inflation, and dividend yield and long-term Treasury yield.[18] These plausible inputs give rise to recently low equity-bond risk premium estimates: near-zero average since the mid-1980s, and negative values between 1997 and 2001.

We propose an alternative set of plausible input assumptions that are somewhat more optimistic for stocks and thus give rise to higher risk premium estimates.[19]

Exhibit 12 summarizes our selections, and Exhibit 13 shows the histories of our inputs (except for yields).

D/P: Since raw dividend yields arguably underestimate recent equity market cash flow yields due to share buy-backs, and since we do not have long histories of net buy-back-adjusted dividend yields, we prefer to use earnings data that have not undergone such a structural change as dividends. We use smoothed earnings yields multiplied by a constant payout rate (0.59) as a proxy for sustainable dividend yields.[20]

G_{real}: As we find limited predictability in long-term real earnings growth, we assume that investors take historical average real earnings growth as a proxy for future G_{real}. The geometric average growth rate is more relevant than the arithmetic average if investors are interested in a long-run wealth accumulation rate.[21]

The historical window length is ambiguous, and we prefer to take an average of the past 10, 20, 30, 40, and 50 years' average growth rates; this choice gives more weight to more recent decades and implies shorter windows than in Arnott and Bernstein [2002]. This approach hopes to capture some slow-moving variation in trend earnings growth rates that may be associated with changing productivity trends and changing inflation or volatility drags.

Since these historical averages are quite unstable over time—the extremes of their range (from −4% to +6%) appear unreasonable for long-run ex ante G views—we take an average of these averages and a 2% anchor for the G_{real} proxy. This admittedly ad hoc approach succeeds in giving a plausible ex ante G_{real} series (a range between 0 and 4% most of the time), while allowing slow variation over time (*see Exhibit 13*). The latest value is 2.5%.

Y: We use the longest available Treasury yield (Ibbotson Associates' roughly 20-year bond until 1951, Salomon Brothers' 20-year or 30-year on-the-run series thereafter), and annualize it. These long bonds' durations are roughly double the ten-year maturity bonds' durations (near seven), and thus are closer to equity durations, although still shorter.

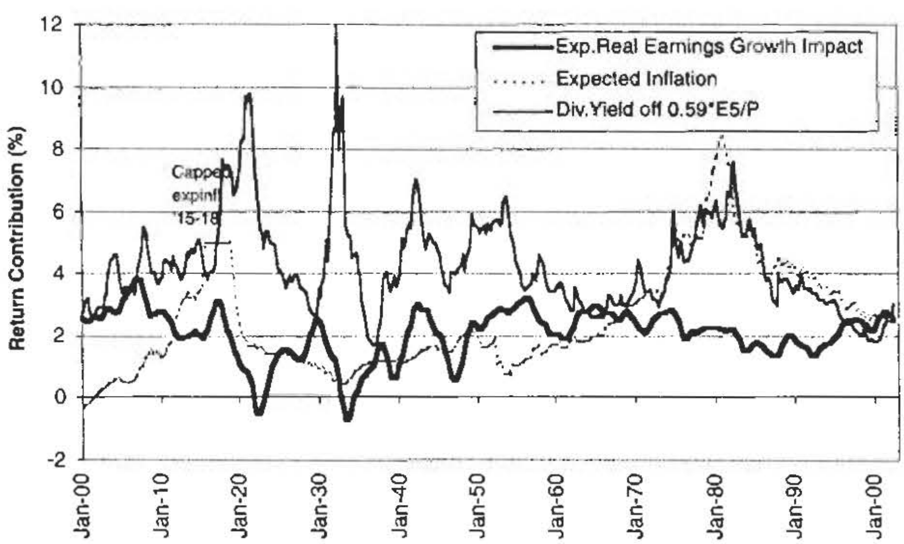
Ex Ante Inflation: We follow Arnott and Bernstein [2002] in regressing each quarter the next-decade inflation on the previous three years' inflation and using the fitted value as a quasi-out-of-sample prediction of the long-term inflation outlook.[22] The regression window length is arbitrary. We use a moving 30-year window and full sample since 1870, averaging the two. We make one exception around World War I; we cap the 1915-1918 expected inflation at 5%, even though our regression proxy rose above it, peaking above 9%.[23]

When survey-based inflation forecasts become available, we incorporate them. After 1951, we use the Livingston survey's median forecast of one-year-ahead inflation as a third component in the average that proxies for expected inflation. And from 1979 when ten-year-ahead

EXHIBIT 12
Estimates of Expected Asset Class Returns and Underlying Input Assumptions

	Input/Assumption:	Mid-2002	*End-99*	(50yr Avg)
Ex Ante Real Stock Return:		5.5%	*4.0%*	(6.2%)
D/P	0.59(5-Year Operating) Earnings Yield	3.0	*1.8*	(3.9)
+ Real Growth (G_{real})	Average of 2% and past 10/20/30/40/50yr real earnings growth adjusted for volatility	2.5	*2.2*	(2.3)
Ex Ante Real Bond Return:		**3.0**	*3.9*	(3.3)
Long Govt Yield	30- or 20-Year Treasury Yield (annualized)	5.6	*6.6*	(6.7)
– Ex Ante Inflation ($E\pi$)	Consensus forecast of decade-ahead inflation since 1979; earlier regression-based long-run inflation forecasts	2.6	*2.7*	(3.4)

EXHIBIT 13
Three Components of Ex Ante Nominal Stock Return—1900–June 2002



Sources: Blue Chip Economic Indicators, FRB Philadelphia, Ibbotson Associates, Arnott, and Schroder Salomon Smith Barney.

survey forecasts are available, we use them as our expected inflation proxy.[24]

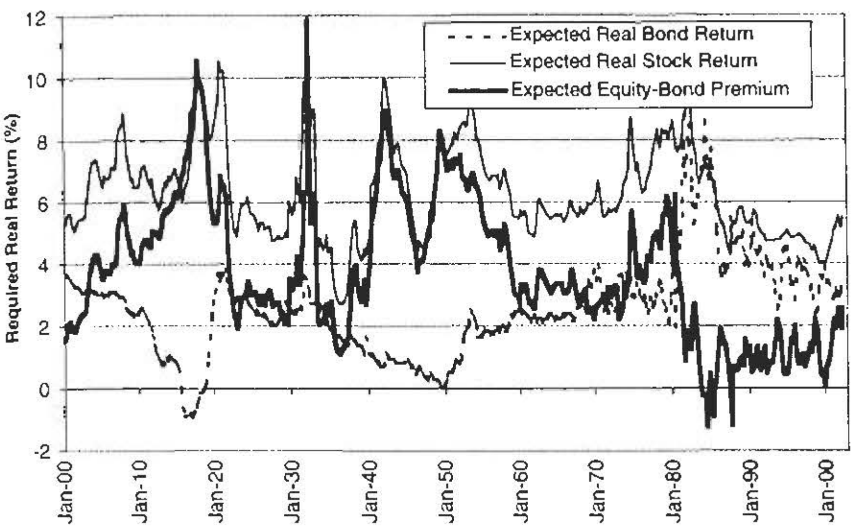
This set of inputs results in the feasible ex ante real long-term stock and bond return series shown in Exhibit 14. The estimated real stock returns varied between 4% and 9% most of the century, sweeping from the top of this range to the bottom between 1982 and 1999. The estimated real bond returns varied between 0% and 5% except for the 1980-1985 period, when ex ante real returns occasionally exceeded 8%. Overall, the post-Second World War pattern of a long upward trend (pre-1982) and a long downward trend (post-1982) in inflation is matched in required real bond returns, although with a short lag.

Bernstein [2002] notes that the great variation in required bond and stock returns in recent decades makes the use of historical returns either irrelevant or, worse, misleading for any kind of future projections.

The equity-bond premium (the difference between the other two series) experienced a clear downward shift 20 years ago. Before 1982, the premium ranged between 2 and 10 percentage points most of the time, while since 1982 the range has mostly been 0 to 2 percentage points.

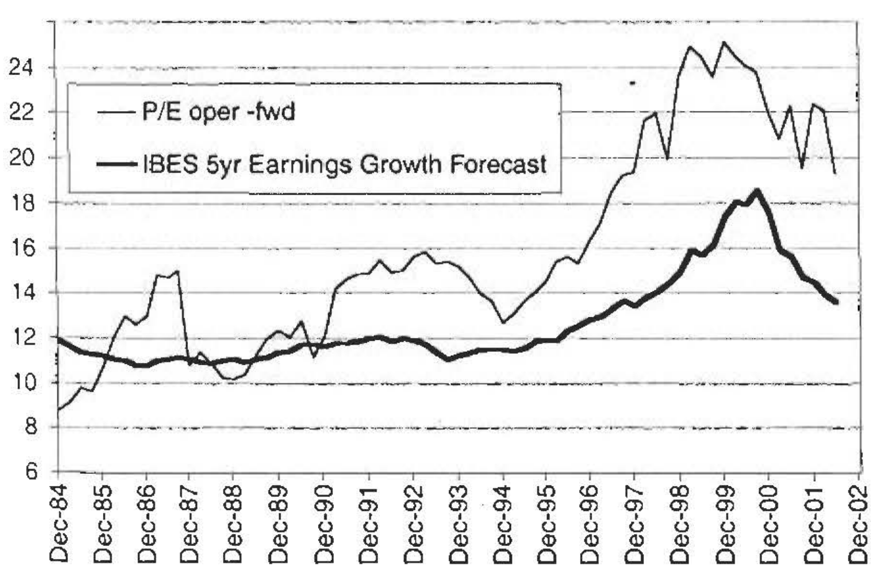
The lowest equity-bond premiums—June 1984, September 1987, and December 1999—coincided with temporary peaks in bond risk premiums. On all three occasions, a Fed tightening triggered a heavy bond market sell-off (year-on-year rises in ten-year yields of 310bp, 220bp, and 180bp, respectively), while equity markets had not yet suffered much. Over

EXHIBIT 14
Estimated Long-Term Real Stock and Bond Returns and their Difference (Ex Ante Premium)—1900–June 2002



Source: Schroder Salomon Smith Barney.

EXHIBIT 15
Forward-Looking P/E Ratio and Analysts' Medium-Term Earnings Growth Forecasts—1985–June 2002



Sources: IBES and Schroder Salomon Smith Barney.

the following year, stocks underperformed bonds by 5, 25, and 26 percentage points, respectively.

It is counter-intuitive that the ex ante equity-bond premium was averaging just 1 percentage point during the great bull market, while realized equity returns between 1982-2001 were 16% per year (*see Exhibit 3*). Using the more conservative Arnott and Bernstein estimates, the ex ante premium was actually negative most of this period.

How could equities outperform bonds by 5 percentage points per year with such a slim ex ante premium? The first answer that comes to mind, a falling equity-bond premium, is not valid for this period; the premium already had shrunk by 1982 and actually edged a bit wider during the 20-year period. A better answer is that discount rates fell (ex ante real returns for stocks fell by 3.5 percentage points, and expected long-run inflation fell even more), and the longest-duration asset class, equities, reaped the greatesr windfall gains from falling rates.

This analysis assigns almost all of the equity outperformance and P/E multiple expansion to lower discount rates rather than greater growth optimism. But recall that our series of feasible ex ante equity returns is based on pretty rational real earnings growth forecasts (that rose just by 1% in the 1990s; see Exhibit 13). Actual subjective growth forecasts probably were much less rational during the Internet boom. Indeed, analysts' medium-term earnings growth forecasts rose from their normally overoptimistic 11%-12% level (of nominal annual growth) to a heady 18%-19% level in 2000, before tailing off (*see Exhibit 15*).

EXHIBIT 16
Forecasting Ability of Various Predictors—Predictive Correlations Based on Quarterly Data

Forecast Horizon and Data Window =>	10yr Return 1900-2001	5yr Return 1900-2001	5yr Return 1960-2001	1yr Return 1900-2001	1yr Return 1983-2001
Predict Real Equity Return Using:					
Trailing Earnings Yield	0.58	0.27	0.17	0.06	0.33
Ex Ante Real Equity Return Estimate	0.40	0.31	0.03	0.25	0.26
Past 5yr Real Equity Return	-0.13	-0.13	0.26	-0.14	-0.40
Predict Real Bond Return Using:					
Nominal Bond Yield	0.54	0.42	0.65	0.29	0.50
Ex Ante Real Bond Return Estimate	0.54	0.61	0.77	0.60	0.62
Past 5yr Real Bond Return	0.08	0.17	0.10	0.04	0.23
Predict Equity-Bond Excess Return Using:					
Earnings Yield Gap (EarnY – GovtY)	0.53	0.32	0.19	0.20	0.56
Ex Ante Equity-Bond Premium Estimate	0.51	0.32	0.05	0.26	0.47
Past 5yr Equity-Bond Excess Return	-0.03	-0.22	-0.28	-0.21	-0.32

Sharpe [2002] uses these growth forecasts, without prejudging their rationality, and estimates that about half of the late-1990s P/E expansion reflects lower discount rates and half greater growth optimism. Thus, part of the late-1990s decline in feasible real equity return in Exhibit 14 likely should be attributed to irrational growth forecasts.

How robust are these estimates of ex ante asset class returns? Details are sensitive to the input assumptions, but the broad contours of such estimates tend to be similar (compare Exhibits 10 and 14), because all are anchored by market yields on equities and bonds.[25] The long-term growth forecasts can vary more widely, and in the basic DDM these forecasts translate one-on-one into higher or lower estimated equity returns or premiums.

Predictive Ability of Equity-Bond Premium Estimates

To assess the usefulness of our ex ante expected return estimates, we use these measures to predict real stock return and real bond return and their difference (excess return) over ten-year, five-year, and one-year horizons. Exhibit 16 displays for each trade the predictive ability of our ex ante expected return measure and two alternative predictors, a simpler yield proxy and a past-return measure.

In all cases, our estimates exhibit reasonable forecasting ability, but they are clearly better predictors than the simple yield measures only at the short (one-year) horizon. The long-horizon correlations are typically higher than short-horizon correlations, mainly because the realized returns are smoother at longer horizons.

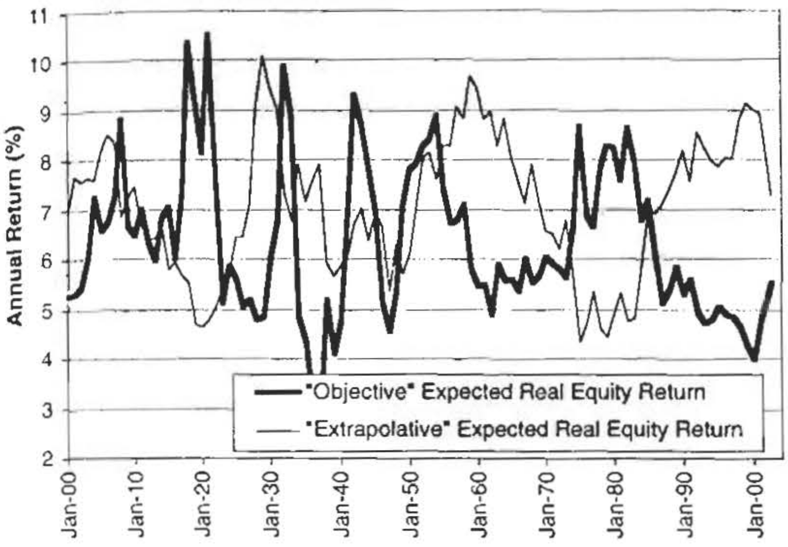
For example, the correlations between the ex ante equity-bond premium and subsequent realized outperformance of equities over bonds are 0.51 for the ten-year horizon, 0.32 for the five-year horizon, and 0.26 for the one-year horizon. In a scatterplot of ex post long-run equity-bond premiums on the ex ante premiums, the 1998-2000 observations show up as major outliers.

Past five-year equity returns (real and excess) have generally been negatively correlated with future returns, consistent with a mild mean-reversion tendency. This pattern underscores the extrapolation risk following an extended period of above-average market returns. Past bond returns on the contrary have been positively related to future returns, consistent with slow-moving variation in required returns.

WHERE DO WE STAND?

While our analysis cannot unambiguously reveal the current extent of the equity-bond premium, our framework does clarify the assumptions needed for various risk premium estimates. Moreover, we argue that

EXHIBIT 17
Gap Between Objectively Feasible (Rational) and Extrapolative (Irrational) Return Expectations—1900–June 2002



Source: Schroder Salomon Smith Barney.

"how high are objectively feasible future stock returns?" is not the only critical question for equity markets' medium-term prospects. Acknowledging the role of irrational expectations, another key question is: "How high returns do investors subjectively expect?" If objective and subjective return expectations are not in balance, equity markets remain vulnerable to disappointments.

There are no directly observable proxies for either return, but we have tried to provide evidence on both. As an illustration only, Exhibit 17 contrasts our estimate of feasible ex ante real equity return with a simple proxy of extrapolative subjective return expectations (75% of a long-run anchor, 7% real equity return, plus 25% of past-decade average real equity return).

Clearly a wide gap arose between the two series in the late 1990s. Just when rising valuations reduced feasible future returns, many investors confused recent windfall gains as a sign of permanently higher equity returns. This gap has narrowed from both sides since the end-1999 peak, but at least in this illustration the gap has not yet been closed.

At a minimum, our framework should give structure to the dialogue about future equity returns. Aggressive return forecasts must be explained by something: high dividend yield, high trend real earnings growth, high inflation, or further multiple expansion. Low dividend yields remain a reality, and from the current above-average valuation levels, further multiple expansion is unlikely.[26]

Since inflation is also likely to remain low, high returns need to be earned the hard way—by very high real profit growth rates.

The mega-bullish equity market view requires throwing away the history books and fully embracing the "this time is different" idea. For example, technology-related arguments might be used to justify a tripling of long-run G_{real} to 4%–5%, which would enable long-run nominal equity returns near 9%–10%. (The finding that the trend earnings growth lags the trend GDP growth does challenge the credibility of such assumptions, given the consensus view of next-decade real GDP growth at 3.1%.)

A moderately constructive case is that feasible and subjectively expected long-run equity returns are in balance near 7%–8%. The deliberately optimistic assumptions we use in Exhibit 12 give rise to 8% feasible (nominal) return, almost as high as the CFO survey forecasts. Stable inflation, low macroeconomic volatility, reduced trading costs, and better diversification opportunities may help sustain the above-average P/E levels. And, given the fall in bond yields, equities again offer more than a negligible risk premium.[27]

A moderately bearish view is that the feasible long-run nominal equity return is closer to 5%–6% than 7%–8%. Such estimates simply follow from using (unadjusted) dividend yields and historical average dividend growth rates.

The most bearish view involves further declines (mean reversion) in the market's P/E multiples. Below-average earnings growth and higher risk aversion are plausible scenarios (see Campbell and Shiller [2001] and Arnott and Asness [2002]). Unwarranted investor optimism, a remnant of the 1990s bull market returns, can also be bad news. Refusal of investors to reconcile themselves to the moderate feasible long-run returns is not sustainable in the medium term.

APPENDIX

Dividend Discount Models and Equity-Bond Premiums

Dividend discount models analyze stocks as if they were perpetual (consol) bonds, with the twist that their coupon rate is expected to grow over time. We describe here the basic Gordon [1962] model with a constant dividend growth rate. Given a constant discount rate R (which can be viewed as a sum of riskless component Y and an equity-bond risk premium component ERP), the stock price can be expressed as the sum of expected discounted future cash flows:

$$P_t = E_t[\sum_{j=1}^{\infty}(\frac{1}{1+R})^j(D_{t+j})]$$

where $R \equiv Y + ERP$.

If we assume a constant growth rate G:

$$E_t(D_{t+j}) = (1 + G)E_t(D_{t+j-1}) = (1 + G)^j D_t$$

we can express the stock price simply as

$$P_t = E_t(D_{t+1})/(R - G) = (1 + G)D_t/(R - G)$$

Thus:

$$E(D_{t+1})/P_t = R - G$$

or as an approximation of the dividend yield:

$$D/P = R - G = Y + ERP - G$$

In equilibrium the equity return that investors require ($R = Y + ERP$) must equal the rationally feasible long-run return ($D/P + G$).

Earnings Discount Model: To express the equation in terms of the E/P ratio, we assume a constant dividend payout rate $k \equiv D/E$. With a constant dividend payout rate, dividend growth rate and earnings growth rate are equal. Then

$$D/P = (E/P)(D/E) = Y + ERP - G$$

Thus:

$$E/P = (Y + ERP - G)/k$$

Real or Nominal: The DDM can be expressed in real terms or in nominal terms. Mechanically, a rise in expected inflation rate raises both the dividend growth rate and the bond yield, without having an impact on the stock price. Empirically, however, the correlation between inflation rates and earnings yields suggests that either real growth rates, payout rates, or equity risk premiums are related to inflation.

Dynamic Models: It is not necessary to assume a constant growth rate. Practical implementations often involve multistage models where growth rate varies over the horizon (see Cornell [1999] and Jagannathan, McGrattan, and Scherbina [2001]). Sharpe [2002] uses a dynamic version of the growth model that allows growth rates and required returns to vary over time. It still follows that low earnings yields are related to high growth prospects or low required returns.

ENDNOTES

The author thanks Robert Arnott, Clifford Asness, Peter Bernstein, Alistair Byrne, and Steven Wieting for helpful discussions and for help in acquiring historical data. This article is largely based on research reports written for Schroder Salomon Smith Barney in May and June 2002. The original disclaimer there applies.

[1]If the payout rate is constant, dividend growth rate and earnings growth rate are equal. We use the latter because payout rates fell in the 1980s and 1990s, and many observers argue that share buy-backs have replaced dividend payments.

[2]The distinction between objective and subjective expectations implies that the subjective expectations can be irrational. In fully rational markets, there is just one expected return that clears the market. The feasible asset return that investors can rationally expect is, by assumption, equal to the required asset return.

[3]Most of our data analysis focuses on U.S. markets because the literature has concentrated on them, partly because of better data availability and reliability. The global leading role of the U.S. economy and asset markets and higher valuation ratios than in most other major equity markets also make the U.S. experience the most interesting topic.

[4]All returns are expressed as annual compound returns, unless otherwise stated.

[5]One reason is that U.S. government bonds were not perceived to be riskless until the 20th century. In addition, yield trends were more favorable for bonds as the 19th century ended with extended deflation. Long yields were then halved from 1802's near-6% level to near 3% at the beginning of 1900, and then doubled back by the end of 2001. Of course, equity and bond markets also were less developed in the 1800s, making data less comprehensive and reliable.

[6]The peso problem refers to infrequent, unlikely events such as currency devaluation that may influence market pricing (e.g., forward bias in peso-dollar pricing) but may not show up, even in a long historical sample.

[7]The CFO survey and our bond investor survey asked for views on the expected annual return of a major equity index over the next decade. The academic survey required some adjustments because it asked for the 30-year equity-bill-premium (and only an arithmetic average in 1998). We first subtract from the 7% consensus view in 1998 0.8 percentage point (the gap between arithmetic and geometric means in the later survey), then add a 5% expected average bill rate (typical long-run view of economists in 1998 from another survey) to get an 11.2% expected nominal return. In 2001, the survey quotes a 4.7 percentage point geometric mean premium over bills; we add 4.7% expected average bill rate to it to get a 9.4% estimate.

[8]The falling consensus views may partly reflect a real change due to the growing literature on the changing equity risk premium, besides simple extrapolation from recent returns.

[9]Specifically, we have found that the negative correlation between stock and bond returns has made government bonds

the ultimate safe haven. The negative beta feature can even justify a negative risk premium for government bonds when the traditional inflation risk premium has fallen to near zero. All else equal, a low or negative bond risk premium (over cash) makes the current equity-bond premium wider. (See Best, Byrne, and Ilmanen [1998] and Ilmanen [2002].)

[10]We use operating earnings rather than reported earnings since the former became available in the early 1980s. Broadly speaking, operating earnings are earnings from continuous operations, excluding non-recurring items. Operating earnings may give a better picture of trend earnings, as they are less influenced by one-off events and cyclical downturns (see Wieting and Peng [2002]).

Findings of aggressive and even illegal earnings accounting practices, however, have made many investors prefer the reported earnings. Stock option expensing and pension return assumptions are other contentious earnings topics. Any adjustments to recent earnings levels would imply lower earnings yields and lower ex ante equity returns in our empirical analysis.

[11]Improving macro stability has not brought along financial market stability, an unattractive outcome for equity investors. Alan Greenspan, among others, highlighted the contrast between low output volatility and high equity market volatility in his annual Jackson Hole speech in August 2002.

[12]Overall, Japan's experience confirms the inflation-dependence of earnings yields but there is a hint of a leaning J-shape. We conjecture that earnings yields could actually rise in a deflationary environment. Low-but-positive inflation is the optimal environment for equity valuations; both higher inflation and deflation can hurt equities and raise E/P ratios. This also suggests that U.S. equity multiples already reflect all the possible gains from disinflation and that the best they can do now is to hold onto these gains (if inflation remains near 2%-4%).

[13]Modigliani and Cohn [1979] argue that investors and analysts incorrectly discount real dividend streams with nominal discount rates, resulting in too low a price for real fundamentals when inflation is high. For a recent review, see Ritter and Warr [2002]. Sharpe [2002] suggests a variant of inflation illusion: Investors and analysts actually discount nominal cash flows using nominal discount rates, but do not make sufficient inflation adjustments to their extrapolative nominal growth forecasts.

[14]Under certain conditions, the earnings yield equals the ex ante real equity return—for example, if the constant retention rate (1 – payout rate) matches the constant dividend growth rate. Intuitively, earnings yield understates expected return because it excludes dividend growth, but it exaggerates expected return because only a part of earnings are paid out as dividends. Unless the two extra terms just balance, the DDM should provide a better ex ante real return measure than the earnings yield.

[15]The equity-cash premium is the difference between the ex ante equity return and the expected average Treasury bill rate over the next decade. The bond risk premium is the difference between the ten-year Treasury yield and the expected average Treasury bill rate over the next decade. The equity-bond premium is the difference between the ex ante equity return and the ten-year Treasury yield.

The nominal ex ante equity return is estimated as a sum of the dividend yield (proxied by a forward-looking earnings yield times a constant assumed payout rate), expected long-run real GDP growth rate, and expected inflation. The main raw material is economists' consensus forecasts of next-decade average real GDP growth, inflation, and Treasury bill rates from the semiannual Blue Chip Economic Indicators survey.

Note that using the current Treasury bill yield in equity premium calculations could be quite misleading when short rates are exceptionally low (or high) and expected to revert to normal levels. For example, the current three-month rate is near 2%, while the expected next-decade average short rate is above 4%.

[16]The theoretical argument is in the "Modigliani-Miller spirit," based on the idea that management retains a greater share of earnings when it sees greater future profit opportunities. The empirical finding that high retention rates predict low earnings growth may reflect management's exuberance or inefficient empire building (see Arnott and Asness [2002]). Alternatively, management may be concerned with dividend smoothing, and will pay higher dividends only when it can afford (or dares) to do so, given its expectation of strong future profit growth.

[17]In the DDM context, the equity market can be viewed as a consol bond with a growing coupon rate. It follows from simple algebra that the modified duration of equities is $1/(R - G)$, which is just the inverse of the dividend yield. For D/P of 2.5%, this duration is 40, but this result is model-dependent; recall that the basic model assumes constant R and G. More generally, equities really are long-duration assets, that is, very sensitive to *permanent* discount rate changes—and more so when dividend yields are low.

[18]Arnott and Bernstein present the real dividend growth rate component in two parts: the predicted long-run growth rate of GDP per capita, and the predicted dilution of dividend growth versus GDP per capita growth.

[19]Our exercise follows in the same spirit as the Arnott-Bernstein study—trying to come up wirh reasonable views on each of the DDM inputs (say, what long-term real growth rate and what inflation rate investors could have expected at the time). There is sufficient uncertainty about these inputs that both sets of assumptions can be deemed plausible. Our assumptions are deliberately more optimistic than those of Arnott and Bernstein, to see how much expected returns rise if we add an implicit adjustment for share buy-backs to dividend yields, and if we use higher, but not outrageous, earnings growth estimates.

[20]Recall that $D/P = (D/E)(E/P)$. Since one-year trailing earnings yields are volatile, we use smoother five-year average earnings.

[21]We do not use geometric averages but rather a closely related procedure proposed in Fama and French [2002]. We reduce arithmetic averages by half the variance difference

between the earnings growth rate and dividend growth rate.

[22]The simple approach we use captures both the past average as an anchor and the varying sensitivity of future expectations to current inflation; this sensitivity increased during the 20th century once inflation became more persistent. We explored other inflation forecasting models with yield and growth indicators. The results were not robust, perhaps because forecasting decade-ahead developments leaves us with few independent observations.

[23]War-related inflations had typically been temporary before the First World War. More generally, inflation had not been persistent in the past, so investors had little reason to raise long-run inflation expectations sky-high (and would have been right, as a deflation soon followed). The 5% cap actually may be too high, given that the 1800s experienced mild net deflation, and given that bond yields stayed below 5% through the 1915-1918 period.

[24]Our proxy series and the consensus forecast are closely related during the overlapping period, and there is no large jump when moving from one series to another.

[25]As we have noted, even these yields are subject to debate about the impact of share buy-backs on dividend yields and about the appropriate Treasury maturity. Our current D/P estimate of 3.0% in Exhibit 12 is especially high, virtually double the raw number. This high level is partly offset in the equity-bond premium by our use of the 30-year Treasury yield (1 percentage point higher than the 10-year yield).

[26]Our analysis ends in mid-2002, but even during the third-quarter 2002 equity sell-off the dividend yield rose only to 2%. The long duration of equities means that feasible returns rise painfully slowly; a 15%-20% price decline may increase the feasible long-term return by about 0.5 percentage point. Yet the 1% fall in long-term Treasury yields in the third quarter had a greater impact on the equity-bond premium, raising our estimate to nearly 4 percentage points. Greater attractiveness versus bonds can benefit equities in the near term, but a wide cushion does not make the absolute level of feasible equity return any higher. It is unclear whether absolute or relative return prospects matter more.

Further disinflation or yield declines are unlikely to boost P/E ratios, because they likely would reflect bad deflation. Moreover, there appears little chance that the late-1990s growth optimism, exuberant sentiment, and risk tolerance will reappear any time soon. Observed empirical patterns (mean reversion, low payout rates) point rather to lower P/E multiples in the future. A cyclical upturn supported by easy monetary policy can of course raise equity valuations and realized returns over a shorter horizon.

[27]Siegel [1999] and Carlson, Pelz, and Wohar [2002] review these arguments. Jones [2002] provides specific evidence of falling trading costs during the past century and notes that the gross equity premium may have fallen by 1 percentage point as a result.

REFERENCES

Arnott, Robert D., and Cliff S. Asness. "Does Dividend Policy Foretell Earnings Growth?" Working paper, First Quadrant, 2002.

Arnott, Robert D., and Peter L. Bernstein. "What Risk Premium is 'Normal'?" *Financial Analysts Journal*, March/April 2002.

Arnott, Robert D., and Ronald Ryan. "The Death of the Risk Premium: Consequences of the 1990s." *The Journal of Portfolio Management*, Spring 2001.

Asness, Cliff. "Bubble Logic or How to Learn to Stop Worrying and Love the Bull." Working paper, AQR, June 2000a.

——. "Stocks vs. Bonds: Explaining the Equity Risk Premium." *Financial Analysts Journal*, March/April 2000b.

Bernstein, Peter L. "The Risk Premium: A Critical Look and Some Pleasant Conclusions." *Economics and Portfolio Strategy*, April 1999.

——. "The Uses of History." *Economics and Portfolio Strategy*, June 2002.

Best, Peter, and Alistair Byrne. "Measuring the Equity Risk Premium." *Journal of Asset Management*, 1 (2001).

Best, Peter, Alistair Byrne, and Antti Ilmanen. "What Really Happened to U.S. Bond Yields?" *Financial Analysts Journal*, May/June 1998.

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

Carlson, John B., Eduard A. Pelz, and Mark E. Wohar. "Will Valuation Ratios Revert to Historical Means?" *The Journal of Portfolio Management*, Summer 2002.

Cornell, Bradford. *The Equity Risk Premium*. New York: Wiley, 1999.

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

Fama, Eugene F., and Kenneth R. French. "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay." *Journal of Financial Economics*, 60 (2001).

——. "The Equity Premium." *Journal of Finance*, 57 (April 2002).

Gordon, Myron J. *The Investment, Financing, and Valuation of the Corporation*. Homewood, IL: Irwin, 1962.

Graham, John R., and Campbell R. Harvey. "Expectations of Equity Risk Premia, Volatility and Asymmetry From a Corporate Finance Perspective." Working paper, NBER, 2001.

Ibbotson, Roger, and Peng Chen. "Stock Market Returns in the Long Run: Participating in the Real Economy." Working paper, Yale ICF, 2002 (forthcoming, *Financial Analysts Journal*).

Ilmanen, Antti. "Stocks Versus Bonds: Balancing Expectations and Reality." Schroder Salomon Smith Barney, June 2002.

Jagannathan, Ravi, Ellen R. McGrattan, and Anna Scherbina. "The Declining U.S. Equity Risk Premium." Working paper, NBER, 2001.

Jones, Charles M. "A Century of Stock Market Liquidity and Trading Costs." Working paper, Columbia University, May 2002.

Liang, J. Nellie, and Steven A. Sharpe. "Share Repurchases and Employee Stock Options and Their Implications for Expected Returns." Working paper, FRB, 1999.

Mehra, Rajnish, and Edward C. Prescott. "The Equity Premium: A Puzzle." *Journal of Monetary Economics*, 15 (March 1985).

Modigliani, Franco, and Richard Cohn. "Inflation, Rational Valuation and the Market." *Financial Analysts Journal*, March/April 1979.

Nordhaus, William D. "The Mildest Recession: Output, Profits and Stock Prices as the U.S. Emerges From the 2001 Recession." Working paper, NBER, 2002.

Poterba, James M. "The Rise of the 'Equity Culture': U.S. Stockownership Patterns 1989-1998." Working paper, MIT, 2001.

"Proceedings of Equity Risk Premium Forum." AIMR and TIAA-CREF, 2002.

Ritter, Jay R., and Richard S. Warr. "The Decline of Inflation and the Bull Market of 1982-1999." *Journal of Financial and Quantitative Analysis*, March 2002.

Sharpe, Steven A. "Re-examining Stock Valuation and Inflation: The Implications of Analysts' Earnings Forecasts." *Review of Economics and Statistics*, August 2002.

Shiller, Robert J. *Irrational Exuberance*. Princeton: Princeton University Press, 2000.

Siegel, Jeremy J. "The Shrinking Equity Risk Premium." *The Journal of Portfolio Management*, Fall 1999.

——. *Stocks for the Long Run*. New York: McGraw-Hill, 1998.

Wadhwani, Sushil B. "The U.S. Stock Market and the Global Economic Crisis." *National Institute Economic Review*, 1999.

Welch, Ivo. "The Equity Premium Consensus Forecast Revisited." Discussion paper, Cowles Foundation, 2001.

——. "Views of Financial Economists on the Equity Risk Premium and on Financial Controversies." *Journal of Business*, October 2000.

Wieting, Steven. "Regress to the Mean or Cause and Effect?" Salomon Smith Barney, 2001.

Wieting, Steven, and Ken Peng. "Inside the S&P 500." Salomon Smith Barney, 2002.

Note

Biases in Arithmetic and Geometric Averages as Estimates of Long-Run Expected Returns and Risk Premia

Daniel C. Indro and Wayne Y. Lee

Daniel C. Indro is an Assistant Professor of Finance and Wayne Y. Lee is Firestone Professor of Corporate Finance at Kent State University.

The empirically documented presence of negative autocorrelation in long-horizon common stock returns magnifies the upward (downward) bias inherent in the use of arithmetic (geometric) averages as estimates of long-run expected returns and risk premia. Failure to account for this autocorrelation can lead to incorrect project accept/reject decisions. Through simulations, we show that a horizon-weighted average of the arithmetic and geometric averages contains a smaller bias and is a more efficient estimator of long-run expected returns.

■ Consider an investment project with an average life (duration) of N months. What rate should be used to discount this project's expected cash flows? In particular, suppose the required return on the N-month investment project is based on a *market* equity-risk premium, that is, the difference between the future expected return on the market index and the risk-free rate of interest. Since risk premia are not constant (Brigham, Shome, and Vinson, 1985; Harris, 1986; Harris and Marston, 1992; Maddox, Pippert, and Sullivan, 1995; and Brennan, 1997) and can depend on the choice of measurement period, averaging method, or portfolio weighting (Carleton and Lakonishok, 1985), how should the historical monthly market return data be used to compute the risk premium? In practice, the arithmetic and geometric average of monthly returns are used as a proxy for determining the future expected N-month market return.[1]

Brealey and Myers (1991) argue that if monthly returns are identically and independently distributed, then the arithmetic average of monthly returns should be used to estimate the long-run expected return. However, the empirical evidence from Fama and French (1988a, 1988b), Lo and MacKinlay (1988), and Poterba and Summers (1988) suggests that there is significant long-term negative autocorrelation in equity returns and that historical monthly returns are not independent draws from a stationary distribution. Based on this evidence, Copeland, Koller, and Murrin (1994) argue that the geometric average is a better estimate of the long-run expected return. Thus, as noted by Fama (1996), when expected returns are autocorrelated, compounding a sequence of one-period returns is problematic for project valuation.

In this paper, we examine the biases obtained by using the arithmetic or geometric sample averages of single-period returns to assess the long-run expected rates of return when there is both a time-varying and a stationary component in those returns. To do this, we adopt the analytical framework outlined in Blume (1974). We find that for long-run expected return and risk premium, the arithmetic average produces an

We wish to thank Michael Hu, the Editors, and especially the referee whose comments and suggestions greatly improved the paper's expositions. We are responsible for any remaining errors.

[1]Alternatively, in deriving the cost of equity estimates, Harris (1986) and Harris and Marston (1992) employ the Discounted Cash Flow (DCF) model, which uses a consensus measure of financial analysts' forecasts of earnings growth as a proxy for investor expectations. Although this alternative is appealing, Timme and Eisemann (1989) caution that it requires a judicious choice of the weight assigned to each forecast to construct the consensus forecast. Otherwise, the DCF model can generate a risk-adjusted discount rate that contains estimation risk and requires an adjustment such as that outlined in Butler and Schachter (1989).

Financial Management, Vol. 26, No. 4, Winter 1997, pages 81 - 90

estimate that is too high relative to the true mean, and that the geometric average produces an estimate that is too low. The magnitude of upward and downward bias is proportional to the total variance underlying the asset's return, and to the length of the investment horizon (N months) relative to the length of the historical sample period ($T \geq N > 1$). In addition, we confirm Blume's finding that there are significant biases associated with the use of the arithmetic and geometric averages, even when returns are independently and identically distributed each period. Finally, simulation results show that the horizon-weighted average of the arithmetic and geometric averages proposed by Blume is less biased and more efficient than alternative estimates.

I. The Bias in the Arithmetic and Geometric Averages

Here, we describe the return generating process and derive the biases in the arithmetic and geometric averages.

A. Return Generating Process

Let R_t denote a one-period total return over a time interval of length dt. Specifically,

$$R_t = 1 + r_t dt = 1 + \mu_t dt + \varepsilon_t \sqrt{dt} \quad (1)$$

where $r_t dt$ is the net return for period $t = 1,2,....,T$; $\mu_t dt$ is the conditional mean, and the deviations from the conditional mean, $\varepsilon_t \sqrt{dt}$ are independently and identically distributed over time with mean zero and variance $\sigma_\varepsilon^2\, dt$. Further, assume that the conditional mean $\mu_t dt$ is distributed as follows. For $t = 1$, the conditional mean is

$$\mu_1 dt = \mu dt + \eta_1 \sqrt{dt} \quad (2)$$

where μdt is the unconditional mean. For $t = 2,3,....,T$, the conditional mean follows a mean-reverting process around the unconditional mean:

$$\mu_{t+1} dt = \mu dt + \rho(\mu_t dt - \mu dt) + \eta_{t+1}\sqrt{dt} = (1 - \rho)\, \mu dt + \rho\mu_t dt + \eta_{t+1}\sqrt{dt} = \mu dt + \Sigma_{j=1}^{t+1} \rho^{t-j+1}\, \eta_j \sqrt{dt} \quad (3)$$

where the single-period autocorrelation between conditional means, $\rho \leq 0$, captures the time variation in expected returns, and $\eta_t \sqrt{dt}$ are independently and identically distributed random variables with mean zero and variance $\sigma_\eta^2\, dt$. From Equations (1) through (3) it follows that

$$r_t dt = \mu dt + \varepsilon_t \sqrt{dt} + \Sigma_{i=1}^{t} \rho^{t-i} \eta_i \sqrt{dt} = \mu dt + \nu_t \sqrt{dt} \quad (4)$$

for all t. The return generating process described by Equation (4) is consistent with that used by Fama and French (1988a) to document significant negative autocorrelations in long-horizon returns.[2] The unconditional mean, $E(r_t dt)$, is μdt. The unconditional variance, $Var(r_t dt)$, is $[(1-\rho^{2T})/(1-\rho^2)]\sigma_\eta^2 dt + \sigma_\varepsilon^2 dt$ for a finite T, and $[1/(1-\rho^2)]\sigma_\eta^2 dt + \sigma_\varepsilon^2 dt$ as $T \to \infty$.

B. The Bias in the Arithmetic Average

From a sample of T observations, we compute the arithmetic average, R_A, as:

$$R_A = 1 + r_A dt = 1 + \mu dt + T^{-1}\Sigma_{t=1}^{T} \nu_t \sqrt{dt} \quad (5)$$

and the estimated N-period return, $R_A^N = (1 + r_A dt)^N$,

$$R_A^N = (1 + \mu dt + T^{-1}\Sigma_{t=1}^{T} \nu_t \sqrt{dt})^N \quad (6)$$

In addition, applying the expected value operators to Equation (6) yields:

$$E(R_A^N) = E(1 + \mu dt + T^{-1}\Sigma_{t=1}^{T} \nu_t \sqrt{dt})^N \quad (7)$$

Since $(1 + \mu dt + T^{-1}\Sigma_{t=1}^{T} \nu_t \sqrt{dt})^N$ is a convex function of $T^{-1}\Sigma_{t=1}^{T} \nu_t \sqrt{dt}$, it follows by Jensen's inequality that for $N > 1$, the arithmetic average is biased upward:

$$E(R_A^N) > (1 + \mu dt + E(T^{-1}\Sigma_{t=1}^{T} \nu_t \sqrt{dt}))^N > (1 + \mu dt)^N \quad (8)$$

Further, by taking a Taylor series expansion of $E(R_A^N)$ around $(1 + \mu dt)$, the extent of the bias is given by:[3]

$$E(R_A^N) = (1 + \mu dt)^N [1 + \frac{N(N-1)}{2} (1 + \mu dt)^{-2} \sigma_\xi^2\, dt] + O(dt^2) \quad (9)$$

---

[2]Specifically, in Fama and French (1988a), p(t), the natural log of a stock price at time t, is the sum of a random walk, q(t), and a stationary component, z(t):

$$p(t) = q(t) + z(t) \text{ and } q(t) = q(t-1) + \mu + \varepsilon(t) \quad (3a)$$

where μ is expected drift and $\varepsilon(t)$ is white noise. z(t) follows a first-order autoregression (AR1) process:

$$z(t) = \phi z(t-1) + \eta(t) \quad (3b)$$

where $\eta(t)$ is white noise and ϕ is less than 1. From Equations (3a) and (3b), we compute a continuously compounded return:

$$p(t) - p(t-1) = [q(t) - q(t-1)] + [z(t) - z(t-1)] = \mu + \varepsilon(t) + \eta(t) + (\phi-1)z(t-1) \quad (3c)$$

Through successive substitutions for $z(\cdot)$ from Equations (3b) into (3c), the consistency between our formulation and that of Fama and French (1988a) follows from a comparison of Equations (3c) and (3).

[3]Derivations of the extent of biases in the arithmetic and geometric averages are available from the authors on request.

where $O(dt^2)$ denotes an order of no greater than dt^2, $\lim O(dt^2) \to 0$ as $dt \to 0$. From Equation (5), $\xi\sqrt{dt} = T^{-1}\Sigma_{t=1}^{T} v_t\sqrt{dt}$, and

$$\sigma_\xi^2 dt = E[(\xi\sqrt{dt})^2] = T^{-2}(T\sigma_\eta^2 dt + \Sigma_{i=1}^{T}(T - i)\rho^{2i}\sigma_\eta^2 dt) + T^{-2}(T\sigma_\xi^2 dt) = T^{-1}(\sigma_\eta^2 dt + \sigma_\varepsilon^2 dt) + T^{-1}\left(\frac{(T+1)}{2}\right)\rho^{2\tau}\sigma_\eta^2 dt \tag{10}$$

since by the mean value theorem there exists a τ, $T > \tau > 1$ such that $\Sigma_{t=1}^{T}(T - i)\rho^{2i} = \Sigma_{t=1}^{T}(T - i)\rho^{2\tau}$.

For $\rho = 0$ and fixed N, it is clear that the estimator R_A^N is asymptotically unbiased and consistent as $T \to \infty$, but for a finite and small T, is upward-biased for $N > 1$ by an amount proportional to the number of periods, $[N(N-1)/2]$, and variance, $T^{-1}(\sigma_\eta^2 dt + \sigma_\varepsilon^2 dt)$. Furthermore, for $\rho < 0$ and fixed N, the estimator R_A^N is asymptotically unbiased and consistent only for $N = 1$. For $N > 1$, the amount of upward bias is proportional to the number of periods, $[N(N-1)/2]$, and either the variance $\frac{1}{2}\rho^{2\tau}\sigma_\eta^2 dt$ for $T \to \infty$, or the variance $T^{-1}(\sigma_\eta^2 dt + \sigma_\varepsilon^2 dt) + T^{-1}[(T+1)/2]\rho^{2\tau}\sigma_\eta^2 dt$ for a finite and small T. Compounding the single-period arithmetic return tends to produce an estimated long-run return, and thus a risk premium, that is too high relative to the true mean $(1 + \mu dt)^N$.

C. The Bias in the Geometric Average

From a sample of T observations, the geometric average, R_G, is computed as:

$$R_G = \left(\Pi_{t=1}^{T} R_t\right)^{1/T} \tag{11}$$

and the estimated N-period return, R_G^N, as:

$$R_G^N = \left(\Pi_{t=1}^{T} R_t\right)^{N/T} = \exp\left\{\frac{N}{T}\Sigma_{t=1}^{T} \ln R_t\right\} \tag{12}$$

Hence, for a fixed N and $T \to \infty$, it is clear from Equation (12) that

$$p\lim R_G^N = \exp\left\{p\lim \frac{N}{T}\Sigma_{t=1}^{T} \ln R_t\right\} = \exp\{NE[\ln R_t]\} < \exp\{N \ln [E(R_t)]\} < 1 + \mu dt \tag{13}$$

The geometric average is asymptotically biased downwards and thus is an inconsistent estimator of the long-run expected return.

To examine the bias for a fixed N and finite T, we rewrite the geometric average as:

$$R_G^N = \left(\Pi_{t=1}^{T} R_t\right)^{N/T} = \Pi_{t=1}^{T} (1 + \mu dt + v_t\sqrt{dt})^{N/T} = [(1 + \mu dt)^T + \zeta\sqrt{dt}]^{N/T} \tag{14}$$

where

$$\zeta\sqrt{dt} = \Pi_{t=1}^{T} (1 + \mu dt + v_t\sqrt{dt}) - (1 + \mu dt)^T \tag{15}$$

Taking the expectation of Equation (14) and a Taylor series expansion around $(1 + \mu dt)^T$ yields:

$$E(R_G^N) = E[(1 + \mu dt)^T + \zeta\sqrt{dt}]^{N/T} = (1 + \mu dt)^N + \left(\frac{N}{T}\right)(1 + \mu dt)^{N-T} E(\zeta\sqrt{dt}) + \left(\frac{N}{T}\right)\left(\frac{N}{T} - 1\right)(1 + \mu dt)^{N-2T} E(\zeta\sqrt{dt})^2 + O(dt^2) \tag{16}$$

where

$$E(\zeta\sqrt{dt}) = (1 + \mu dt)^{T-2}[\Sigma_{i=1}^{T-1}\rho^{2i-1}\Sigma_{j=1}^{T-i} j\rho^{T-i-j}]\sigma_\eta^2 dt + O(dt^2) \tag{17}$$

and

$$E(\zeta\sqrt{dt})^2 = (1 + \mu dt)^{2(T-1)}[T(\sigma_\varepsilon^2 dt + \sigma_\eta^2 dt) + \rho^2\sigma_\eta^2 dt\, \Sigma_{i=1}^{T-1}(T-i)\rho^i + 2\sigma_\eta^2 dt\, \Sigma_{i=1}^{T-1}\rho^{2i-1}\Sigma_{j=1}^{T-i} j\rho^{T-i-j}] + O(dt^2) \tag{18}$$

Observe that for $\rho = 0$,

$$E(R_G^N) = (1 + \mu dt)^N \left\{1 + (1 + \mu dt)^{-2}\left(\frac{N}{T} - 1\right)[T(\sigma_\varepsilon^2 dt + \sigma_\eta^2 dt)]\right\} \tag{19}$$

the geometric average is downward-biased for $N < T$ but unbiased as $N \to T$. For $\rho < 0$,

$$E(R_G^N) = (1 + \mu dt)^N \left\{1 + \left(\frac{N}{T}\right)(1 + \mu dt)^{-2} [E(\zeta\sqrt{dt}) + \left(\frac{N}{T} - 1\right)E(\zeta\sqrt{dt})^2]\right\} \tag{20}$$

By definition, $E(\zeta\sqrt{dt})^2 = \mathrm{Var}(\zeta\sqrt{dt}) > 0$, and it can be shown that $E(\zeta\sqrt{dt}) \le 0$ for $\rho \le 0$.[4] Hence, from Equation (20), the geometric average is always biased downward for $\rho < 0$, even as $N \to T$. It is also clear from Equation (20) that an increase in the stationary variance $\sigma_\varepsilon^2 dt$ raises the magnitude of the downward bias. The effect on the bias of changes in the parameters governing the temporal variation in expected returns, namely, ρ and $\sigma_\eta^2 dt$, is generally ambiguous. However, when $N \to T$,

$$E(R_G^N) = (1 + \mu dt)^N\{1 + (1 + \mu dt)^{-2}[1 + (T - 2)\rho]\rho\sigma_\eta^2 dt + O(\rho^3)\sigma_\eta^2 dt\} \tag{21}$$

the downward bias at the limit is an increasing function of ρ and $\sigma_\eta^2 dt$.

[4]The sketch of the proof is as follows. Let T = 5. Compute and sum the five variances and ten covariances of $v_t\sqrt{dt}$. Examining the covariance sum for $\rho \le 0$ results in $E(\zeta\sqrt{dt}) \le 0$. The general result is obtained by induction. The formal derivation is available from the authors on request.

II. Simulation Results

We use simulations to assess the severity of the biases in the arithmetic and geometric averages. In addition, we present two other estimates of expected return, as suggested in Blume (1974): a weighted average and an overlapping average.

We calculate the weighted average as a horizon-weighted average of the arithmetic and geometric averages:

$$E(W^N) = \frac{T-N}{T-1} R_A^N + \frac{N-1}{T-1} R_G^N \tag{22}$$

where the weights sum to one. When N=1, the arithmetic average receives all the weight. As $N \to T$, more weight is given to the geometric average.

We construct the overlapping average as follows. We compute an N-period total return, T-N+1 in number, by multiplying the first through the N^{th} one-period total returns together, the second through the $(N+1)^{st}$ one-period returns together, and so on. We then average the overlapped total returns.

To examine the empirical properties of each estimator, we use the return generating process described in Equation (3). For a benchmark monthly return, $\mu = 0.01$, and alternative values of autocorrelations $\rho = 0$, -0.05, -0.25, we draw a total of 250,000 random values of $\varepsilon_t\sqrt{dt}$ and $\eta_t\sqrt{dt}$ from zero mean normal variates with variances ranging from zero to 0.0081 for σ_ε^2 and zero to 0.0045 for σ_η^2, respectively. We then partition the 250,000 returns into 1,000 samples of 250 observations (T =250), and calculate the values of the four estimators for horizons N = 12,24,60,84,120.

Table 1 presents the simulation results when the autocorrelation and time-varying variance components are absent, i.e., $\rho = 0$ and $\sigma_\eta^2 = 0$. Simulation results in the presence of both time-varying and stationary variance as well as negative autocorrelation components appear in Table 2 (ρ =-0.05) and Table 3 (ρ = -0.25).

For the four estimators, the patterns of bias (direction and magnitude) and efficiency (standard deviation or the 0.05-0.95 fractile values) that appear in Table 1 are similar to those found in Blume (1974). Notice from Table 1 that for any investment horizon and stationary variance, the geometric average is always biased downward. For longer horizons N (=60,84,120), the arithmetic average is upward-biased, regardless of the stationary variance. For shorter horizons, N (=12,24), the arithmetic average is downward-biased for a small value of stationary variance, σ_ε^2 (= 0.0036), but upward-biased for a large value of stationary variance, σ_ε^2 (= 0.0081). For a small value of stationary variance, σ_ε^2 (= 0.0036), the overlapping estimator is downward-biased for any horizon, but for a large value of stationary variance, σ_ε^2 (= 0.0081), the estimator is upward-biased for shorter horizons, N (=12,24), and downward-biased for longer horizons, N (=60,84,120). Finally, for any horizon, the weighted average estimator is downward-biased for a small value of stationary variance, σ_ε^2 (= 0.0036),and upward-biased for a large value of stationary variance, σ_ε^2 (= 0.0081).

The magnitude of the bias is the largest for the geometric average. In addition, observe that for the smaller value of stationary variance, σ_ε^2 (= 0.0036), the arithmetic average has the least bias for shorter horizons, N (= 12,24), and the overlapping average the least bias for longer horizons, N (= 60,84,120). For the large value of stationary variance, σ_ε^2 (= 0.0081), and any horizon, the weighted and overlapping averages have less bias than the arithmetic and geometric averages. Overall, the geometric average is the most efficient estimator, and the overlapping average is the least efficient. The weighted average is consistently more efficient than the arithmetic and overlapping averages.

If we compare both Panel A's in Tables 1 and 2, we see that the arithmetic and geometric averages are more upward- and less downward-biased, respectively, and that both averages are less efficient. This represents the combined effect of a small negative autocorrelation (ρ = -0.05) and time-varying variance (σ_ε^2 = 0.0036),which is greater than that of σ_ε^2 alone. Moreover, although the bias for all estimators increases with N, the weighted average is not only the least biased, but is also more efficient than the overlapping average.

Similarly, if we compare Panels A and B of Table 2, introducing σ_ε^2 (= 0.0045) to a small negative autocorrelation (ρ = -0.05) and time-varying variance (σ_ε^2 = 0.0036) magnifies the magnitude of bias for all estimators. The overlapping average is the least biased, but least efficient, estimator. The weighted average is only slightly more biased, but is more efficient than the overlapping average.

Finally, the relative impact of σ_ε^2 and σ_η^2 is evident when we compare Panels B and C of Table 2. When $\sigma_\eta^2 > \sigma_\varepsilon^2$, the weighted average contains consistently smaller biases than when $\sigma_\eta^2 < \sigma_\varepsilon^2$, and its efficiency improves as N increases. Although the overlapping average is still the least biased, it is also the least efficient estimator. The weighted average is only slightly more biased, but is more efficient, than the overlapping average.

In general, the direction and magnitude of the biases reported in Table 2 are also observed in Table 3. In the majority of the cases reported in Table 3, however, the weighted average is the least biased of all estimators, although this improvement is achieved at the expense of efficiency. If we compare Panels A and C, we also

Table 1. Simulation Results in the Absence of Autocorrelation and Time-Varying Variance, $\rho = 0$ and $\sigma_\eta^2 = 0$

Monthly benchmark return is 1%. Horizon is stated in the number of months. Wt. Ave. is the horizon-weighted average of the arithmetic and geometric averages. Overlap is the overlapping average.

Panel A. $\rho = 0,\ \sigma_\eta^2 = 0, \sigma_\varepsilon^2 = 0.0036$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	Fractiles 0.50	Fractiles 0.95
Arithmetic	12	1.1268	1.1254	0.0507	1.0427	1.1246	1.2076
Geometric			1.1018	0.0499	1.0209	1.1013	1.1831
Wt. Ave.			1.1243	0.0507	1.0417	1.1237	1.2064
Overlap			1.1251	0.0516	1.0427	1.1248	1.2090
Arithmetic	24	1.2697	1.2691	0.1146	1.0872	1.2648	1.4582
Geometric			1.2165	0.1104	1.0422	1.2128	1.3998
Wt. Ave.			1.2640	0.1142	1.0831	1.2604	1.4526
Overlap			1.2657	0.1191	1.0786	1.2610	1.4682
Arithmetic	60	1.8167	1.8422	0.4198	1.2325	1.7990	2.5677
Geometric			1.6575	0.3796	1.1088	1.6198	2.3181
Wt. Ave.			1.7966	0.4098	1.2036	1.7567	2.5050
Overlap			1.8022	0.4725	1.1562	1.7383	2.6531
Arithmetic	84	2.3067	2.3858	0.7693	1.3400	2.2752	3.7442
Geometric			2.0580	0.6672	1.1556	1.9645	3.2448
Wt. Ave.			2.2719	0.7337	1.2796	2.1701	3.5650
Overlap			2.2851	0.8909	1.1991	2.1236	3.9425
Arithmetic	120	3.3004	3.5698	1.6822	1.5190	3.2362	6.5931
Geometric			2.8912	1.3714	1.2295	2.6239	5.3736
Wt. Ave.			3.2319	1.5270	1.3830	2.9328	5.9712
Overlap			3.2528	1.9440	1.2160	2.7965	6.8591

Panel B. $\rho = 0,\ \sigma_\eta^2 = 0, \sigma_\varepsilon^2 = 0.0081$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	Fractiles 0.50	Fractiles 0.95
Arithmetic	12	1.1268	1.1306	0.0760	1.0079	1.1284	1.2583
Geometric			1.0774	0.0730	0.9599	1.0745	1.2022
Wt. Ave.			1.1281	0.0758	1.0059	1.1261	1.2556
Overlap			1.1283	0.0780	1.0047	1.1260	1.2605
Arithmetic	24	1.2697	1.2839	0.1727	1.0159	1.2734	1.5833
Geometric			1.1662	0.1581	0.9214	1.1544	1.4452
Wt. Ave.			1.2726	0.1713	1.0071	1.2624	1.5697
Overlap			1.2703	0.1791	0.9944	1.2607	1.5759
Arithmetic	60	1.8167	1.9316	0.6610	1.0403	1.8298	3.1544
Geometric			1.5195	0.5241	0.8149	1.4320	2.5107
Wt. Ave.			1.8299	0.6269	0.9857	1.7356	2.9926
Overlap			1.8074	0.6846	0.8913	1.6954	3.1078
Arithmetic	84	2.3067	2.5929	1.2706	1.0569	2.3301	4.9944
Geometric			1.8540	0.9167	0.7508	1.6531	3.6284
Wt. Ave.			2.3363	1.1471	0.9532	2.1020	4.5182
Overlap			2.2787	1.2826	0.7824	2.0096	4.7529
Arithmetic	120	3.3004	4.1676	3.0671	1.0823	3.3482	9.9503
Geometric			2.5834	1.9241	0.6640	2.0506	6.3036
Wt. Ave.			3.3788	2.4961	0.8798	2.7156	8.1821
Overlap			3.2201	2.7834	0.6314	2.4351	8.7221

Table 2. Simulation Results with a Small Autocorrelation ρ = -0.05

Monthly benchmark return is 1%. Horizon is stated in the number of months. Wt. Ave. is the horizon-weighted average of the arithmetic and geometric averages. Overlap is the overlapping average.

Panel A. $\rho = -0.05$, $\sigma^2_\eta = 0.036$ $\sigma^2_\varepsilon = 0$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	0.50	0.95
Arithmetic	12	1.1268	1.1269	0.0515	1.0446	1.1237	1.2166
Geometric			1.1032	0.0506	1.0246	1.1003	1.1917
Wt. Ave.			1.1258	0.0515	1.0437	1.1226	1.2156
Overlap			1.1236	0.0527	1.0383	1.1221	1.2165
Arithmetic	24	1.2697	1.2724	0.1171	1.0913	1.2627	1.4801
Geometric			1.2195	0.1125	1.0499	1.2107	1.4201
Wt. Ave.			1.2674	0.1167	1.0872	1.2574	1.4748
Overlap			1.2621	0.1216	1.0743	1.2546	1.4707
Arithmetic	60	1.8167	1.8556	0.4393	1.2440	1.7918	2.6651
Geometric			1.6687	0.3962	1.1294	1.6127	2.4032
Wt. Ave.			1.8095	0.4286	1.2159	1.7476	2.6018
Overlap			1.7869	0.4676	1.1393	1.7179	2.6344
Arithmetic	84	2.3067	2.4123	0.8214	1.3575	2.2626	3.9446
Geometric			2.0793	0.7102	1.1858	1.9524	3.4127
Wt. Ave.			2.2966	0.7826	1.2986	2.1572	3.7665
Overlap			2.2608	0.8839	1.1510	2.1064	4.0036
Arithmetic	120	3.3004	3.6361	1.8669	1.5475	3.2106	7.1027
Geometric			2.9415	1.5153	1.2756	2.6007	5.7753
Wt. Ave.			3.2902	1.6915	1.4119	2.9204	6.4632
Overlap			3.2330	1.9575	1.1754	2.7698	6.8499

Panel B. $\rho = -0.05$, $\sigma^2_\eta = 0.036$, $\sigma^2_\varepsilon = 0.0045$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	0.50	0.95
Arithmetic	12	1.1268	1.1319	0.0748	1.0164	1.1283	1.2568
Geometric			1.0786	0.0720	0.9662	1.0763	1.1971
Wt. Ave.			1.1294	0.0747	1.0143	1.1259	1.2544
Overlap			1.1278	0.0771	1.0077	1.1238	1.2610
Arithmetic	24	1.2697	1.2867	0.1713	1.0331	1.2732	1.5796
Geometric			1.1686	0.1571	0.9335	1.1585	1.4330
Wt. Ave.			1.2754	0.1669	1.0239	1.2617	1.5668
Overlap			1.2720	0.1819	1.0056	1.2590	1.6056
Arithmetic	60	1.8167	1.9412	0.6685	1.0847	1.8290	3.1359
Geometric			1.5266	0.5307	0.8419	1.4446	2.4583
Wt. Ave.			1.8388	0.6343	1.0243	1.7300	2.9745
Overlap			1.8159	0.7385	0.9271	1.6760	3.1844
Arithmetic	84	2.3067	2.6111	1.3023	1.1206	2.3285	4.9536
Geometric			1.8663	0.9401	0.7859	1.6736	3.5227
Wt. Ave.			2.3524	1.1760	1.0025	2.0926	4.4684
Overlap			2.3005	1.4391	0.8698	1.9396	4.7906
Arithmetic	120	3.3004	4.2146	3.2132	1.1767	3.3451	9.8342
Geometric			2.6119	2.0128	0.7088	2.0869	6.0431
Wt. Ave.			3.4166	2.6141	0.9468	2.6988	7.9694
Overlap			3.3191	3.4287	0.7108	2.3538	8.5702

Table 2. Simulation Results with a Small Autocorrelation ρ = -0.05 (*Continued*)

Panel C. $\rho = -0.05$, $\sigma_\eta^2 = 0.0045$ $\sigma_\varepsilon^2 = 0.0036$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	0.50	0.95
Arithmetic	12	1.1268	1.1306	0.0749	1.0085	1.1289	1.2550
Geometric			1.0779	0.0720	0.9603	1.0771	1.1963
Wt. Ave.			1.1282	0.0747	1.0064	1.1265	1.2522
Overlap			1.1266	0.0779	0.9985	1.1242	1.2583
Arithmetic	24	1.2697	1.2839	0.1701	1.0172	1.2744	1.5750
Geometric			1.1670	0.1559	0.9223	1.1602	1.4312
Wt. Ave.			1.2727	0.1687	1.0084	1.2632	1.5609
Overlap			1.2689	0.1828	0.9850	1.2568	1.5954
Arithmetic	60	1.8167	1.9297	0.6472	1.0435	1.8333	3.1133
Geometric			1.5206	0.5141	0.8168	1.4500	2.4503
Wt. Ave.			1.8287	0.6141	0.9896	1.7368	2.9461
Overlap			1.8123	0.7192	0.8688	1.6657	3.1331
Arithmetic	84	2.3067	2.5865	1.2395	1.0614	2.3363	4.9036
Geometric			1.8538	0.8962	0.7533	1.6824	3.5067
Wt. Ave.			2.3320	1.1197	0.9580	2.1085	4.4085
Overlap			2.2913	1.3224	0.7811	1.9445	4.7278
Arithmetic	120	3.3004	4.1422	2.9827	1.0888	3.3611	9.6930
Geometric			2.5764	1.8779	0.6672	2.1025	6.0039
Wt. Ave.			3.3626	2.4308	0.8854	2.7379	7.8210
Overlap			3.2489	2.8583	0.6348	2.3838	8.1933

Table 3. Simulation Results with a Large Autocorrelation ρ = -0.25

Monthly benchmark return is 1%. Horizon is stated in the number of months. Wt. Ave. is the horizon-weighted average of the arithmetic and geometric averages. Overlap is the overlapping average.

Panel A. $\rho = -0.25$, $\sigma_\eta^2 = 0.00108$ $\sigma_\varepsilon^2 = 0.00252$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	0.50	0.95
Arithmetic	12	1.1268	1.1262	0.0487	1.0448	1.1266	1.2077
Geometric			1.1021	0.0478	1.0213	1.1024	1.1816
Wt. Ave.			1.1251	0.0486	1.0437	1.1254	1.2065
Overlap			1.1225	0.0494	1.0386	1.1221	1.2011
Arithmetic	24	1.2697	1.2708	0.1097	1.0915	1.2692	1.4585
Geometric			1.2169	0.1054	1.0431	1.2152	1.3962
Wt. Ave.			1.2656	0.1092	1.0869	1.2638	1.4527
Overlap			1.2603	0.1136	1.0728	1.2567	1.4536
Arithmetic	60	1.8167	1.8458	0.3996	1.2447	1.8149	2.5689
Geometric			1.6565	0.3602	1.1113	1.6280	2.3034
Wt. Ave.			1.7991	0.3898	1.2134	1.7704	2.5056
Overlap			1.7895	0.4342	1.1623	1.7311	2.5611
Arithmetic	84	2.3067	2.3891	0.7302	1.3586	2.3035	3.7467
Geometric			2.0536	0.6308	1.1592	1.9784	3.2159
Wt. Ave.			2.2726	0.6955	1.2935	2.1953	3.5686
Overlap			2.2606	0.7989	1.1846	2.1236	3.7313
Arithmetic	120	3.3004	3.5665	1.5918	1.5493	3.2937	6.5994
Geometric			2.8738	1.2908	1.2349	2.6504	5.3055
Wt. Ave.			3.2216	1.4415	1.3994	2.9794	5.9669
Overlap			3.2091	1.6643	1.1889	2.8265	6.4095

Table 3. Simulation Results with a Large Autocorrelation ρ = -0.25 (*Continued*)

Panel B. $\rho = -0.25$, $\sigma^2_\eta = 0.000405$ $\sigma^2_\varepsilon = 0.007695$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	Fractiles 0.50	Fractiles 0.95
Arithmetic	12	1.1268	1.1299	0.0785	1.0006	1.1268	1.2676
Geometric			1.0768	0.0756	0.9512	1.0737	1.2076
Wt. Ave.			1.1275	0.0783	0.9980	1.1244	1.2646
Overlap			1.1264	0.0812	0.9936	1.1230	1.2652
Arithmetic	24	1.2697	1.2829	0.1789	1.0011	1.2696	1.6069
Geometric			1.1652	0.1643	0.9049	1.1528	1.4583
Wt. Ave.			1.2715	0.1775	0.9908	1.2584	1.5910
Overlap			1.2679	0.1898	0.9755	1.2511	1.5983
Arithmetic	60	1.8167	1.9326	0.6969	1.0028	1.8162	3.2732
Geometric			1.5208	0.5546	0.7788	1.4267	2.5679
Wt. Ave.			1.8309	0.6615	0.9445	1.7202	3.0817
Overlap			1.8186	0.7458	0.8661	1.6569	3.2862
Arithmetic	84	2.3067	2.6022	1.3673	1.0040	2.3058	5.2596
Geometric			1.8619	0.9902	0.7047	1.6447	3.7447
Wt. Ave.			2.3451	1.2358	0.8964	2.0758	4.6840
Overlap			2.3242	1.4276	0.7842	1.9571	5.1075
Arithmetic	120	3.3004	4.2200	3.4602	1.0057	3.2985	10.7135
Geometric			2.6200	2.1793	0.6066	2.0356	6.5943
Wt. Ave.			3.4233	2.8210	0.8030	2.6675	8.5390
Overlap			3.3601	3.1676	0.6356	2.3754	9.7576

Panel C. $\rho = -0.25$, $\sigma^2_\eta = 0.00243$ $\sigma^2_\varepsilon = 0.00567$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	Fractiles 0.50	Fractiles 0.95
Arithmetic	12	1.1268	1.1294	0.0721	1.0199	1.1252	1.2561
Geometric			1.0753	0.0694	0.9690	1.0721	1.1970
Wt. Ave.			1.1269	0.0719	1.0174	1.1225	1.2533
Overlap			1.1200	0.0738	1.0113	1.1146	1.2504
Arithmetic	24	1.2697	1.2808	0.1641	1.0403	1.2661	1.5779
Geometric			1.1611	0.1505	0.9390	1.1493	1.4329
Wt. Ave.			1.2693	0.1628	1.0296	1.2543	1.5632
Overlap			1.2529	0.1700	1.0132	1.2368	1.5553
Arithmetic	60	1.8167	1.9141	0.6252	1.1038	1.8038	3.1274
Geometric			1.4987	0.4957	0.8545	1.4161	2.4576
Wt. Ave.			1.8115	0.5930	1.0404	1.7044	2.9563
Overlap			1.7524	0.6358	0.9180	1.6407	2.9633
Arithmetic	84	2.3067	2.5532	1.1906	1.1483	2.2839	4.9347
Geometric			1.8140	0.8578	0.8024	1.6276	3.5213
Wt. Ave.			2.2965	1.0745	1.0309	2.0482	4.4316
Overlap			2.1744	1.1431	0.8366	1.9151	4.4332
Arithmetic	120	3.3004	4.0541	2.8088	1.2184	3.2539	9.7808
Geometric			2.4915	1.7562	0.7301	2.0054	6.0396
Wt. Ave.			3.2761	2.2832	0.9765	2.6212	7.8862
Overlap			2.9808	2.3220	0.6750	2.2822	7.5861

Table 3. Simulation Results with a Large Autocorrelation ρ = -0.25 (*Continued*)

Panel D. $\rho = -0.25,\ \sigma^2_\eta = 0.0036\ \sigma^2_\varepsilon = 0.0045$

Estimator	Horizon	Benchmk Return	Average	Standard Error	Fractiles 0.05	0.50	0.95
Arithmetic	12	1.1268	1.1275	0.0709	1.0146	1.1272	1.2492
Geometric			1.0730	0.0684	0.9633	1.0725	1.1877
Wt. Ave.			1.1250	0.0708	1.0125	1.1247	1.2467
Overlap			1.1158	0.0724	1.0008	1.1168	1.2410
Arithmetic	24	1.2697	1.2762	0.1605	1.0295	1.2705	1.5606
Geometric			1.1560	0.1474	0.9280	1.1503	1.4107
Wt. Ave.			1.2646	0.1592	1.0207	1.2593	1.5468
Overlap			1.2446	0.1662	0.9894	1.2401	1.5459
Arithmetic	60	1.8167	1.8947	0.6019	1.0754	1.8196	3.0423
Geometric			1.4809	0.4767	0.8296	1.4190	2.3638
Wt. Ave.			1.7925	0.5707	1.0183	1.7202	2.8760
Overlap			1.7249	0.6193	0.8986	1.6286	2.9045
Arithmetic	84	2.3067	2.5137	1.1352	1.1072	2.3119	4.7477
Geometric			1.7816	0.8146	0.7699	1.6323	3.3347
Wt. Ave.			2.2595	1.0233	0.9959	2.0773	4.2567
Overlap			2.1478	1.1423	0.8072	1.8783	4.4142
Arithmetic	120	3.3004	3.9518	2.6400	1.1565	3.3109	9.2557
Geometric			2.4201	1.6346	0.6883	2.0137	5.5876
Wt. Ave.			3.1891	2.1377	0.9301	2.6705	7.4157
Overlap			2.9632	2.3759	0.6444	2.2599	7.7379

observe that when σ^2_ε and σ^2_η both increase by the same proportion, the weighted average experiences a smaller bias relative to the other three estimators. Furthermore, we see from Panels B and C that a reduction in σ^2_ε that is offset by a corresponding increase in σ^2_η improves the weighted average's efficiency.

The effect of higher negative autocorrelation is evident when we compare Panel D in Table 3 with Panel B in Table 2. Even though we obtain a higher efficiency for all estimators, a higher negative autocorrelation ρ leads to a smaller bias in the arithmetic and weighted averages, but a larger bias for the geometric and overlapping averages. Moreover, although Table 3 shows that the weighted average is the second most efficient estimator, it is overall the least biased when negative autocorrelation, time-varying, and stationary variance components are all present.

III. Concluding Remarks

We show that both the arithmetic and geometric averages are biased estimates of long-run expected returns, and the bias increases with the length of the investment horizons. The existence of negative autocorrelation in long-horizon returns documented by Fama and French (1988a, 1988b), Lo and MacKinlay (1988), and Poterba and Summers (1988) exacerbates the bias. The implication is that without making an adjustment, we are likely to obtain an estimate of long-run expected return (and risk premium) that is either too high or too low, and this can result in an inappropriate decision to reject a good project or accept a bad project.

The horizon-weighted average of the arithmetic and geometric averages, proposed by Blume (1974), is an alternative estimate of long-run expected returns. Our simulation results indicate that in general, the horizon-weighted average contains the least bias. It is also more efficient than other estimators in the presence of negative autocorrelation, time-varying, and stationary variances. This conclusion contrasts with Blume's conjecture that "...if one cannot assume independence of successive one-period relatives or if there is even a slight chance that these relatives are dependent, the simple average of *N*-period relatives would appear preferable to the nonlinear estimators which, even under ideal conditions, yield only a modest increase in efficiency." ■

References

Blume, M.E., 1974, "Unbiased Estimates of Long-Run Expected Rates of Return," *Journal of the American Statistical Association* (September), 634-638.

Brealey, R.A. and S.C. Myers, 1991, *Principles of Corporate Finance*, New York, NY, McGraw-Hill.

Brennan, M.J., 1997, "The Term Structure of Discount Rates," *Financial Management* (Spring), 81-90.

Brigham, E.F., D.K. Shome, and S.R. Vinson, 1985, "The Risk Premium Approach to Measuring a Utility's Cost of Equity," *Financial Management* (Spring), 33-45.

Butler, J.S. and B. Schachter, 1989, "The Investment Decision: Estimation Risk and Risk-Adjusted Discount Rates," *Financial Management* (Winter), 13-22.

Carleton, W.T. and J. Lakonishok, 1985, "Risk and Return on Equity: The Use and Misuse of Historical Estimates," *Financial Analysts Journal* (January-February), 38-47.

Copeland, T., T. Koller, and J. Murrin, 1994, *Valuation: Measuring and Managing the Value of Companies*, Second Edition, New York, NY, John Wiley & Sons.

Fama, E.F., 1996, "Discounting Under Uncertainty," *Journal of Business* (October), 415-428

Fama, E.F. and K.R. French, 1988a, "Permanent and Transitory Components of Stock Prices," *Journal of Political Economy* (April), 246-273.

Fama, E.F. and K.R. French, 1988b, "Dividend Yields and Expected Stock Returns," *Journal of Financial Economics* (October), 3-25.

Harris, R.S., 1986, "Using Analysts' Growth Forecasts to Estimate Shareholder Required Rates of Return," *Financial Management* (Spring), 58-67.

Harris, R.S. and F.C. Marston, 1992, "Estimating Shareholder Risk Premia Using Analysts' Growth Forecasts," *Financial Management* (Summer), 63-70.

Lo, A.W. and A.C. MacKinlay, 1988, "Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test," *Review of Financial Studies* (Spring), 41-66.

Maddox, F.M., D.T. Pippert, and R.N. Sullivan, 1995, "An Empirical Study of Ex-Ante Risk Premiums for the Electric Utility Industry," *Financial Management* (Autumn), 89-95.

Poterba, J.M. and L.H. Summers, 1988, "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics* (October), 27-59.

Timme, S.G. and P.C. Eisemann, 1989, "On the Use of Consensus Forecasts of Growth in the Constant Growth Model: The Case of Electric Utilities, *Financial Management* (Winter), 23-35.

NBER WORKING PAPER SERIES

THE CROSS-SECTION OF HURDLE RATES FOR CAPITAL BUDGETING: AN EMPIRICAL ANALYSIS OF SURVEY DATA

Ravi Jagannathan
Iwan Meier
Vefa Tarhan

Working Paper 16770
http://www.nber.org/papers/w16770

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2011

We thank Michael Fishman, Robert Korajczyk, Deborah Lucas, Robert McDonald, David Matsa, Mitchell Petersen, Artur Raviv, Ernst Schaumburg, Michael Sher, Paul Spindt, and Timothy Thompson for helpful comments. We also thank the focus group participants William A. Colaianni, James J. Cowhey, Pavel A. Dorosevich, Gaea Gomez, Art Mollenhauer, and Don Porth for their input to improve the design of the survey. This research is sponsored by the Zell Center for Risk Research and we are indebted to its Director and Dean Emeritus of the Kellogg School of Management Donald Jacobs for his support and advice. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Cross-Section of Hurdle Rates for Capital Budgeting: An Empirical Analysis of Survey Data
Ravi Jagannathan, Iwan Meier, and Vefa Tarhan
NBER Working Paper No. 16770
February 2011
JEL No. G12,G3,G31

ABSTRACT

Whereas Poterba and Summers (1995) find that firms use hurdle rates that are unrelated to their CAPM betas, Graham and Harvey (2001) find that 74% of their survey firms use the CAPM for capital budgeting. We provide an explanation for these two apparently contradictory conclusions. We find that firms behave as though they add a hurdle premium to their CAPM based cost of capital. Following McDonald and Siegel (1986), we argue that the hurdle premium depends on the value of the option to defer investments. While CAPM explains only 10% of the cross-sectional variation in hurdle rates across firms, variables that proxy for the benefits from the option to wait for potentially better investment opportunities explain 35%. Estimates of our hurdle premium model parameters imply an equity premium of 3.8% per year, a figure that is essentially the same as that reported in the survey by Graham and Harvey (2005). Consistent with our model, growth firms use a higher hurdle rate when compared to value firms, even though they have a lower cost of capital.

Ravi Jagannathan
Kellogg Graduate School of Management
Northwestern University
2001 Sheridan Road
Leverone/Anderson Complex
Evanston, IL 60208-2001
and NBER
rjaganna@northwestern.edu

Iwan Meier
HEC Montreal
3000, chemin de la Cote-Sainte-Catherine
Montreal (Quebec) H3T 2A7
Canada
iwan.meier@hec.ca

Vefa Tarhan
Loyola University Chicago
Graduate School of Business
1 E. Pearson St., Maguire Hall
Chicago, IL 60611
v-tarhan@luc.edu

I. Introduction

According to a survey by Womack and Zhang (2005) 38% of the total class time of the core finance courses at major MBA programs is devoted to capital budgeting decisions, computing net present value (NPV) and cost of capital. The tuition fees of the top 30 ranked MBA programs by Business Week total 1.6 billion in 2010. Thus, it appears that business schools generate considerable revenues in return for an education of the principles of corporate finance.

A number of studies document that when computing the net present value of a project, the majority of firms discount future cash flows using hurdle rates that reflect their weighted average cost of capital (WACC) (e.g. Bierman (1993), Bruner, Eades, Harris, and Higgins (1998)) and thus indeed follow the standard approach as taught in MBA programs. Additionally, surveys over the past four decades report that since Sharpe (1964) and Lintner (1965) introduced the capital asset pricing model (CAPM), firms have increasingly adopted its framework to determine their cost of equity. In fact, Graham and Harvey (2001) find that three out of four CFOs rely on the CAPM. Thus, in spite of its criticism in the literature, it appears that CAPM is widely used in practice.

In a survey that we conducted, we ask firms what they use for hurdle rates in their capital budgeting decisions. Since we know the identity of the respondents to our survey, we can match firms with fundamental Barra betas and data from Compustat and CRSP to compute their WACC. We document that hurdle rates firms use in practice exceed their computed WACC, i.e., firms add a hurdle premium to their cost of capital.[2] The hurdle premium is substantial and

[2] In a roundtable discussion on capital structure and payout policy, Jon Anda from the investment banking division of Morgan Stanley states that “my feeling is that a large number of companies today are using hurdle rates that are well above their weighted average cost of capital” (see Smith, Ikenberry, Nayar,

accounts, on average, for about half of the hurdle rate. We also find that the presence of the hurdle rate premium is independent of whether the cost of equity is inferred from the single-factor CAPM, the Fama-French three-factor model, or computed by making assumptions about the size of the equity premium.

Poterba and Summers (1995) also find hurdle rates to be on the high side. They document an average real hurdle rate of 12.2%, at a time when the long-term inflation expectation was around 5%. They argue that the hurdle rates are higher than both the cost of debt and the cost of equity of firms in their survey sample. Moreover, they find that hurdle rates are not related to CAPM betas. How is it that firms claim to use CAPM and WACC, and yet their hurdle rates are not systematically related to beta, and are also much higher than firms' computed WACC? In this paper, we provide an explanation based on high growth prospects that make options to wait for better investment opportunities valuable when firms cannot undertake all positive net present value projects due to limited availability of organization capital. We propose a model that explains the determinants of hurdle rates and at the same time produces results that are consistent with the previous survey findings that firms indeed use CAPM and WACC. While WACC is an important determinant of the hurdle rate, it is not its only component.

The key to our model is that firms with high growth opportunities incorporate a premium associated with an option to wait to their hurdle rates. This insight is provided by McDonald and Siegel (1986). In addressing the investment timing problem they observe that investing in a current positive NPV project is irreversible, while the decision to defer the investment is reversible. They argue that the correct decision is reached by comparing the NPV of the current project with the NPV (as of the current period) that can be obtained if the investment is made in

---

Anda and McVey (2005, p. 52)). Additionally, Antill and Arnott (2004) claim that the hurdle rates of the twelve oil companies they examine exceed their WACC.

the future. This option to wait is valuable to growth firms since it may enable them to take future projects that possibly have higher NPVs than the (positive) NPV projects they have in the current period. Such firms may behave in this manner due to managerial and other human capital constraints in the current period. At the same time, these firms may fear facing adverse conditions in capital markets in the future when highly valuable projects materialize. We hypothesize that in order to avoid this possibility, in the current period these firms would put themselves in a financial position to undertake the highly valuable projects that they may encounter in the future. In other words, current period financial flexibility concerns are likely to be important for firms with high growth prospects. This suggests that firms with high cash reserves would have high hurdle premia.

It is important to emphasize that the option to wait for future projects that have higher expected values than the current period positive NPV investments, is different from a traditional real option attached to a specific project. If firms consider a project to be strategic, then they judge that investing in such a project has the potential to generate additional future cash flows that are currently not incorporated in the valuation of the project. For instance, the first investment in a foreign country might pave the way for other positive NPV projects in the future. In such cases, firms could use decision trees to incorporate future cash flows. However, survey evidence shows that firms often incorporate such real options associated with strategic projects by using *lower* hurdle rates (e.g. Bruner, Eades, Harris, and Higgins (1998)). In contrast, firms that are in a position to take advantage of options to wait would use *higher* discount rates in screening projects in the current period. When firms uncover a new positive NPV project, they have to decide whether to take it or to wait for a potentially better future opportunity. The decision can be characterized as an optimal stopping problem. Given a number of future projects with a distribution of NPVs, where only the approximate distribution is known, the firm has to decide whether it is optimal to take a currently available positive NPV project or to wait for a better opportunity. The average expected NPV of the future projects depends on the growth prospects of the industry, while the dispersion is driven by the riskiness of the industry. This suggests that both recent period industry returns and the unpredicted fraction of industry returns would be positively correlated with hurdle premia.

If firms do not face any constraints and capital markets are well functioning, every positive NPV project in the current period would be funded. However, firms with high-growth prospects may not want to take every positive NPV project in the current period since they may find even better opportunities in the future. For this reason, firms with high growth prospects may pass up on some good current period projects by using hurdle rates that exceed their WACC. The difference between the hurdle rates they use and their computed WACC would represent the premium associated with the option to wait. The option to wait is more valuable to firms with high growth prospects who operate in an environment where the NPV distribution of possible projects are likely to have a wider dispersion than those faced by mature firms.

Jagannathan and Meier (2002) argue that organizational and managerial constraints may represent another reason why firms with valuable options to wait, i.e., firms with ample growth opportunities, would use higher hurdle rates. Since in corporate finance growth is about the sales variable, we use sales growth per employee as a proxy to measure the presence of managerial constraints. Jagannathan and Meier (2002) use a real options framework that builds on McDonald (1999) to demonstrate that depending on growth prospects and the dispersion of the NPV distribution of future projects, the hurdle rate premium can be substantial. The optimal solution for when to take a positive NPV project can be found using the classical stopping problem (also known as parking or secretary problem).

In this paper we make several contributions. First, we document that there is a hurdle rate premium. Second, we develop a model where hurdle rates have two components: WACC, and variables that represent firm characteristics that proxy for the value of the option to wait. The model enables us to estimate the equity premium, along with the loadings on firm characteristics. Our estimate for the equity premium is identical to the figure found by Graham and Harvey (2005) from a survey they conducted at about the same date of our survey (3.8% in both cases). Also, unlike Poterba and Summers (1995) who do not find a significant relation between historical beta and hurdle rates, we find that fundamental beta is positively correlated with hurdle rates in our sample. Third, we find that actual WACC constitutes about half of the value of the average hurdle rate, while the remaining half of the variation in hurdle rates can be explained by variables that proxy for the value of options to wait. Furthermore, we find that dispersion of hurdle premia is three times the dispersion of WACC. Fourth, as hypothesized, financial flexibility considerations play an important role: firms with high levels of cash use higher hurdle rates. Fifth, we find that firms with high growth opportunities use higher hurdle rates (they load negatively on the Fama-French HML factor) even though their stocks earn lower returns. Additionally, the R-square obtained from the estimation of the market model for firms that are in the same industry (2 digit SIC) as the sample firms, is negatively correlated with hurdle rates. Finally, we confirm Jagannathan and Meier (2002) that managerial and organizational constraints play an important role in investment decisions: the estimate for the sales growth per employee variable is positive and is significantly related to hurdle rates.

The remainder of the paper proceeds as follows. Section II describes the experimental design and data. Section III discusses survey results. Section IV presents the model. Empirical findings are discussed in Section V. Finally, Section VI concludes.

II. Experimental Design and Data

Figure 1 gives an overview of the results from the survey literature. Apparently, starting in the 1990s an overwhelming fraction of firms use discounted cash flow (DCF) methods. Similarly, starting in the 1980s the use of WACC and CAPM has increased dramatically. Interestingly, the use of company-wide hurdle rates has not declined over time. In order to examine how hurdle rates are related to cost of capital and to test whether the hurdle premium is related to options to wait, we combine survey questions with archival data from Barra, CRSP, and Compustat. Hurdle rates cannot be observed directly in archival databases and require a survey. Besides Poterba and Summers (1995), to the best of our knowledge, ours is the only survey on hurdle rates that knows the identity of the respondents. Combining survey data with financial databases enables us to examine the determinants of the hurdle premium.

The survey was completed by the CFOs of 127 companies in October 2003. A high percentage of the respondents reveal their identity (83.5%). Almost all surveys are filled out completely and there is no decline in the number of responses towards the end of the four-page questionnaire. Survey data has strengths and weaknesses. Surveys are the only way to obtain hurdle rates used in practice. On the downside, surveys do not produce as many observations as databases such as Compustat. Additionally, if survey questions are not phrased carefully, tests based on survey responses could be misleading. In designing the survey, we carefully followed the advice of experts in the fields of psychology and marketing.[3] We designed the questions in such a way that we minimize the use of technical terms and names of models that are taught in a typical MBA course. For example, we avoid terms such as "cost of capital" and "CAPM" in our questionnaire. Instead, the survey participants were asked questions on their "hurdle rates." It is a well documented observation in psychology, known as the social desirability hypothesis (see e.g. Singer and Presser (1989)), that respondents to surveys tend to try to please the conductor of the survey by providing the answers they think the survey's author expects. Therefore, in designing the survey questions we tried to avoid using technical terms. The input from numerous finance academics helped to further improve the content of the questions. Additionally, in order to test the survey with practitioners, we invited six CFOs from the Chicago area to a focus group meeting. After filling out the survey, we discussed each question to assure that the wording was not ambiguous. The survey was sent out together with a cover letter from the Dean Emeritus of the Kellogg School of Management, Donald Jacobs, along with a postage-paid return envelope to a total of 4,600 CFOs of U.S. companies listed in the Compustat name file. We asked the participants to return the questionnaire within ten days. A week after the initial mailing we sent a follow-up mailing to remind the potential participants.

We have some evidence that the surveys were actually filled out by CFOs as we received a number of e-mails from the CFOs requesting an advance copy of the survey results. In addition, many respondents provided elaborate comments to open questions. The survey responses appear to be accurate. For example, when we compare self-reported sales figures with the numbers retrieved from Compustat, we find that a reassuring 92.3% of the respondents checked the correct sales range.

[3] Among others, Gillman (2000) and Morgan (1988) provide guidelines for surveys and focus group meetings.

Table I compares the breakdown by industry, hurdle rate statistics, and the use of CAPM/WACC to previous surveys. Except for the fact that our sample excludes financial firms, the distribution across industries are comparable to other surveys.[4] In all surveys and in the Compustat sample manufacturing exceeds 50% of the sample. In our survey manufacturing firms make up 66% of the sample.[5] Firms in the wholesale and retail sectors make-up 11.6% of our sample, while mining and construction and transportation/communication sectors are equally represented (10.7% each). In Table I, in the Compustat sample we compute the weights by including only the sectors that we have mailed our survey to. While our sample size is a third of Graham and Harvey (2001), we know the identity of 106 out of 127 firms and are able to match 93 firms with Barra and CRSP/Compustat. Summary statistics of the hurdle rates in our survey match those of Poterba and Summers (1995), and the use of WACC is comparable to Bruner, Eades, Harris and Higgins (1998). Other characteristics (not reported in the table) of survey firms are as follows: Firm size measured by (self-reported) sales is below $100 million for 35.2% of the companies and 31.2% of the responding firms report sales in excess of $1 billion. The majority of the firms (72.0%) have multiple product lines.

Table II compares the characteristics of the 93 responding firms for which we can match Compustat data and the Compustat sample of firms. Based on mean values it appears that the two samples are similar except for four variables. Survey firms have higher market value of

[4] Financial firms account for 15% of the respondents in Graham and Harvey (2001). We exclude all finance and insurance companies with the major SIC code in the ranges 6000-6499, 6700-6799; and utilities (4900-4999) in order to exclude regulated firms. We also discard radio and TV broadcasting, cable, and other pay TV services (4840-4949), as these firms might be driven by non-commercial interests, e.g. religious radio stations. Finally, we exclude health, education, social services, and museums (7200+).

[5] In a number of surveys the fraction of manufacturing firms is even more pronounced. For example, in Gitman and Mercurio (1982) this ratio is 93.8%, while in Gitman and Forrester (1977) it is 74%.

assets (even though the mean book values are not statistically different).[6] They also have a higher ratio of cash-to-book assets. The book assets of survey firms also generate higher operating profits. Finally, the survey firms are more capital expenditure intensive. Given that manufacturing firms are somewhat overrepresented in the survey sample, this is not surprising. Other important financial variables, such as, leverage ratio (total debt divided by book value of assets), current ratio, total asset turnover, and return on book equity are comparable.

III. Survey Findings

Since Poterba and Summers (1995) is the only other study where the identity of survey firms are known, it is useful to compare their findings with ours. They comment that hurdle rates in their sample appear to be too high compared to cost of capital. We confirm this observation for our survey sample. As can be seen in Panel B of Table I, while our average nominal hurdle rate of 14.8% is somewhat lower than their implied nominal rate of 17.8% (12.2% real and inflation expectation of 5%), their median rate that we construct from their data is 10% in real terms and 15.5% in nominal terms, which is very close to our median of 15%. The standard deviations of the two samples are also similar. Taken together, these stylized facts suggest that, the real discount rates used by firms have not changed much even though the two surveys were conducted 14 years apart.

As we discussed in Section I, Poterba and Summers (1995) find no relation between hurdle rates and systematic risk as measured by historical betas. This is puzzling since it appears to contradict the evidence from the survey literature that firms use CAPM along with WACC to

[6] For variable definitions, see the caption of Table II.

compute cost of equity and cost of capital.[7] For this reason, we repeat the exercise of Poterba and Summers (1995) for our sample by regressing self-reported hurdle rates on the same set of financial variables they use. Figure 2 illustrates the results from kernel-weighted local polynomial regressions for our sample firms. We use a non-parametric kernel method to minimize the effect of outliers and to account for the presence of non-linearities. The figures suggest that the relation between hurdle rates and all the explanatory variables, except for the current ratio, are essentially flat. Even in the case of the current ratio, it appears that the relationship is dominated by some firms which have high current ratios and high hurdle rates.

Table III summarizes the bivariate OLS coefficients for the same set of explanatory variables using the two survey samples in question. The table indicates that the similarity between the two surveys extends beyond having similar summary statistics: The regression coefficients obtained from the two samples are also comparable. In neither of the samples the explanatory financial variables, except for current ratio, is related to hurdle rates.[8] In our sample, even the current ratio turns out to be insignificant (*p*-value of 0.12) once the two firms with current ratios in excess of 10 (the cutoff rate as e.g., in Cleary (1999)) are excluded from the analysis. Using fundamental beta from Barra instead of historical beta (estimated from five years of monthly data) slightly increases the coefficient estimates for both the full sample and manufacturing sector sub samples. In the case of manufacturing firms, the positive relationship between fundamental beta and hurdle rates cannot be rejected at the 10% level. Given that

[7] Graham and Harvey (2001) find that three out of four CFOs use CAPM and 85% of the firms that Bruner, Eades, Harris, and Higgins (1998) interview use WACC.

[8] The coefficients for total equity return have the same sign as in Poterba and Summers (1995) but differ in size. Over the 10 years preceding the survey date (1993-2003) the S&P 500 index increased by 138%, whereas over the period 1980-1990 considered in Poterba and Summers (1995) the index increased by 227.4%.

historical beta coefficients for individual firms from an index model tend to have low R-squares, and hence provide noisy estimates, in the remainder of this paper we rely on fundamental betas.

The bar chart in Panel A of Figure 3 shows what survey participants use as their hurdle rate. Of the 117 firms that responded to the question on what their hurdle rate represents, a significant percentage of the CFOs (71.8%) claim that the hurdle rate they use is their weighted average cost of capital (WACC). In the case of 7 firms (6.0%), the hurdle rate represents their cost of levered equity, while for 9 firms (7.7%) it reflects their unlevered cost of equity. For 17 firms (14.5%), the hurdle rate falls into the "other" category.[9] The widespread use of WACC in our sample is consistent with the findings of Gitman and Vandenberg (2000), Bruner, Eades, Harris, and Higgins (1998), and Bierman (1993) who report that even larger fractions of firms use WACC. As displayed in Figure 1, similar to the increased use of discounted cash flow (DCF) techniques and CAPM, the use of WACC has also increased over time. For example, in a survey conducted 30 years ago, Petty, Scott, and Bird (1975) document that only 30% of the Fortune 500 firms that responded to their survey use WACC. In contrast, in later surveys, such as the one by Bruner, Eades, Harris, and Higgins (1998), this figure is over 80%.

In the survey, we ask the participants for the nominal hurdle rate that they have used for a typical project during the two years preceding the survey date. Since hurdle rates represent firms' WACC by a substantial margin, in the case of the small number of firms which use their levered or unlevered cost of equity, we convert their hurdle rates to their WACC equivalents. In doing this, we use data on debt/asset ratios and tax rates from Compustat, and cost of debt information we obtain from the survey responses. The details of how we convert the 16 levered/unlevered cost equity responses to their WACC equivalents are described in the

[9] This category consists of firms which provide their hurdle rates without indicating what they represent.

Appendix. Panel B of Figure 3 displays the distribution of hurdle rates (WACC and its equivalents sample) used by survey firms.

Panel A of Table IV displays summary statistics on self-reported hurdle rates for various samples: The sample of firms which indicated what their hurdle rates are (all respondents), the sample of firms which indicated what their hurdle rates are, but did not state what they represent (the "other" category), the WACC equivalent sample (those who marked WACC as their hurdle rates plus the WACC of the levered/unlevered cost of equity subsample), finally, the sample for which we can match with Compustat, CRSP, and Barra data bases. In the next section we analyze the determinants of the hurdle premium using this last sample. The summary statistics for all respondents in Panel A show that the mean hurdle rate is 14.8% in nominal terms (the median is 15.0%). In this sample none of the numbers is less than 5%, and the maximum hurdle rate used is 40%. Furthermore, the skewness coefficient of 1.7 indicates that the distribution is fairly symmetric, and the kurtosis coefficient of 9.6 confirms that the distribution is centered around the mean and median. Adjusting for the average realized inflation of 2.2% during the two years preceding the survey date (January 2001 to December 2003) produces an average real hurdle rate of 12.3%, which is essentially same as the 12.2% real hurdle rate reported by Poterba and Summers (1995). The mean and median of the WACC equivalent sample are 14.1%, and 14.0%, respectively. Next, we look at those firms for which we can match Barra betas and CRSP/Compustat data. Again, the means and medians are very close to those for the full sample. Thus, sample selection does not change the characteristics of the hurdle rate distribution.

Panel B of Table IV reports the industry composition of firms in each sample. Comparing the first (full) sample, and the sample we use in our tests (the last sample), suggest that there is no industry related bias. Examination of Panel C leads to the conclusion that other than the standard deviation for the manufacturing firms (which is somewhat higher), the summary statistics across industries are similar.

IV. Modeling Hurdle Rates

In order to test our hypothesis that firms screen projects by adding a hurdle premium to their cost of capital and to explore the determinants of the premium, we propose a model that explains hurdle rates by the weighted average cost of capital plus a linear combination of firm characteristics that are likely to be related to the value of the option to wait. We use nonlinear least squares estimation to solve simultaneously for the equity premium that firms use to compute their cost of equity and WACC, and the loadings on firm characteristics that proxy for the value of the option to wait.

(1) $$\text{Hurdle} = WACC + a + \sum_{j=1}^{k} b_j \text{Char}_j + \varepsilon$$

where,

(2) $$\text{WACC} = \frac{D}{D+E} r_D (1 - \text{Tax}) + \frac{E}{D+E} r_E$$

(3a) $$r_E = r_F + \beta_{MKT} P_{MKT}$$

(3b) $$r_E = r_F + \beta_{MKT} P_{MKT} + \beta_{SMB} P_{SMB} + \beta_{HML} P_{HML}$$

In the CAPM specification (3a) we use the fundamental Barra beta. In the three factor specification (3b), in order to get the beta coefficients for SMB and HML we first subtract $\beta_{Barra} r_{MKT}$ from monthly returns to get a time series of residual returns in excess of what can be explained by market returns.

(4) $r_E - r_F + \beta_{Barra} r_{MKT}$

We then regress five years of monthly residual returns prior to the survey date on the returns of the factor-mimicking portfolios for SMB and HML.

The firm characteristics variables that we include in our model are: cash-to-assets ratio, average industry stock returns during the five years prior to the survey date, the average R-squares of the market model in the industry that the firm belongs (again using 5 years worth of monthly observations), sales growth per employee, and Altman's Z- score.

Due to tax related costs of holding excess cash and agency costs, we expect growth firms to have high cash-to-assets ratio. There is ample evidence that shareholders force non-growth firms to distribute their cash holdings. For example, Nohel and Tarhan (1998) show that firms with low Q ratios improve their operating performance by distributing cash via share repurchases. The value of the option to wait should be higher for high-growth firms, since it may enable these firms to undertake future projects that are more valuable than the positive NPV projects they have in the present period. These firms are likely to screen projects using a hurdle rate that exceeds their WACC. At the same time, due to the possibility that they may face difficulties in the future when valuable projects materialize, they are likely to maintain high financial flexibility in the current period by having a high cash-to-assets ratio. Thus, we expect cash-to-assets to have a positive sign.

Financially healthy firms are likely to have higher growth prospects. Thus, measures of financial health, such as Altman's Z-score, are expected to have a positive estimated coefficient.[10] Systematic risk is also likely to be positively related to hurdle rates. Holding other

[10] For financially unhealthy firms, a measure of how close the firm is to bankruptcy is likely to be positively correlated with hurdle rates. As probability of bankruptcy increases, provided that the firm has time to wait before chapter 11 or liquidation, the higher is the value of option to wait. This represents a lottery type of situation. Rather than accepting a project which has a positive NPV where the NPV is not

firm characteristics constant, fundamental Barra beta will be positively correlated with hurdle rates since it would mean a higher WACC.

Since stock prices reflect anticipated future growth, industries with high past returns are likely to have high growth prospects in the future. The average expected NPV of future projects, in turn, is likely to be positively correlated with the growth prospects of the industry. For this reason, firms that belong to industries with high average returns are likely to have high hurdle premia.

Dispersion of the distribution of future NPVs is driven by the riskiness of the industry. The firm has to decide whether it is optimal to accept a current positive NPV project or wait for a possibly better one by using a hurdle rate with two components – WACC and the hurdle premium. Holding the point estimate of beta constant, the lower is the R-squares of the market model, the wider is the dispersion, thus, the higher is the value of the option for waiting.[11]

Finally, managerial and other human capital constraints will influence hurdle rates in the positive direction. High-growth firms are likely to have high opportunity costs of not waiting for possible better projects in the future due to limited managerial talent. These firms are likely to place a high value on the option to wait. Since in corporate finance the term "growth" concerns the sales variable, we use a categorical variable sales growth per employee to capture human capital constraints.

V. Empirical Findings

---

high enough to materially change the firm's situation, it would be reasonable for the firm to reject the project by using a high hurdle rate in hopes of encountering a project with a high enough NPV that would make a difference in the firm's value.

[11] There is also the possibility that unsystematic risk may also play a role (Goyal and Santa-Clara (2003)). First, managers may feel that shareholders are not fully diversified and price this risk in their hurdle rates. Second, lower R-squares involve a wider confidence around the point estimate for beta and, to be on the safe, side managers may use higher rather than lower hurdle rates when the R-squares is low.

Table V displays the results from various models that we use to determine the relative importance of WACC, and variables related to the option to wait, in explaining the cross-sectional variation in hurdle rates. In Columns 1 and 2 we show the results from estimating (1), (2), for the single factor CAPM (equation 3a), and the Fama-French three factor model (3b), respectively. The 3.8% equity premium estimate obtained from the single factor CAPM is identical to Graham and Harvey (2005), who in a survey they conduct at approximately the same date as our survey, find the average expected equity premium to be 3.8% (median 3.6%).

The cash-to-assets is positively correlated with hurdle rates (at 1% level of significance). Simutin (2010) finds that firms with high cash balances generate higher future stock returns. Based on this finding, he argues that excess cash holdings proxy for high growth opportunities. Since high growth opportunities imply a high valuation for the option to wait, the positive correlation between cash and hurdle rate is as expected.[12]

The dispersion of the distribution of future NPVs is driven by the riskiness of the industry, and since low R-squares obtained from estimating the market model of individual firms in the same industry imply a wider dispersion, the expected correlation between average industry R-squares and hurdle rates is negative. This expectation is confirmed by the highly significant negative coefficient for the R-squares variable. The positive estimate (significant at the 1% level) for the sales growth per employee variable is also as expected. We use this variable as a proxy for managerial and organizational constraints. Growth firms are more likely to find this constraint to be binding. As a result, they would put a high value for the option to wait. The

[12] Opler, Pinkowitz, Stulz, and Williamson (1999) analyze the tradeoff between the benefits and costs of cash holdings. While cash holdings create value by providing financial flexibility to take advantage of future profitable projects, cash holdings also involve tax related costs and agency costs (e.g., by enabling managers to engage in empire building types of activities). In fact, Pinkowitz, Stulz, and Williamson (2006) find that in countries with weak investor protection, cash is discounted at a higher rate. However, in countries with intense shareholder activism (such as the U.S.), benefits of cash exceed its potential costs (especially in the case of growth firms).

positive and highly significant estimate for the variable in question is consistent with this interpretation.

Thus, the three variables discussed above each have the expected sign and are statistically significant. However, even though the other two variables – average industry return, and the financial health of a firm as measured by Altman's Z-score – are, as expected, positively correlated with the value of the option to wait, the financial health variable is not statistically significant when CAPM is used (it is significant at the 5% level when the three-factor model is used). The model estimated in Column 1, explains 45% of the cross-sectional variation in hurdle rates. Furthermore, Panel A of Figure 4 shows the relation between the predicted values of hurdle rates (horizontal axis), and the actual hurdle rates (vertical axis). The 45 degree line in the figure is superimposed. However, when we run a regression of predicted values on actual hurdle rates we obtain a slope that is not statistically different from one (estimated slope coefficient is 0.87 with a standard error of 0.15), and the estimate for the intercept is 0.025 (with a *p-value of* 0.27).

In Table V, Column (2) displays the results from estimating (1), and (2) using the three-factor model (3b). An interesting result is that the estimated loading on the HML factor is negative. The literature finds that value stocks earn *higher returns* than growth stocks. The negative estimated coefficient for the HML factor indicates that growth firms use *higher hurdle rates* than value firms. Thus, while value firms earn higher returns, growth firms expect to earn more on their future projects and use higher hurdle rates. At the same time, the estimated equity premium becomes smaller in this specification. However, the 3.8% equity premium estimate of Column 1 is still within one standard deviation of the estimate for the equity premium in Column 2. The results also show that the SMB loading is unrelated to hurdle rates. Given that small firms are more likely to suffer from financial constraints, this suggests that capital rationing cannot explain the high hurdle rates. Another finding is that estimated coefficients for variables that proxy for the value of option to wait are robust with respect to whether the single-factor CAPM or the three-factor model is used. The three factor model has slightly higher explanatory power than CAPM (0.49 vs. 0.45). Finally, we find that in our models the intercept coefficient is not statistically different from zero. This suggests that we are not missing any systematic adjustments managers may be making to hurdle rates, such as using a higher hurdle rate to account for possible optimism in the cash flow projections.

It is possible that the results in columns 1 and 2 may be driven by the non-linear specification and also by simultaneously solving for the implied equity premium. To see whether or not this is the case, in Columns 3 and 4, we repeat the two exercises by including the three components of WACC in linear regression models for the single and three-factor models without simultaneously inferring the equity premium. The results displayed are remarkably similar to those in columns 1 and 2 in terms of magnitudes, statistical significance, and explanatory power. The similarity of the options related coefficients across the four columns indicate that the results are robust not just with respect to the non-linear and linear specifications, but also with respect to CAPM vs. the three-factor model. Taken together, this suggests that the variables we use to proxy for the option value to wait are orthogonal to the cost of capital component of hurdle rates.

This observation is confirmed by Panels B and C of Figure 4 which break up the two components of hurdle rates. As in Panel A, both Panels B and C have the 45 degree line superimposed. In Panel B the horizontal axis is the predicted WACC, while the hurdle rate minus the predicted hurdle premium (i.e., cost of capital plus the error term) is plotted on the vertical axis. The estimated slope coefficient is not statistically different from one (0.93, with standard errors of 0.30), and the intercept is not different from zero (0.011 with a *p*-value of 0.63). Panel C examines the hurdle rate premium by plotting the predicted hurdle premium (horizontal axis) against hurdle rate minus implied WACC using 3.8% as the equity premium (vertical axis). As in Panels A and B, the slope and intercept terms in Panel C are not different from one and zero, respectively.

In Table VI we pursue the relative importance of cost of capital and the option value to wait components of hurdle rates in explaining both the levels of and the cross-sectional variation in hurdle rates. In (5) and (6) we examine the cost of capital component using CAPM and the three-factor model, respectively. Judging by the R-squares of 0.11 and 0.17, we conclude that cost of capital is an important component. In fact using beta alone (Model 7) results in an R-square of only 0.03. The failure of (5) to satisfactorily explain hurdle rates can also be seen in Panel A of Figure 5: only one of the observations is below the 45 degree line. Apparently, this situation cannot be attributed to the inferred equity premium of 3.8% since using the historical risk premium of 6.6% (Panel B) does not produce a material improvement.[13]

Two additional comments are in order: One, the intercept estimates in (5) and (6) indicate that 6.3% to 7.7% of the average levels of hurdle rates cannot be explained by WACC. Two, while the cost of capital component belongs in the specification of hurdle rates, it is less important in explaining the variation in hurdle rates than the option to wait component. The linear model in (9) has an R-square of 0.37 suggesting that the premium component has approximately three times the explanatory power of the cost of capital component. However, in spite of this, based on the estimated intercept of 0.079, this component alone is not sufficient in

[13] Welch (2000) reports that academic financial economists forecast an arithmetic average equity premium over a 10-year horizon of 7%.

explaining the hurdle rates either. The implication that emerges from Table VI is that the specification of hurdle rates needs to include variables that capture both components. Combining the findings of Tables V and VI reveals that our non-linear models which simultaneously infer the equity premium (Models 1 and 2 of Table V) are superior to the two linear models that incorporate both components (Models 3 and 4 in Table V). Our models have the highest explanatory power (0.45 vs. 0.41 when CAPM is used and 0.49 vs. 0.48 when the three-factor model is used). At the same time, our two models have intercept estimates that are undistinguishable from zero. In sum, our models succeed in explaining both the average levels of hurdle rates and also the cross-sectional variation of hurdle rates.

VI. Conclusion

We examine the cross-sectional variation in hurdle rates that firms use in their capital budgeting decisions. We find that managers systematically add a hurdle premium to their CAPM based cost of capital. The size of this premium is substantial; it makes up about one half of the average hurdle rate used in practice. Following McDonald and Siegel (1986) we argue that the option to defer investments can explain the hurdle premium. This option to wait is most valuable to firms with growth opportunities facing organizational capital constraints that limit the rate of growth.

We develop a model of hurdle rates where the CAPM beta enters nonlinearly through the weighted average cost of capital (WACC) and variables that proxy for the option to wait that enter linearly. The coefficient estimates corresponding to the variables that proxy for the value of the option to wait for better future investment opportunities have the right signs and are statistically significant. We find that firms with higher hurdle rates keep higher cash balances, which is consistent with maintaining financial flexibility to undertake future valuable projects when they materialize. Such firms tend to be growth firms loading negatively on the Fama and French (1993) HML factor, which is also consistent with our hypothesis that the option to wait is more valuable to growth firms.

The model explains the level of hurdle rates and 45% of is cross-sectional variation across firms. The implied equity premium of 3.8% that we infer from the model is identical to the average equity premium that Graham and Harvey (2005) report in their survey of CFOs. The specification of our model is robust to whether we use CAPM or the Fama-French three-factor model. Since small firms are more likely to suffer from capital rationing, the insignificant factor loading for the Fama and French (1993) SMB factor suggests that the high hurdle rates are not driven by capital market constraints. Furthermore, the zero intercept of the model suggests that managers do not use higher hurdle rates to compensate for optimistic cash flow projections.

While we find both the cost of capital and the hurdle premium components to be important, cost of capital can only explain 10% of the variation in hurdle rates across firms, whereas proxies for the option to wait explain 35%. Further, the variation of the hurdle premium across firms is three times the variation in cost of capital.

Our analysis reconciles two seemingly contradictory findings in the literature. Since the hurdle premium (the difference between the hurdle rate used by a firm and its CAPM based cost of capital) varies substantially more than the cost of capital across firms, it masks the relation between the hurdle rate and the CAPM beta. This may explain why Poterba and Summers (1995) do not find CAPM betas to be significant in explaining hurdle rates. We also find that the CAPM based cost of capital is an important determinant of the hurdle rate that a firm uses. This is consistent with Graham and Harvey (2001) who report that most managers use the CAPM.

We hope that our findings – that the hurdle premium is about the same as the cost of capital and varies much more across firms – will stimulate further research that will help understand how firms arrive at what hurdle premium to use.

Appendix

Converting Levered/Unlevered Cost of Equity Hurdle Rates into WACC Equivalents

In 13.7% of the cases where survey participants indicate that they use either levered or unlevered cost of equity as their hurdle rate, we transform these cost of equity figures to their weighted average cost of capital (WACC) equivalents. If they indicate that the hurdle rate represents their cost of levered equity, we use this rate as the cost of equity and average it with their after-tax cost of debt and market value weights to compute their WACC. If they indicate that the hurdle rate represents their cost of unlevered equity, we check if these firms have any debt. Obviously, for the four firms that do not have any debt, unlevered cost of equity and WACC are identical. For firms with debt in their balance sheets, we lever up the reported cost of unlevered equity to obtain their cost of levered equity, and then compute WACC.

To compute WACC we use Compustat data to infer the market value-based weights for cost of debt and cost of equity. To compute the weight of debt, we divide total debt (Compustat items DLTT + DLC) by total debt plus market value of common and book value of preferred equity (CSHO × PRCC_F + PSTK). For the weight of equity we use (1 – weight of debt).

The mean life of a typical project for firms in our survey sample is 6.8 years. For this reason, we use the 10-year Treasury bond rate, which was 4.3% at the time of our survey, as a proxy for the risk-free rate.[14] For the before-tax cost of debt we use the survey participants' answers to our question regarding what the interest rate on their senior debt is.[15] The survey

[14] This choice seems to be justified for other reasons as well: In their survey of 27 highly regarded corporations, Bruner, Eades, Harris, and Higgins (1998) find that more than 70% use a 10-year or longer-term Treasury rate as the proxy for the risk-free rate. They report that only 4% of the firms in their survey use the 90-day T-bill rate.

[15] We do not know whether their answers refer to the coupon rate or the yield to maturity of their senior bonds. Thus, for firms that have not issued debt recently, it is possible that their answers do not reflect the marginal cost of debt if they report coupon rates. However, given the secular decline of interest rates

provides data on the before-tax cost of debt for 88 firms. Using Compustat data, we check whether firms that left the interest rate question blank had any debt. Out of the 39 non-responding firms we can match Compustat data for 28, and 16 of these firms turn out to have no debt. The remaining 12 firms with debt left the interest rate question blank.[16] For these firms we use their Altman's Z-score and the default spreads at the time of the survey to assign interest rates. If a firm's Z-score is greater than 3, a score that indicates a very low probability of default (8 firms), we assign the 10-year Treasury bond rate in effect at the time of the survey plus 1 percent (5.3%). For the two firms with Z-scores of less than 1.81 (financially unhealthy firms), we assign the 10-year Treasury rate plus 4 percent (8.3%). Firms that have Z-scores in the interval between 1.81 and 3 (2 firms) are assigned a before-tax cost of debt of 6.3. Finally, for firms that report a rate below the 10-year Treasury rate (4.3% at the time of the survey) we add a spread of 0.5% to the Treasury rate. Therefore, all our WACC calculations assume cost of debt of at least 4.8%.

We calculate a firm's tax rate by dividing total income taxes (Compustat item TXT) by income before taxes PI). When item TXT or PI is negative (tax credits and negative profits, respectively), we set the tax rate to zero. Additionally, we cap the tax rate at 34 percent.[17]

---

that started in the late 1990s and continued during the early 2000s, this should work against finding a hurdle rate premium.

[16] Out of these 12 firms, 2 have less than 1% debt (as a fraction of market value of assets) and another 6 less than 5%.

[17] The tax rate we obtain in this manner reflects a firm's average and not marginal tax rate. However, we were unable to obtain a sufficient number of observations on marginal tax rates.

References

Antill, Nick, and Robert Arnott, 2004, Creating value in the oil industry, *Journal of Applied Corporate Finance* 16(1), 18-31.

Bierman, Harold J., 1993, Capital budgeting in 1992: A survey, *Financial Management* 22(3), 24.

Brigham, Eugene F., 1975, Hurdle rates for screening capital expenditure proposals, *Financial Management* 4, 17-26.

Bruner, Robert F., Eades, Kenneth M., Harris, Robert S., and Robert C. Higgins, 1998, Best practices in estimating the cost of capital: survey and synthesis, *Financial Practice and Education* 8, 13-28.

Cleary, Sean, 1999, The relationship between firm investment and financial status, *Journal of Finance* 54(2), 673-692.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33 (1), 3-56.

Fremgen, J. M., 1973, Capital budgeting practices: A survey, *Management Accounting* 54(11), 19-25.

Gillham, Bill, 2000, Developing a questionnaire, Continuum, New York, NY.

Gitman, Lawrence J., and John R. Forrester Jr., 1977, A survey of capital budgeting techniques used by major U.S. firms, *Financial Management* 6, 66-71.

Gitman, Lawrence J., and Vincent A. Mercurio, 1982, Cost of capital techniques used by major U.S. firms: Survey and analysis of Fortune's 1000, *Financial Management* 11, 21-29.

Gitman, Lawrence J., and Pieter A. Vandenberg, 2000, Cost of capital techniques used by major U.S. firms: 1997 vs. 1980, *Financial Practice and Education* 10(2), 53-68.

Goyal, Amit K., and Pedro Santa-Clara, 2003, Idiosyncratic risk matters! *Journal of Finance* 58(3), 975-1007.

Graham, John R., and Campbell R. Harvey, 2001, The theory and practice of corporate finance: evidence from the field, *Journal of Financial Economics* 60, 187-243.

Graham, John R., and Campbell R. Harvey, 2005, The long-run equity risk premium, *Finance Research Letters* 2, 185-194.

Ibbotson, 2004. Yearbook, Ibbotson Associates, Chicago, IL.

Istvan, Donald F., 1961, The economic evaluation of capital expenditures, *Journal of Business* 36, 45-51.

Jagannathan, Ravi, and Iwan Meier (2002). Do we need CAPM for capital budgeting? *Financial Management* 31(4), 55-77.

Kim, Suk H., Trevor Crick, and Edward J. Farragher, Foreign capital budgeting practices used by the U.S. and Non-U.S. multinational companies, *The Engineering Economist* 25(3), 207-215.

Klammer, Thomas P., 1972, Empirical evidence of the adoption of sophisticated capital budgeting techniques, *Journal of Business* 45(3), 387-397.

Klammer, Thomas P., and Michael C. Walker, 1984, The continuing increase in the use of sophisticated capital budgeting techniques, *California Management Review* 27(1), 137-148.

Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.

Mao, James C.T., 1970, Survey of capital budgeting: Theory and practice, *Journal of Finance* 25(2), 349-360.

McDonald, Robert, 1998. Real options and rules of thumb in capital budgeting. In: Brennan, M.J., Trigeorgis, L. (Eds.), Innovation, Infrastructure, and Strategic Options. Oxford University Press, London.

McDonald, Robert, and Daniel Siegel, 1986, The value of waiting to invest, *Quarterly Journal of Economics* 101(4), 707-728.

Moore, James S., and Alan K. Reichert, 1983, An analysis of the financial management techniques currently employed by large US corporations, *Journal of Business, Finance, and Accounting* 10, 623-645.

Morgan, David L., 1988. Focus groups as qualitative research, Sage Publications, Newbury Park, CA.

Nohel, Tom, and Vefa Tarhan, 1998, Share repurchases and firm performance: new evidence on the agency costs of free cash flow, *Journal of Financial Economics* 49(2), 187-222.

Opler, Tim, Pinkowitz, Lee, Stulz, René, and Rohan Williamson, 1999, The determinants and implications of corporate cash holdings, *Journal of Financial Economics* 52, 3-46.

Petry, Glenn H., 1975, Effective use of capital budgeting tools, *Business Horizons* (October), 57-65.

Petty, William J., Scott, David F., and Monroe M. Bird, 1975, The capital expenditure decision-making process of large corporations, *The Engineering Economist* 20(3), 159-172.

Petty, William J., and Oswald D. Bowlin, The financial manager and quantitative decision models, *Financial Management* 5(4), 32-41é

Pinkowitz, Lee, René Stulz, and Rohan Williamson, 2006, Does the contribution of corporate cash holdings and dividends to firm value depend on governance? A cross-country analysis, *Journal of Finance* 61(6), 2725-2751.

Poterba, James M., and Lawrence H. Summers., 1995, A CEO survey of U.S. companies' time horizon and hurdle rates, *Sloan Management Review* (Fall), 43-53.

Schall, Lawrence D., Sundem,Gary L., and William R. Geijsbeek Jr., 1978, Survey and analysis of capital budgeting methods, *Journal of Finance* 38(1), 281-287.

Sharpe, William F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Simutin, Mikhail, 2010, Excess cash and stock returns, *Financial Management* 39(3), 1197-1222.

Singer, Eleanor, and Stanley Presser, 1989, Survey research methods: A reader, University of Chicago Press, Chicago, IL.

Smith, Clifford, David Ikenberry, Arun Nayar, Jon Anda, and Henry McVey, 2005, Morgan Stanley roundtable on capital structure and payout policy, *Journal of Applied Corporate Finance* 17(1), 36-53.

Trahan, Emery A., and Lawrence J. Gitman, 1995, Bridging the theory-practice gap in corporate finance: A survey of chief financial officers, *Quarterly Review of Economics and Finance* 35(1), 73-87.

Welch, Ivo, 2000, Views of financial economists on the equity premium and on professional controversies, *Journal of Business* 73 (4), 501-537.

Womack, Kent L., and Ying Zhang, 2005, Core Finance Course Trends in the Top MBA Programs in 2005, working paper, Dartmouth College.

Table I: Comparison of survey samples.

Panel A shows the industry breakdown using 2-digit SIC codes. "-" indicates that these sectors were excluded from the survey/sample or not listed as a category in the questionnaire. Panel B shows summary statistics on hurdle rate and the percentage of survey respondents that use CAPM and WACC.

Panel A

Industry	SIC Code	Poterba and Summers (1995)	Bruner, Eades, Harris, and Higgins (1998)	Graham and Harvey (2001)	Compustat (2003)	Jagannathan, Meier, and Tarhan (2010)
Agriculture, forestry, fishing	01 - 09	0.0	3.7	-	0.6	0.0
Mining, construction	10 - 17	4.4	0.0	4.0	10.5	10.7
Manufacturing	20 - 39	60.6	77.8	51.3[a]	64.5	66.0
Transportation, communication	40 - 49	12.5	11.1	18.2[b]	10.1 [c]	10.7 [c]
Wholesale and retail trade	50 - 59	6.9	3.7	11.1	13.7	11.6
Finance, insurance, and real estate	60 - 67	6.9	-	15.4	-	-
Services	70 - 89	5.6	3.7	-	0.6[d]	1.0[d]
Total obs.		228[e]	27[f]	392[g]	5,108	127

Panel B

Hurdle Rate						
Mean		12.2% (real) [h] =17.8% nom				14.8% (nominal)
						1
Median		10.0%[i]				5.0%
Standard deviation		~5.6%[k]				5.0%
Use CAPM			81%[l]	74%		
Use WACC			85%[m]			71.8%

Notes to Table I:

[a] Combines the survey categories “manufacturing” and “high-tech”; excludes “energy” (see footnote c below) which is reported in the survey category “transportation/energy”.

[b] Including “energy”; SIC codes 46, 49 (5540 and 5541).

[c] Excluding radio/TV and utilities providing gas, electricity, and water supply (SIC codes 4830, 4941).

[d] Only SIC code 70 (hotels, other lodging places).

[e] 160 respondents identified their firms. The questionnaire was sent to each CEO in the 1990 Fortune 1,000 list.

[f] Firms that were selected by their peers for best financial management practices according to Business International Corporation (1992), “Creating World-Class Financial Management: Strategies of 50 Leading Companies,” Research Report 1-110, New York, NY, 7-8. From the 50 companies, 18 with headquarters outside the US were excluded, 5 declined to participate.

[g] Questionnaires were sent by mail to each CFO in the 1998 Fortune 500 list and faxed out to 4,400 Financial Executives International (FEI) member firms. The raw data and a detailed description of the dataset are available on Campbell R. Harvey’s website.

[h] 66.2% of the respondents report nominal rates and the authors convert these to real rates using a long-term expected inflation rate of 5%.

[i] Page 46: 1/3 of all firms use <10% and the most common rate, used by 1/5 of the firms, is 10%.

[k] This is an approximation based on the midpoints of the categories and the frequencies shown in Figure 2 (page 46).

[l] An additional 4% use sometimes WACC, only 4% answered no (2 firms did not answer this question). 89% use some form of cost of capital as their discount rate (an additional 7% sometimes).

[m] An additional 4% use a modified version of CAPM.

Table II: Firm characteristics of surveyed firms.

The mean and median firm characteristics are tabulated for the 93 responding firms for which we can match with Compustat data in 2003 and for the 3,832 non-responding firms in Compustat. We exclude utilities, radio/TV broadcasting, cable, and other pay TV services (4840-4999), finance and insurance companies (SIC codes 6000-6499, 6700-6799), and health/education/social services, and museums (7200+). Book value of assets is Compustat item AT. Market value of assets is defined as book value of liabilities (LT) plus market value of assets, which is the sum of preferred stock (PSTK) and market value of common equity (PRCC_F × CSHO). Current ratio is current assets divided by current liabilities (ACT / LCT), total debt is the sum of debt in current liabilities plus long-term debt (DLC + DLTT), and return on book equity is the ratio between net income and book equity (NI / CEQ). For the characteristics that are expressed as fractions of book assets, we trim the top and bottom 0.5% of all Compustat firms, and then report the characteristics for responding survey firms and non-responding Compustat firms. The last two columns show the *p*-values for the difference in mean *t*-test and Fishers's exact test for differences in medians under the null hypothesis of zero mean and median, respectively.

	Survey N = 93		Computstat N = 3,832		Difference tests p-values	
	Mean	Median	Mean	Median	Mean	Median
Book assets	4,293	524	2,556	158	0.23	0.00
Market assets	8,821	680	4,168	279	0.02	0.00
Sales	4,142	373	2,392	144	0.15	0.01
Market/book assets	2.37	1.69	3.09	1.60	0.21	0.46
Cash/book assets	0.15	0.07	0.20	0.11	0.05	0.09
Sales/book assets	0.66	0.47	0.70	0.51	0.27	0.75
Current ratio	2.53	1.80	2.87	1.97	0.27	0.25
Total debt/book assets	0.29	0.24	0.29	0.19	0.99	0.40
Capital expenditures/book assets	0.06	0.04	0.04	0.03	0.00	0.09
Operating income/book assets	0.05	0.07	0.02	0.05	0.03	0.01
Return on book equity	-0.01	0.05	0.07	0.07	0.38	0.40

Table III: Hurdle rates and financial characteristics.

The table shows coefficients and standard errors (in brackets below) for bivariate regressions. The dependent variable in all regressions is self-reported hurdle rate. All explanatory variables are defined as in Figure 2 above, with the exception of the dividend payout ratio that is expressed in % to make the coefficients comparable to Poterba and Sommers (1995). *** indicates significance at the 1% level, ** at the 5% level, and * at the 10% level. Standard errors are below in brackets.

Characteristics	Poterba and Summers (1995)		Jagannathan, Meier, and Tarhan (2010)	
	All firms	Manufacturing	All firms	Manufacturing
P/E ratio	– 0.008	– 0.018	– 0.019	-0.018
	(0.031)	(0.036)	(0.020)	(0.029)
Dividend payout ratio (in %)	– 0.002	– 0.002	– 0.026	– 0.027
	(0.002)	(0.002)	(0.020)	(0.026)
Current ratio	1.889***	1.891***	0.791***	0.776***
	(0.633)	(0.746)	(0.196)	(0.240)
% change in EPS (past 10 years)	0.062	0.072	-0.035	-0.221
	(0.051)	(0.055)	(0.219)	(0.321)
Total equity return (past 10 years)	– 0.052	– 0.039	– 0.168	– 0.663
	(0.052)	(0.057)	(0.433)	(0.745)
Historical beta	– 0.102	– 0.067	0.754	2.048
	(1.411)	(2.038)	(0.834)	(1.375)
Fundamental beta			1.950	3.127*
			(1.249)	(1.884)
Equity market-to-book	– 0.187	– 0.287	0.127	0.046
	(0.170)	(0.307)	(0.168)	(0.217)
Tobin's q ratio	– 0.043	– 0.336	1.229	1.384
	(0.622)	(0.777)	(0.715)	(1.054)
Stock turnover rate	0.003	0.008	0.459	1.005*
	(0.007)	(0.008)	(0.381)	(0.560)

Table IV: Statistics on hurdle rates and industry affiliation.

Panel A shows summary statistics of self-reported hurdle rates for three samples (in percent). The hurdle rates represent the nominal rate that the company has used for a typical project during the previous two years. In the column "WACC equivalent sample" we drop firms do not use WACC or cost of levered/unlevered equity (category "other"). We convert self-reported hurdle rates that represent the cost of levered or unlevered equity are to their weighted average cost of capital (WACC) equivalents. This conversion procedure is explained in Section III.C. For two out of the 17 firms that use either cost of equity or unlevered cost of equity we cannot match the debt-equity ratio from Compustat to calculate the WACC equivalent. Therefore, we report the 101 WACC equivalent hurdle rates. The last column shows the sample statistics for WACC equivalent hurdle rates for which we can match beta from Barra and information from CRSP/Compustat. Panel B tabulates the fractions of firms in each industry.

Panel A

Hurdle rate	All respondents	Category "other"	WACC equivalent sample	Sample matched with Barra and CRSP/Compustat
Mean	14.8	17.6	14.1	14.5
Median	15.0	15.0	14.0	14.9
Minimum	5.0	9.0	5.0	7.0
Maximum	40.0	40.0	30.0	30.0
Std. dev.	5.3	6.4	4.9	4.3
25th percentile	12.0	12.0	10.8	12.0
75th percentile	16.0	22.5	15.0	16.0
Skewness	1.4	0.7	1.7	1.0
Kurtosis	6.7	2.2	9.6	4.6
N	119	18	101	73

Panel B

Industry	All respondents	Category "other"	WACC equivalent sample	Sample matched with Barra and CRSP/Compustat
Mining, construction	10.7	28.6	8.3	8.1
Manufacturing	66.0	50.0	67.9	66.2
Transport, communication	10.7	14.3	10.7	12.2
Wholesale and retail trade	11.6	0.0	11.9	12.2
Services	1.0	7.1	1.2	1.3

Panel C

Industry	N	Mean	Median	Standard deviation	Min	Max
Mining, construction	6	13.1	12.5	3.8	9.0	20.0
Manufacturing	48	15.2	15.0	4.3	7.0	30.0
Transport, communication	9	12.4	12.0	2.2	9.0	15.0
Wholesale and retail trade	9	14.2	15.0	2.2	8.5	16.0
Services	1	14.0	14.0	-	14.0	14.0

Table V: Model to explain hurdle rates.

The dependent variable in all models is hurdle rate (WACC equivalent). The values for the equity premium and SMB and HML show implied premia from the model estimation. Beta is the fundamental Barra beta. Debt-to-assets is total debt (Compustat items DLC + DLTT) divided by market value of assets, which is book value of total liabilities and preferred stock plus shares of common stock outstanding times price (LT + PSTK + PRCC_F × CSHO). Cash/assets is CHE to market value of assets, industry return is the average monthly return of the firms in the same 2-digit SIC industry over the past 5 years, and the industry R-square is the average R-square from the index model of firms in the same 2-digit SIC industry (using 5 years of monthly returns and the S&P 500 as the index). Sales growth/employee ([($SALE_t$ – $SALE_{t-1}$)/SALE $_{t-1}$]/EMP is a categorical variable where firms are assigned to 1 if the value is lower than mean – 2 standard deviations across all firms; the next category is from mean – 2 std. dev. to mean – 1.5 std. dev., for which we assign 2, etc. For values larger than mean + 2 std.dev. we assign 10. Financial health (Altman's Z-score) is a categorical variable which is 1 if z-score < 1.81 (financially unhealthy), 2 if z-score ≥ 1.81 and < 3 (neutral), and 3 if ≥ 3 (financially very healthy firms).

	Nonlinear model		Linear model	
	(1) CAPM	(2) Fama-French 3-factor model	(3) WACC components	(4) WACC components and Fama-French factors
Intercept	0.014	0.034	0.062*	0.074*
	(0.022)	(0.023)	(0.038)	(0.038)
Equity premium	0.038***	0.028**		
	(0.011)	(0.012)		
SMB		0.004		0.010
		(0.008)		(0.006)
HML		-0.012***		-0.012***
		(0.005)		(0.004)
Beta			0.027**	0.020*
			(0.010)	(0.010)
Debt-to-assets			0.008	0.015
			(0.025)	(0.023)
Cost of debt			-0.172	-0.056
			(0.360)	(0.349)
Cash/assets	0.119***	0.098***	0.126***	0.096***
	(0.037)	(0.037)	(0.039)	(0.038)
Industry return	0.054	0.052	0.071	0.055
	(0.042)	(0.041)	(0.046)	(0.044)
Industry R-square	-0.374***	-0.398***	-0.361***	-0.406***
	(0.097)	(0.095)	(0.105)	(0.100)
Sales growth/employee	0.008***	0.007***	0.008***	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)
Financial health	0.007	0.010**	0.009	0.012*
	(0.005)	(0.005)	(0.006)	(0.006)
Adjusted R^2	0.452	0.494	0.410	0.482

Table VI: Separating WACC and the explanatory variables for hurdle premium.

The dependent variable is hurdle rate (WACC equivalent). Variable definitions are the same as in Table V.

	Nonlinear model		Linear model		
	(5) CAPM	(6) Fama-French 3-factor model	(7) Only beta	(8) WACC components	(9) Model without WACC
Intercept	0.063***	0.077***	0.124***	0.158***	0.079***
	(0.011)	(0.038)	(0.012	(0.024)	(0.024)
Equity premium	0.047***	0.031**			
	(0.011)	(0.014)			
SMB		0.009			
		(0.008)			
HML		-0.014**			
		(0.006)			
Beta			0.020*	0.026**	
			(0.011)	(0.011)	
Debt-to-assets				-0.035	
				(0.025)	
Cost of debt				-0.052	
				(0.410)	
Cash/assets					0.158***
					(0.038)
Industry return					0.055
					(0.045)
Industry R-square					-0.284***
					(0.100)
Sales growth/employee					0.008***
					(0.003)
Financial health					0.006
					(0.005)
Adjusted R^2	0.114	0.165	0.030	0.094	0.371

Figure 1: Adoption of DCF methods, WACC, CAPM, and company-wide hurdle rates over time.

The surveys on capital budgeting practices of U.S. firms are listed in chronological order below the horizontal time axis. The scatter plot summarizes their findings regarding the percentage of firms that: (i) Use discounted cash flow (DCF) methods, including net present value (NPV), adjusted present value (APV), internal rate of return (IRR), and the profitability index (PI); (ii) Use the weighted average cost of capital (WACC) to discount cash flows; (iii) Employ the Capital Asset Pricing Model (CAPM) to compute cost of equity; and (iv) Use a company-wide hurdle rate.

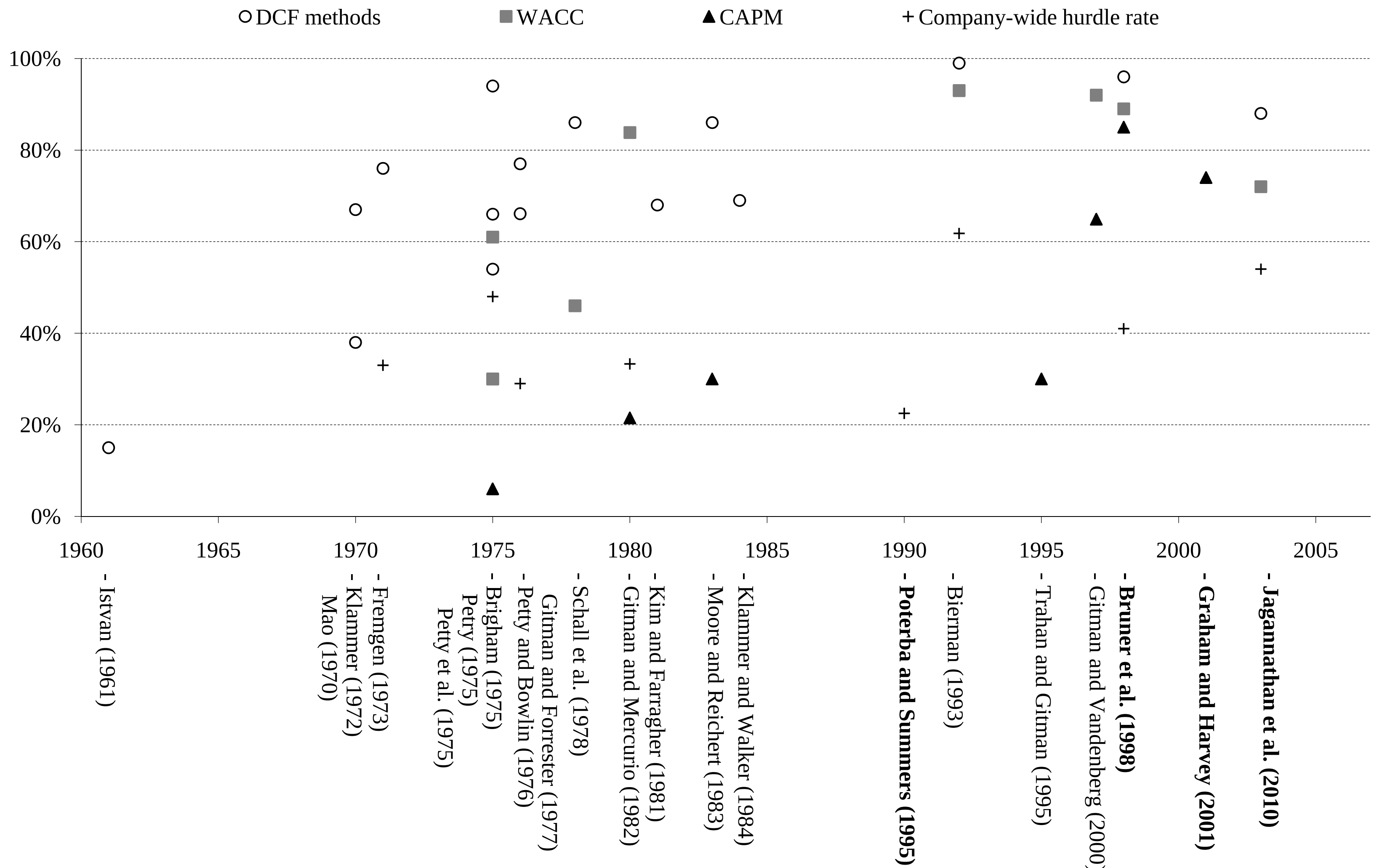
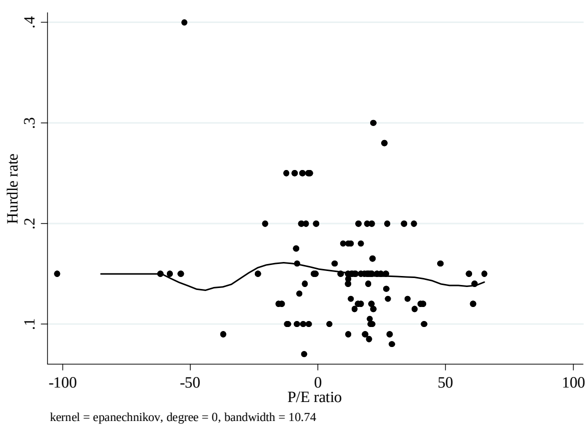


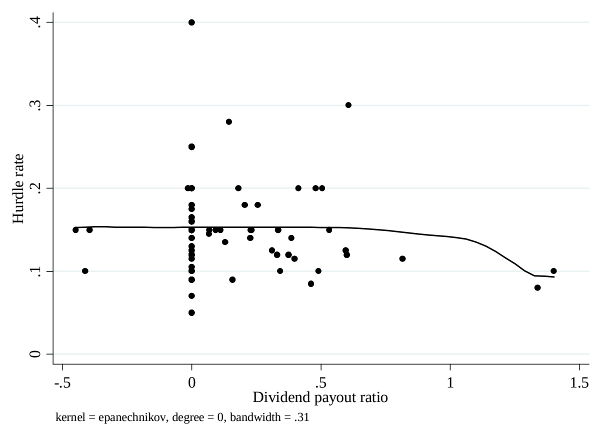
Figure 2: Hurdle rates and firm characteristics.

Kernel-weighted local polynomial regressions of hurdle rate on various firm characteristics. For the local mean smoothing we apply the Epanechnikov kernel function with a rule-of-thumb bandwidth estimator (the bandwidth is shown below the graphs). The characteristics are P/E ratio (Compustat items PRCC_F/EPSPX), dividend ratio (DVC/IBAD), current ratio (ACT/LCT), percentage change in earnings per share ([$EPSPX_t$ – $EPSPX_{t-10}$] / $EPSPX_{t-10}$), total past equity return over 10 years ([$PRCC\_F_t/CUMADJ_t$) – $PRCC\_F_{t-10}/CUMADJ_{t-10}$] / [$PRCC\_F_{t-10}/CUMADJ_{t-10}$]), historical beta (regressing five years of monthly total stock returns on stock market returns), fundamental beta from Barra, market-to-book equity ratio ([CSHO × PRCC_F] / CEQ), Tobin's q ([AT + CSHO × PRCC_F – CEQ – TXDB] / [0.9 × AT + 0.1 × MKVAL]), and stock turnover rate (SHSTRD/CSHOQ). The footnote below indicates outliers that have been removed from the graphs.

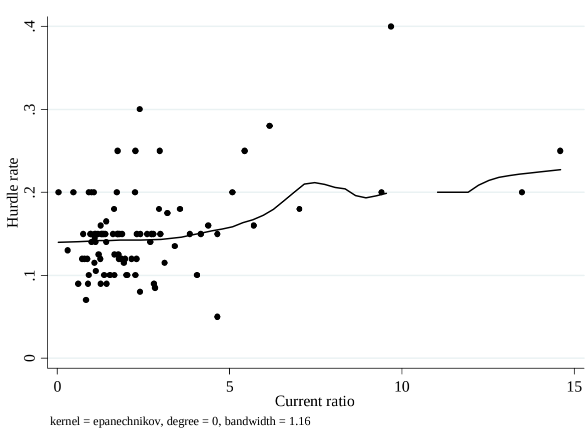
A: Price-earnings ratio



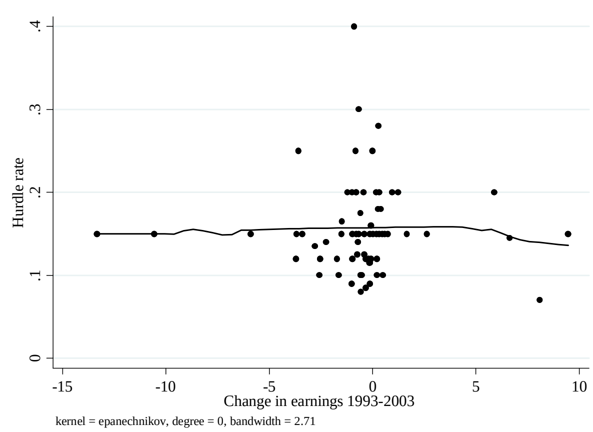
B: Dividend payout ratio



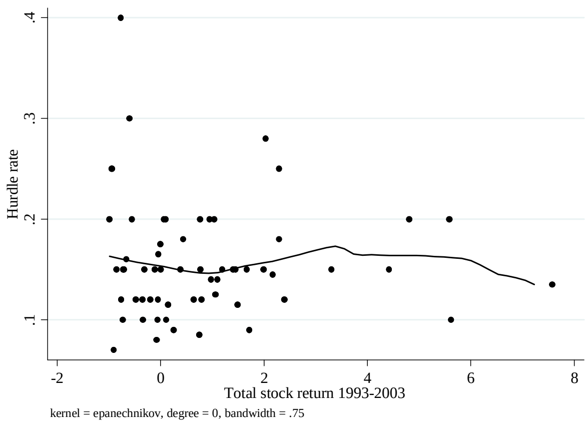
C: Current ratio



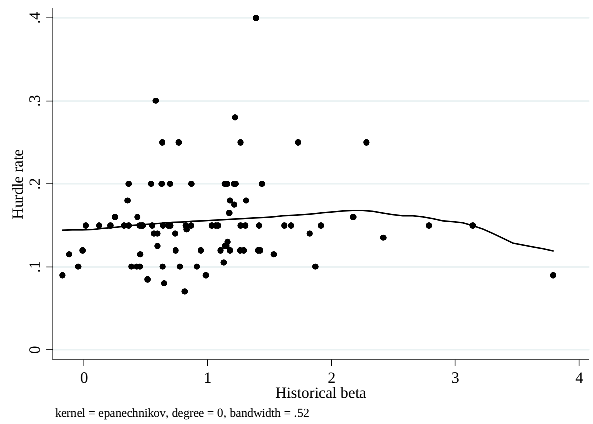
D: Percentage change in EPS



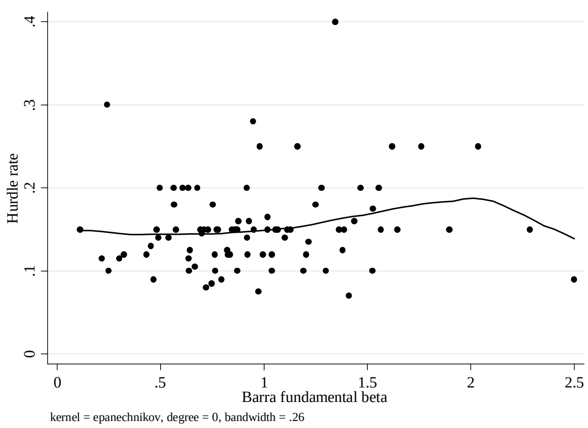
E: Total equity return



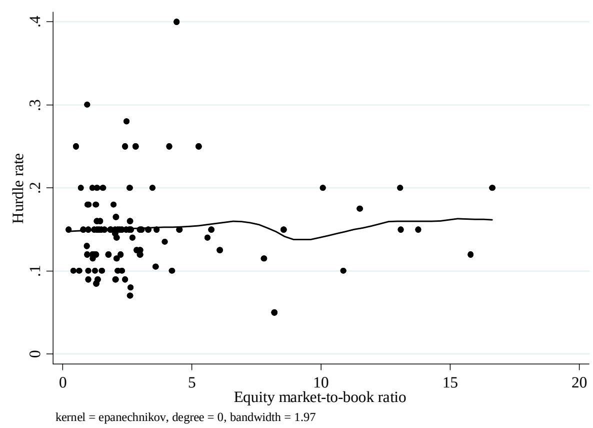
F: Historical beta



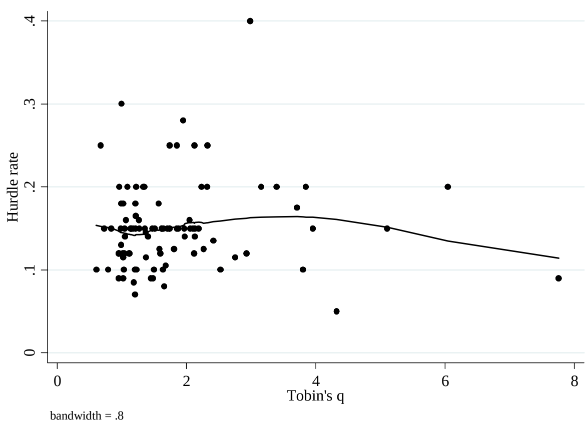
G: Fundamental beta



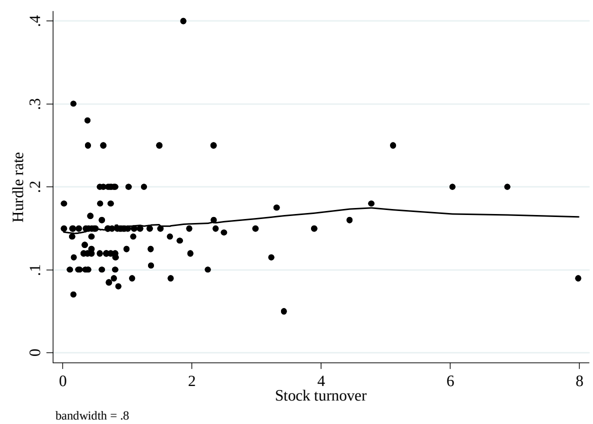
H: Market-to-book ratio



I: Tobin’s q



J: Stock turnover rate



Notes to Figure II:

The following observations in each Panel have been excluded for the local polynomial fitting and are not shown in the graph:

A: (hurdle rate 0.12, P/E ratio 467) and (0.14, 479). Additionally, the observation with hurdle rate = 0.40 shown in the graph is excluded when fitting the curve.

B: (hurdle rate 0.15, dividend payout ratio -2.8).

C: (hurdle rate 0.20, current ratio 25.2). Additionally, the observation (0.40, 9.7) is shown in the graph but excluded when fitting the curve.

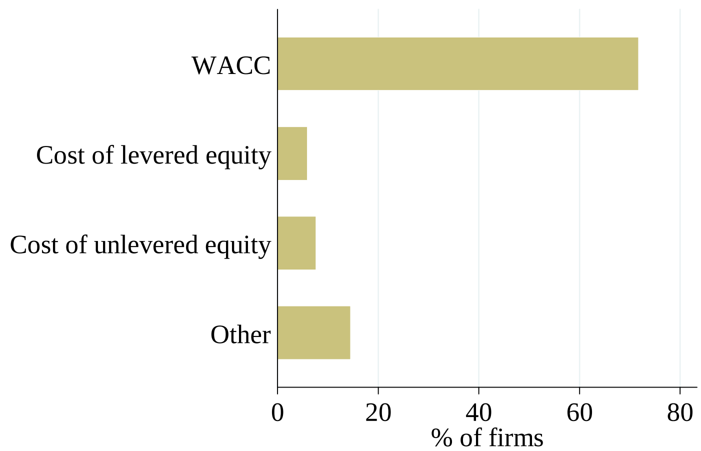
E: (hurdle rate 0.15 and total equity return 11.7) and (0.09, 82.5).

H: Negative ratios and ratios larger than 20: (WACC equivalent hurdle rate 0.20 and equity market-to-book ratio 25.6) and (0.14, -14.6).

Figure 3: What self-reported hurdle rate represents.

A total of 117 firms responded to the question what the firm's hurdle rate represents (Panel A). The eleven firms that explicitly indicate that they add a premium to the weighted average cost of capital (WACC) to assess their hurdle rate are included in the category WACC. Panel B shows summary statistics of self-reported hurdle rates. The hurdle rates represent the nominal rate that the company has used for a typical project during the previous two years. Self-reported hurdle rates that represent the cost of levered or unlevered equity are converted to their weighted average cost of capital (WACC) equivalents (see Appendix A for details) and firms in the "other" category are dropped from the sample. We report the hurdle rates for the remaining 101 firms.

Panel A



Panel B

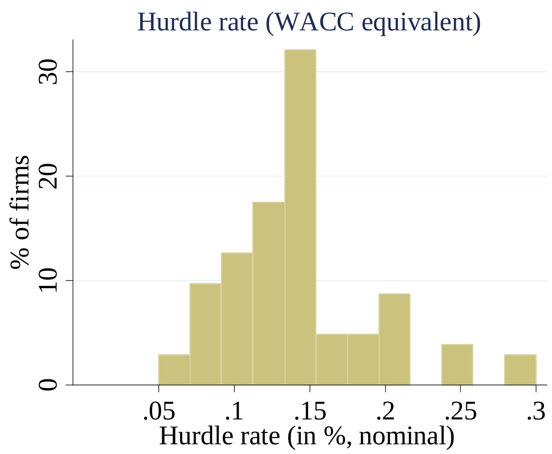
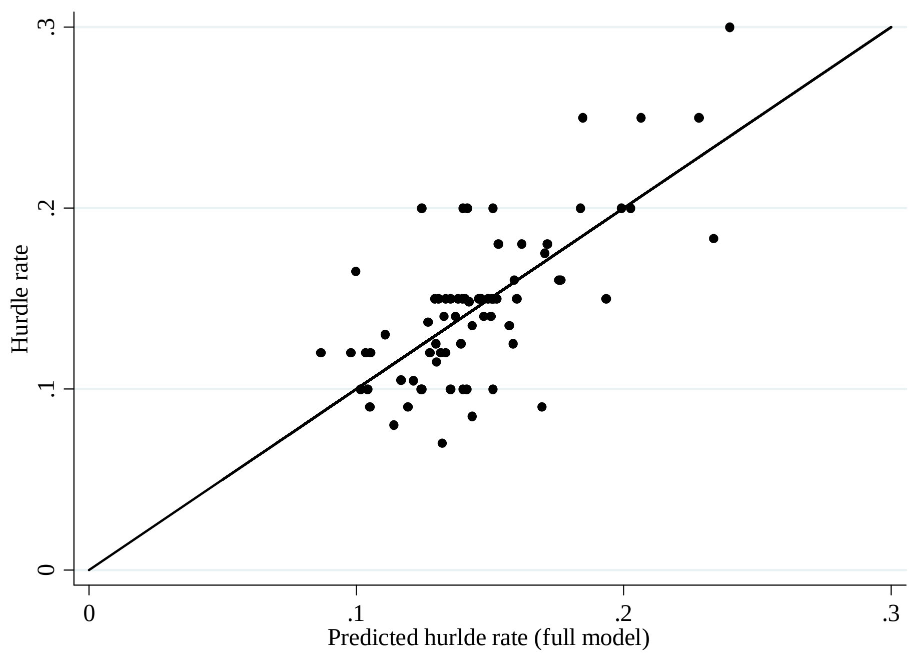


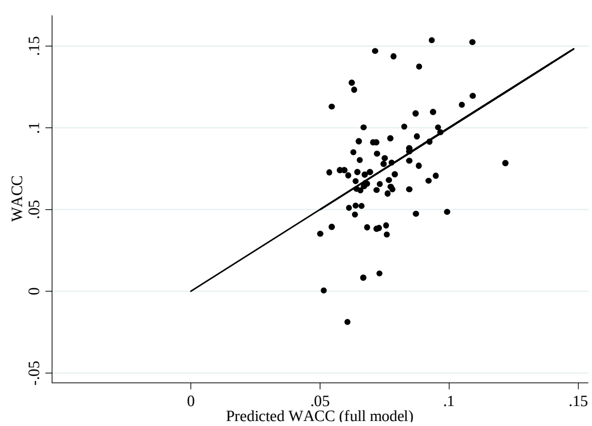
Figure 4: Comparison of the predictions of the full model with self-reported hurdle rates.

Panel A compares predicted hurdle rate from the full model on the horizontal axis with self-reported hurdle rates shown on the vertical axis. Panels B and C decompose the predicted values in two components: Predicted WACC against the WACC = hurdle rate – predicted premium and predicted premium against premium = hurdle rate – computed WACC. The solid line in all three panels is the 45-degree line.

Panel A



Panel B



Panel C

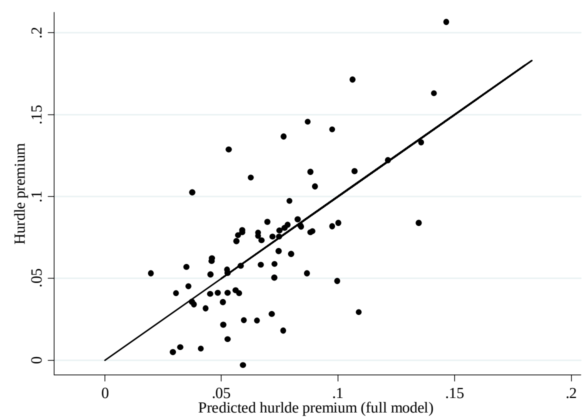
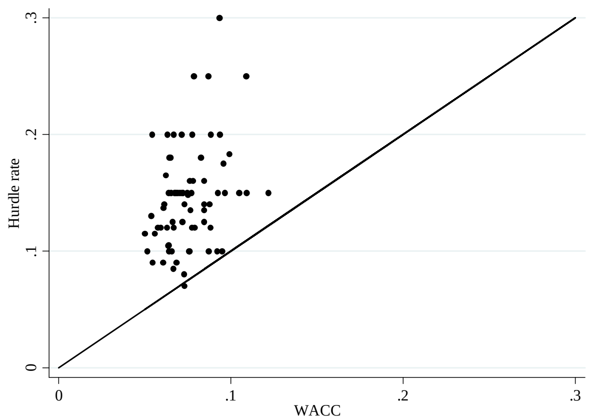


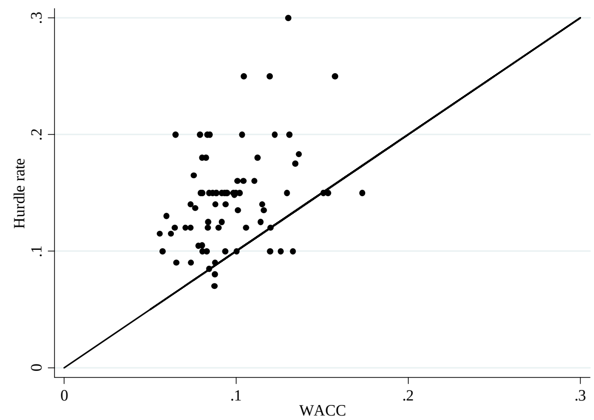
Figure 5: Relationship between hurdle rates and WACC.

The two scatter plots show predicted hurdle rates when using WACC plus a constant (Model 1 in Table VI). Panel A uses the implied equity premium of 3.8% and Panel B assumes an equity premium of 6.6% based on a historical average from Ibboston (2004).

Panel A



Panel B



May 2008

The Most Important Number in Finance

The Quest for the Market Risk Premium

Marc Zenner
marc.p.zenner@jpmorgan.com
(212) 834-4330

Scott Hill
scott.d.hill@jpmorgan.com
(415) 315-8842

John Clark
john.hs.clark@jpmorgan.com
(212) 834-2156

Nishant Mago
nishant.x.mago@jpmorgan.com
(212) 834-2172

We would like to thank Andy Chi, Amra Coralic, and Jessica Vega for their invaluable contributions to the analytics and construction of this report. We would further like to thank Ben Berinstein, Tomer Berkovitz, Jennifer Conrad, John Graham, Thomas Hagerstrom, Chris Harvey, Sean Nossel, Bob Principe, James Rothschild, Mark Shifke, Mark Solomons, Jerry Topitzer, Chris Ventresca, Stephen Wolf, and Janet Wiener for their helpful comments and suggestions. Special thanks are due to Colleen Galle for her support and guidance in writing and publishing this report.

1. The most important number in finance

You will not find it in section C of *The Wall Street Journal*. *CNBC* will not mention it in its morning market recap. *The Economist* will not provide it in its back pages with other financial data. Yet it is one of the most critical metrics in finance, a figure implicit in the evaluation of financing and investment opportunities: the market risk premium. What is it? How and where should it be used? What is the right number to use? Does it change over time?

In this report, we (1) estimate a current range of risk premiums; (2) explain how the risk premium has increased since the beginning of the subprime crisis; (3) discuss how, thanks to Federal Reserve intervention, a higher risk premium does not necessarily lead to a higher cost of capital; and (4) debate how possible divergence between equity and credit markets since last summer may affect strategic and financial decision-making. In addition, we review some common methods used to estimate the market risk premium.

What is the market risk premium?

The *market risk premium* (MRP) reflects the incremental premium required by investors, relative to a risk-free asset like U.S. Treasury bonds, to invest in a globally diversified market portfolio. Below is a simple and generally accepted equation:

Expected return on the market portfolio = Risk-free rate of return + market risk premium

Should the market risk premium be higher for some assets and lower for others? Most likely yes, but how should the adjustment be made? The Capital Asset Pricing Model (CAPM) proposes one such adjustment. CAPM states that the expected return on an asset is the risk-free rate plus an MRP that is adjusted, through beta, to reflect the market risk of the asset:

*Expected return on an asset = Risk-free rate of return + beta * market risk premium*

The beta is a calibration factor that is higher (lower) than one if the asset has a systematic, or non-diversifiable, risk that is higher (lower) than the market's risk. In the CAPM framework, the MRP should apply to all assets, including bonds, real estate, art, etc. In practice, however, the risk premium is mostly used to estimate the expected return on equity (also referred to as the cost of equity). Bond markets rely on their own risk premium concept, the credit spread, which is the difference between the yield on a bond and the maturity-matched Treasury rate.

From a macroeconomic perspective, the MRP reflects the broader outlook on the whole economy. Factors influencing investors' views on market risk include outlooks for economic growth, consumer demand, inflation, interest rates, and geopolitical risks. As such, the MRP is a single metric that reflects these inputs in the expected returns of various asset classes.

Why is the market risk premium so important?

While many finance professionals and executives actively manage their debt and debate the incremental basis points their firm may have to pay on new bonds, they do not tend to focus much on the cost of equity. Is it that debt financing is so much more prevalent than equity financing? Not really. Even with a tax system that favors debt financing, equity financing constitutes over 80% of the total market capitalization for a typical non-financial S&P 500 firm today.

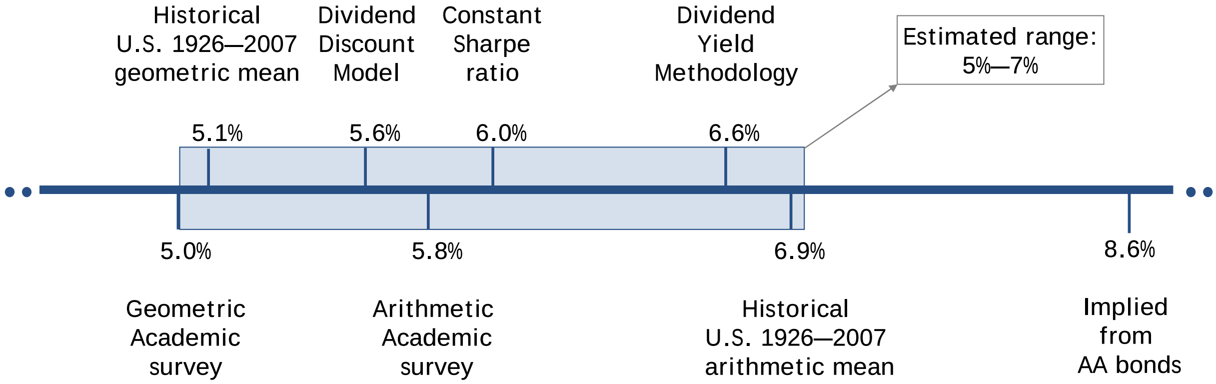
Why then is there less focus on the cost of equity? Maybe because most firms manage debt actively and equity only passively; or because an economic cost of equity of 12% does not translate into an actual cash outlay of 12%; or perhaps because there is no consensus on how to estimate the market risk premium.

Practical Application: Understanding and quantifying the MRP is critical to the value-creation process. With most of their capitalization in the form of equity, decision-makers require an estimate of the MRP to determine their cost of capital, identify projects that create shareholder value, decide how much to pay for acquisition targets, evaluate their capital structure, and compare the costs of various sources of financing. Not adjusting the cost of equity to new market realities may lead firms to (1) over or under-invest or (2) forgo capital-structure opportunities.

What is the market risk premium today?

No single method to estimate the MRP is used universally. Our review of various methods (detailed in Section 2) suggests that they each have strengths and weaknesses. They also generate a wide range of results as summarized in the figure below. We therefore recommend thinking about the MRP in terms of a range rather than a unique number. Based on our results, the MRP probably falls within a range of 5% - 7% today.

Figure 1: Summary of risk premium estimates



Source: JPMorgan

If I do not use CAPM, should I still focus on the market risk premium?

Most practitioners use CAPM as their method of choice to estimate the cost of capital. Interestingly, while academics often emphasize the limitations of CAPM, they still tend to focus on it when teaching about the cost of capital. Two of the risk premium estimation methods we used rely on CAPM (the Dividend Yield and the bond-based methods). The Dividend Discount and Sharpe ratio methods, as well as the historical analysis, do not rely on CAPM. Practitioners who do not use CAPM can still use the risk premium range we suggest by using the low (high) end of the range for projects they perceive to be at the low (high) end of their risk spectrum.

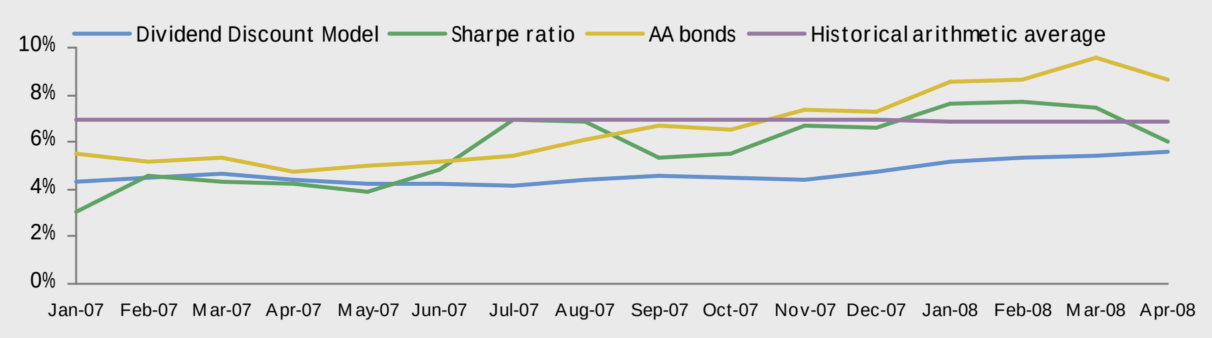
My firm is global, so should I focus on a risk premium based on U.S. data?

The results we present are based on U.S. market data. Can you use these risk premium estimates for investments in other countries? We believe that the U.S.-based MRP is a reasonable estimate for developed economies for a couple of reasons. First, an unconstrained investor would not freely invest in a market in which he/she would earn a lower risk-adjusted return. Hence risk premiums should gravitate to each other across open developed markets, and the U.S.-based risk premium should serve as a good estimate for this. The situation may be different in emerging markets, however, where non-market risks may exist (e.g., political risk) or where investor segmentation and constraints limit the free flow of capital into and out of the country. Second, the U.S. market has some data advantages, namely very broad markets with long data histories. Many other markets tend to be over-weighted in some sectors (e.g., banking, shipping, energy, telecommunications) or have data series that have been interrupted by political events in the 20th century.

Has the risk premium changed since last summer?

Are we in a new risk premium environment? The figure below shows that the answer depends on the methodology. The historical method, as expected, suggests no change in the risk premium. On the other hand, methods that rely on current market information (which we discuss in detail later) signal that the risk premium has increased since the credit crisis began last summer, but that it has declined from its peak in February/March.

Figure 2: Comparing risk premium estimates since last summer



Source: JPMorgan, SBBI Market Report-Morningstar, Bloomberg

Should executives change their hurdle rates for capital allocation?

Boards of Directors and senior executives implicitly use the MRP when determining hurdle rates for new projects and acquisitions. There is a preference for hurdle rates that do not change often, possibly because stable hurdle rates facilitate communication with regional and divisional management. In some cases, however, it is critical to understand whether changing market conditions affect how the market prices risk. Financial decision-makers examine day-to-day data when they look at debt financing, so why not also for equity, often the biggest component of the capital structure? We believe that today's environment warrants re-estimating the cost of capital using new market information, in particular when considering large capital projects or acquisitions.

Practical Application: The cost of capital for many S&P 500 firms has not increased since last summer. Why? While risk premiums increased in both credit and equity markets, the Fed's policy of lowering interest rates has succeeded in offsetting this increase for the largest firms in the economy. It is worth noting that, even in today's environment, many firms tend to use a hurdle rate that is a few percentage points higher than their true cost of capital, which may lead them to forgo valuable investment opportunities.

Which is right—equity or credit markets?

Many market observers have focused on how the equity and credit markets have behaved differently since last summer. While credit markets lost significant liquidity and experienced dramatic pricing changes, the non-financial component of equity markets remained relatively unaffected until the beginning of this year. Have credit markets overreacted, and should they revert to more normalized pricing? Have the equity markets failed to completely absorb the effects of the financial crisis, and should we expect a further decline in equity values, along with an increase in the MRP? Or do credit markets reflect a higher overall premium combining both a heightened risk premium and an increased liquidity premium? In many segments of the credit markets, liquidity diminished significantly over the last few months, but not so in the equity markets. We believe that both effects have taken place; i.e., the equity risk premium has increased, but the credit markets have been affected even more because they are also pricing in an additional premium for liquidity.

Practical Application: Executives should consider this debt vs. equity market premium dynamic when making funding decisions. For example, the after-tax cost of hybrids should be compared to an updated after-tax cost of equity. Furthermore, as discussed above, given that low Treasury rates have offset rising risk premia for the largest firms, executives should consider locking in a low long-term cost of capital, especially if they have near-term refinancing, capital or liquidity needs, or if they expect rates to increase because of inflationary pressures.

2. Different methods to estimate the MRP

A. Historical average realized returns

A common way to estimate the MRP has been to compare realized annual equity returns to average returns of U.S. Treasury bonds over some historical time period.

MRP = average annual equity index return - average return on Treasury bonds

This method is widely used in practice but has a few weaknesses which diminish its usefulness.

Choice of averaging method: The choice of arithmetic vs. geometric averaging methods can lead to significant differences in MRP estimates. For example, if $100 grows to $110 in one year and then drops back to $100 the next, the arithmetic average annual return is [+10.0% - 9.1%]/2, or 0.5%. The arithmetic average represents the best estimate of annual expected return. The geometric mean, however, will be 0%, which is the compounded annual return the investor actually earned. Many academics prefer the arithmetic average because it represents an investor's expected return at any given point in time. But the geometric mean better reflects asset returns investors should expect over long horizons.

Time horizon: As evidenced in Figure 3 below, different time horizons also yield different MRP estimates. For example, an observer examining the U.S. data since 1978 using the geometric mean would determine that the MRP is 4.9%, whereas an observer viewing the data since 1946 would instead conclude it is 5.7%.

Figure 3: Historical risk premium estimates across various time periods

Large company stocks - Intermediate T bonds	Arithmetic	Geometric
1926-2007	6.9%	5.1%
1946-2007	6.8%	5.7%
1978-2007	5.7%	4.9%

Source: Morningstar, JPMorgan

Reaction to changing risk premium: In a changing risk-premium environment, this method can produce counterintuitive results. For example, if the risk premium increases and cash-flow projections remain unchanged, equity prices will drop. This drop in equity prices reflects investors' demand for higher *future* expected returns in the riskier environment. But the drop would cause lower realized returns, which in turn would lower the average *historical* returns, thereby suggesting a lower instead of higher risk premium. Though this backward-looking method may not capture the direction of the change in risk premium well, it may still be a viable long-term estimate of the risk premium investors expect to earn by investing in equity.

Figure 4: Pros and cons of using the historical method

Pros	Cons
Easy to compute	Estimate depends on historical window
Has been a standard in business schools	Estimate depends on averaging method
Does not change often and rapidly	Does not change often and rapidly; i.e., does not incorporate new market realities
Can be sourced by a third-party provider such as Ibbotson Associates	Responds in a counterintuitive way to changes in actual risk premium

Source: JPMorgan

B. Dividend Discount Model

Another means of estimating the MRP is through the Dividend Discount Model (DDM), which can be used to calculate the current market cost of equity. The model solves for an internal rate of return (cost of equity) based on the price level and expected dividend stream of an index (often the S&P 500 as a proxy for the broad market). Dividends are projected by applying an expected payout ratio to forecasted earnings. Earnings are forecasted, in turn, by combining near-term (i.e., 5 years) market estimates with a perpetuity growth rate equivalent to long-term nominal GDP growth. The dividend payout ratio is initially assumed to be the average of recent historical payout ratios, but increases over the long-term towards 80% in the terminal period as reinvestment opportunities are assumed to subside. Simplistically, the formula for the market cost of equity is:

$$Price_0 = \sum_{t=1}^{\infty} \frac{Dividend_t}{(1 + Cost\ of\ Equity)^t}$$

where t is time from now to infinity. Subtracting the 10-year government bond yield from the market cost of equity then provides the market risk premium. Thus, the MRP formula is as follows:

MRP = Cost of equity implied by DDM – 10-year government bond yield

Figure 5: Pros and cons of risk premium implied from Dividend Discount Model

Pros	Cons
Implied from equity market values	Price variable changes daily
Changes and responds to current market environment	Highly dependent on future dividend/cash flow estimates
Forward-looking; not heavily reliant on historical data	Dividend forecasts not updated frequently; may not take market cycles into account

Source: JPMorgan

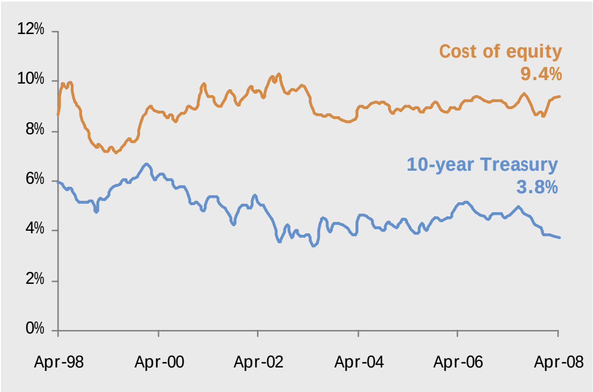
Changes over time: The market cost of equity varies primarily with movements in the level of the index, but also with changes in expectations for future dividends. The chart below shows the market cost of equity based on the S&P 500, as well as the 10-year Treasury yield, over the last 10 years. The resulting MRP, shown to the right, varies from a low of 1.3% at the peak of the market to a high of over 6% in the post-9/11 era. After 2003, the MRP stabilized in the 4% range until the recent credit crisis, which has led to a re-pricing of risk and a higher MRP.

Figure 6: Dividend Discount Model implied risk premium over time

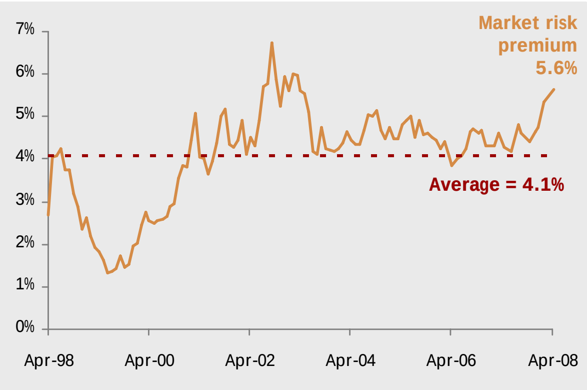
Yearly arithmetic average market risk premium since 1998

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Average MRP	3.4%	1.8%	2.8%	4.4%	5.2%	5.0%	4.6%	4.7%	4.3%	4.4%	5.2%

U.S.—S&P 500 expected return %



U.S.—S&P 500 market risk premium %



Source: JPMorgan, Bloomberg

C. Constant Sharpe ratio method

Another useful metric to estimate the risk premium, the Sharpe ratio, has been inherited from portfolio management theory. The Sharpe ratio measures a portfolio's excess return per unit of risk and can be used to estimate the MRP:

$$\text{Market Sharpe ratio} = S_M = \frac{\text{Portfolio MRP}}{\text{Volatility of MRP}}$$

We estimate that, over the last 50 years, the Sharpe ratio for the broad market (using the S&P 500 index as a proxy) has been about 0.3, which is consistent with academic research. Assuming that this ratio is constant going forward, we can then solve for the forward-looking MRP by multiplying the S&P 500 Sharpe ratio by a measure of future market volatility. We estimate future market volatility via the VIX index, which measures the volatility implied from options on the S&P 500 index. Thus, the Sharpe ratio-implied MRP is:

*MRP = Market (S&P 500) Sharpe ratio * Market (S&P 500) implied volatility*

Figure 7: Pros and cons of the Sharpe ratio method

Pros	Cons
■ Estimate of Sharpe ratio based on more than 50 years of historical data; consistent with academic research ■ VIX component is forward-looking; captures shifts in investor sentiment very quickly	■ Some evidence that Sharpe ratio may change over time instead of remaining constant ■ VIX measures short-term volatility (<1 year), whereas risk premium is generally viewed long-term (10+ years)

Source: JPMorgan

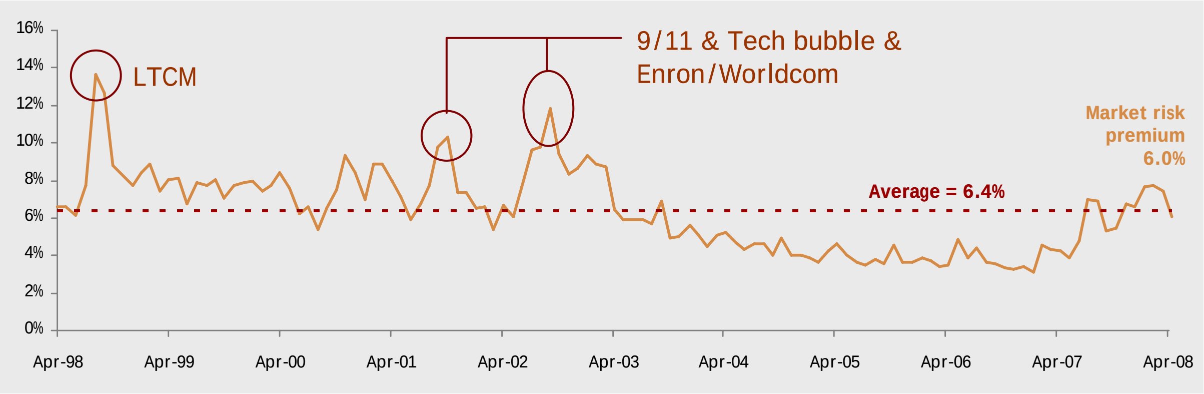
Changes over time: Figure 8 displays the Sharpe ratio-implied MRP over the last 10 years. By definition, the Sharpe ratio-implied MRP moves proportionally with the VIX volatility index. At times of greater uncertainty and market panic, including the Long-Term Capital Management fallout in 1998, the 2000-2002 recession/tech bubble burst, and the current credit crisis, investors have fled to safer securities and demanded a greater MRP to keep their investments in riskier assets. Such shifts in risk preferences have been accompanied by spikes in volatility.

Figure 8: Historical risk premiums computed from the Sharpe ratio method

Yearly arithmetic average market risk premium since 1998

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Average MRP	8.2%	7.8%	7.4%	7.9%	8.0%	6.6%	4.6%	3.9%	3.7%	5.2%	7.2%

U.S.—S&P 500 market risk premium %



Source: Bloomberg, Federal Reserve Data

D. Bond-market implied risk premium

Most of us think of the MRP in the context of cost of equity. Risk premiums do, however, also exist for corporate bonds. The expected return of a bond can therefore be expressed using the Capital Asset Pricing Model, as:

*AA yield ≈ AA expected return = risk-free rate + beta * market risk premium*

Therefore, if we know the expected return on the bond and its beta, we can estimate the implied MRP. For high-yield bonds, we know the yield, but the expected return is likely to be significantly lower than the promised yield. For AA rated corporate bonds, on the other hand, the default probabilities are very low and we can use the yield as a proxy for expected returns. Hence, we use the price series of AA corporate bonds to estimate the MRP. The beta of AA bonds is between 0.15 and 0.20, depending on the estimation period. Using a beta of 0.15, we estimate that the bond-implied MRP was below 4% in 1998 and 2004-2005 but recently rose to about 8.6%.

Figure 9: Pros and cons of the bond-market data methodology

Pros	Cons
■ Based on daily market feedback regarding risk premium on bonds	■ Possibility that expected default rates change at the same rating
■ Assumes no capital-structure arbitrage; i.e., when bonds demand a higher risk premium, other assets such as equity also demand a higher return	■ Depends on CAPM and an assumption about bond betas
	■ Implied risk premium captures both a liquidity and risk premium

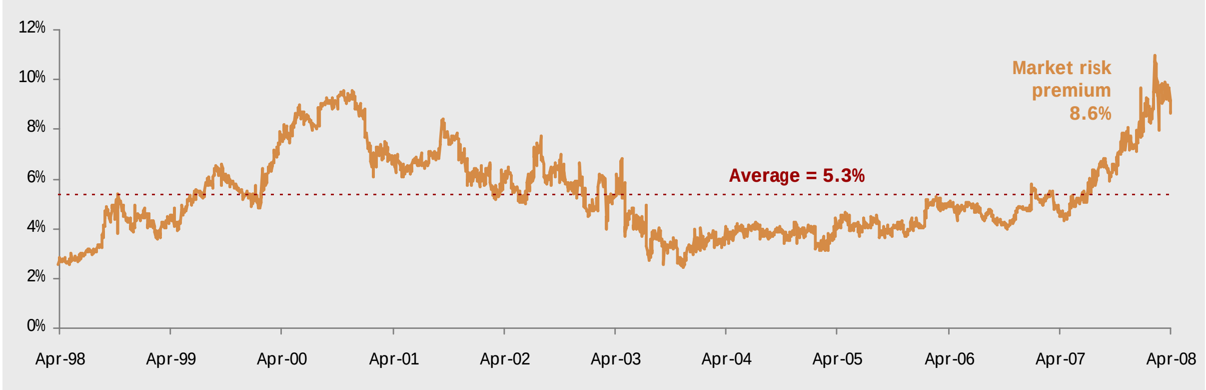
Source: JPMorgan

Figure 10: Recent changes in the bond-market implied risk premium

Yearly arithmetic average market risk premium since 1998

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Average MRP	3.5%	5.1%	8.0%	7.0%	6.0%	4.2%	3.8%	4.0%	4.7%	5.7%	8.9%

AA industrial index spread implied market risk premium %



Source: JPMorgan, Bloomberg

E. Dividend Yield Method

A methodology that is closely related to the Dividend Discount Model method uses the dividend yield as a starting point. The price of a dividend-paying stock can be estimated using the constant-growth valuation model. This model assumes that the dividend will grow at a constant rate forever. We rewrite this model as a function of the cost of equity, stating that the cost of equity is the dividend yield plus the long-term growth rates. The formulas are:

$$Price = \frac{Div_1}{Cost\ of\ Equity - GrowthRate}\text{, and therefore } Cost\ of\ Equity = \frac{Div_1}{Price} + GrowthRate$$

This approach works well in sectors with large and steadily growing dividends. We applied the methodology to three industries known for their focus on dividend yields: Real Estate Investment Trusts (REITs), Master Limited Partnerships (MLPs), and regulated utilities. In the regulated utilities industry, regulators accept this method as a way to estimate the cost of equity. Another useful feature of the model is its closeness to the cash cost of the equity. In fact, some practitioners look at the dividend yield only and ignore the growth component of the equation.

MRP = (Cost of equity implied by Dividend Yield Method - 10-year government bond yield) / beta

Figure 11: MRP implied by dividend yields in dividend-heavy sectors

	Dividend yield	IBES 5-yr EPS growth	Overall growth[1]	Cost of equity	Equity beta	Implied MRP
Regulated utilities	4.1%	6.1%	4.5%	8.6%	0.78	6.6%
MLPs	6.4%	6.5%	5.1%	11.4%	0.61	13.0%
REITs	5.4%	6.5%	4.7%	10.5%	1.13	6.1%
Mean	**5.3%**	**6.4%**	**4.8%**	**10.2%**	**0.84**	**8.6%**
Median	**5.4%**	**6.5%**	**4.7%**	**10.5%**	**0.78**	**6.6%**

Source: JPMorgan, FactSet

[1] Overall growth is weighted combination of 5-yr EPS growth and 4% perpetuity growth assumptions

We use EPS estimates and an assumption of constant payout ratios to forecast the dividend growth over the next five years, and an assumption that dividends will grow at 4% thereafter (long-term real growth plus inflation). Our results suggest that the cost of equity for these sectors is in the 9% to 12% range. The figures also display two clear weaknesses: (1) the need for assumptions to estimate overall or long-term growth, estimated in this case as a weighted-average of the 5-year EPS growth projection followed thereafter by a 4% perpetuity growth rate; and (2) the need to rely on CAPM and a beta estimate to extract the MRP implied by our cost-of-equity estimates. Today, this approach yields an MRP in the 6% range for REITs and utilities, and a higher number for MLPs.

Figure 12: Pros and cons of MRP implied from Dividend Yield Method

Pros	Cons
Intuitive: cost of equity equals dividend yield plus a growth rate	Only applicable in a few dividend-heavy sectors
Widely accepted in dividend-heavy sectors	Capital structures of these sectors may not represent those of the market at large
Close to the actual cash cost on equity	Relies on perpetuity growth rate assumption
Dividend yield changes daily	Depends on CAPM and assumption about industry or firm beta

Source: JPMorgan

F. Survey evidence

One relatively basic method for determining the MRP is to survey market participants for their views on required returns. Such surveys have polled academics, investors, and other corporate-finance practitioners such as CFOs.

An academic survey by Ivo Welch from Brown University provides useful insights on MRP estimates.[1] The typical finance professor responding to Welch's survey estimates that the long-term market risk premium is 5% on a geometric basis and 5.8% on an arithmetic basis. Interestingly, these numbers are very close to the MRP estimates of the historical realized returns methodology, suggesting that finance professors still primarily rely on that approach.

A similar survey conducted quarterly from 2000 to 2007 by John Graham and Campbell Harvey of Duke University compiled the views of U.S. CFOs regarding the current risk premium.[1] Their average risk premium in 1Q07 was 3.2%, and the range from 2000 to 2007 was 2.4% to 4.7%.

Relying on these survey results has some advantages. First and foremost, in the case of finance professors, participants may be biased in their preferred methodology, but they are typically unbiased in their MRP estimates—that is, they do not have any specific incentive to make low or high estimates. Secondly, academics tend to spend a lot of time on the subject and have significant influence on how regulators, practitioners, and even investors look at the MRP.

On the other hand, survey respondents can provide wide differences of opinion and express views that may be extreme (such as a negative MRP). Surveys can also reflect the collective views of the constituent base. As an example, academics' reliance on the historical-data approach suggests that their estimates will not change very often. This may be an advantage for executives looking for a MRP estimate that is robust through time, but it may not capture the realities of a new market environment (such as structural shifts, tax changes, etc.). Conversely, the CFO-based survey is different in that its results are quite volatile and might represent current market conditions and concerns.

Figure 13: Pros and cons of surveys

Pros	Cons
■ Significant time researching this topic	■ Wide differences in opinion
■ Academic estimate unbiased (no reasons for it being high or low)	■ Does not change often and rapidly; i.e., does not incorporate new market realities (e.g., tax rate changes)
■ Does not change often and rapidly	

Source: JPMorgan

As stated above, none of these six estimation methods are used universally. Taken together, however, they provide an understanding of the drivers of the market risk premium, and allow decision-makers to consider using a method that reflects today's volatile market environment.

[1] Ivo Welch, "The Consensus Estimate for the Equity Premium by Academic Financial Economists in December 2007", January 18, 2008.
[2] John Graham and Campbell Harvey, "The Equity Risk Premium in January 2007: Evidence from the Global CFO Outlook Survey", January 18, 2007.

The Effect of Issuing Biased Earnings Forecasts on Analysts' Access to Management and Survival

by

Bin Ke*
Smeal College of Business
Pennsylvania State University

and

Yong Yu
McCombs School of Business
University of Texas at Austin

Abstract

This study offers evidence on the earnings forecast bias analysts use to please firm management and the associated benefits they obtain from issuing such biased forecasts in the years prior to Regulation Fair Disclosure. Analysts who issue initial optimistic earnings forecasts followed by pessimistic earnings forecasts before the earnings announcement produce more accurate earnings forecasts and are less likely to be fired by their employers. The effect of such biased earnings forecasts on forecast accuracy and firing is stronger for analysts who follow firms with heavy insider selling and hard-to-predict earnings. The above results hold regardless of whether a brokerage firm has investment banking business or not. These results are consistent with the hypothesis that analysts use biased earnings forecasts to curry favor with firm management in order to obtain better access to management's private information.

First draft: January 28, 2004
Current draft: April 25, 2006

We thank an anonymous reviewer, Anil Arya, Larry Brown, Steve Huddart, Artur Hugon, Henry Y. Lo, Jim McKeown, Drew Newman, and workshop participants at the 2005 American Accounting Association FARS mid-year meeting and annual meeting, Chinese University of Hong Kong, the 12th Conference on the Theories and Practices of Securities and Financial Markets held at the National Sun Yat-sen University (Kaohsiung, Taiwan), University of Colorado (Boulder), Ohio State University and Pennsylvania State University for helpful comments. We thank IBES for providing the earnings forecast data. Bin Ke acknowledges the research support of the Smeal College of Business.

*Corresponding author
Bin Ke
Pennsylvania State University
354 Business Building
University Park, PA 16802
814-865-0572 (phone)
814-863-8393 (fax)
bxk127@psu.edu (email)

1. Introduction

Prior research finds that financial analysts often issue biased earnings forecasts to please firm management (see e.g., Richardson et al., 2004; Francis and Philbrick, 1993; Das et al., 1998; Lim, 2001; Matsumoto, 2002), but it is not well understood why analysts have incentives to do so. In addition, the form of the forecast bias analysts are assumed to use to please management varies across studies. Some studies assume managers prefer optimistic earnings forecasts (e.g., Francis and Philbrick, 1993) while others assume managers prefer pessimistic forecasts (e.g., Matsumoto, 2002). Recently Richardson et al. (2004) argue that managers prefer initial optimistic forecasts followed by pessimistic forecasts immediately before the earnings announcement.

The objective of this study is to identify the form of the earnings forecast bias analysts use to please management and the associated benefits analysts receive from such biased earnings forecasts. We consider both annual and quarterly earnings forecast biases because analysts often issue both forecasts. To our knowledge, we are the first study that simultaneously examines annual and quarterly earnings forecast biases at the individual analyst level. Because earnings forecast accuracy is important to analysts and their brokerage firms (Mikhail et al., 1999; Leone and Wu, 2002), we hypothesize that analysts issue biased earnings forecasts to curry favor with management so that they can obtain more private information from management to improve their earnings forecast accuracy relative to other analysts (H1). In addition, we hypothesize that analysts who issue biased earnings forecasts are less likely to be fired by their employers (H2).

In light of previous research's conflicting results on the form of the forecast bias analysts use to please management, we consider four possible forms of earnings forecast biases that capture the intertemporal pattern of each analyst's earnings forecasts (denoted OP, OO, PO, PP). For annual earnings forecasts, the four forecast biases are defined using each analyst's first and last one-year ahead annual earnings forecasts issued between two consecutive annual earnings announcement dates. OP denotes the analysts whose first one-year ahead annual earnings forecast issued after the prior fiscal year's earnings announcement is optimistic (i.e., forecast is greater than the realized earnings), but whose last one-year ahead annual earnings forecast issued before the current year's earnings announcement is pessimistic (i.e., forecast is less than or equal to the realized earnings); OO denotes the analysts whose first and last annual earnings forecasts are always optimistic; PP denotes the analysts whose first and last annual earnings forecasts are always pessimistic; finally, PO represents the analysts whose annual earnings forecasts switch from initial pessimism to later optimism. For quarterly earnings forecasts, the four forecast biases for each analyst are defined similarly except that the first earnings forecast for the current quarter is defined as the first two-quarters ahead earnings forecast issued after the announcement of the quarterly earnings two quarters prior and the last earnings forecast is defined as the last one-quarter ahead earnings forecast issued before the current quarter's earnings announcement. The difference in the definitions of the four forecast biases for annual and quarterly earnings forecasts reflects the reality that the majority of analysts issue at least two one-year-ahead annual earnings forecasts between two consecutive annual earnings announcements while only one one-quarter-ahead quarterly earnings forecast between two consecutive quarterly earnings announcements.[1]

We test our hypotheses over the period January 1, 1983-June 30, 2000.[2] For both annual and quarterly earnings forecasts, we find that OP analysts are associated with more accurate earnings

[1] For all the stocks with nonmissing data included in the IBES database over calendar years 1983-1999, 73% of the analyst firm years issue at least two one-year ahead annual earnings forecasts between two consecutive annual earnings announcement dates, while only 23% of the analyst firm year quarters issue at least two one-quarter ahead quarterly earnings forecasts between two consecutive quarterly earnings announcement dates.

[2] Our sample ends on June 30, 2000 because Regulation Fair Disclosure (FD) became effective on October 23, 2000, which prohibits firm management from disclosing material nonpublic information to select individuals, and our variable definitions are measured from July 1, year t to June 30, year t+1. We leave to future research to study the effect of the regulation on the private communication between firm management and analysts.

forecasts and a smaller likelihood of being fired by their employers than other analysts, suggesting that it is the OP bias that analysts use to obtain better access to firm management.

Richardson et al. (2004) find that the OP bias based on consensus earnings forecasts is more severe for firms whose managements wish to sell their personal equity holdings in the firm. Das et al. (1998) argue that access to management is more valuable to analysts when a firm's earnings are difficult to forecast. Therefore, we conjecture that analysts who cover firms with heavy insider trading or hard-to-forecast earnings benefit more from issuing OP earnings forecasts. Consistent with this conjecture, we find that the predicted effect of OP on forecast accuracy and firing is stronger for firms whose earnings are more difficult to forecast and whose managements engage in heavy insider selling. Overall, these results are consistent with the hypothesis that analysts use OP earnings forecasts to gain better access to managers' private information.

Further analyses indicate that the predicted effects of H1 and H2 exist for analysts employed by both investment banks and pure brokerage firms (i.e., those without investment banking businesses). Thus, our results cannot be solely driven by the alleged investment banking incentive. However, we cannot rule out the possibility that the predicted effects for H1 and H2 for the investment bank analysts are partially driven by the investment banking incentive.

Given the documented benefits from issuing OP earnings forecasts, why don't all analysts issue OP forecasts for all firms? We believe there are several reasons. First, as Hong and Kubik (2003) argue, some analysts may not be willing to issue biased forecasts given their good conscience and what they know. Second, firm managers do not have incentives to play the biased earnings forecast game. For example, as we have shown above, managers who do not plan to sell stocks in their own firms do not have as strong an incentive as managers who do to pressure analysts to issue biased forecasts. Furthermore, even if both analysts and managers have incentives to play the biased forecast game, it seems reasonable to assume that managers prefer to cooperate with analysts who have a significant influence on capital market investors (hereafter referred to as the “bang for the buck” hypothesis). We find empirical support for this hypothesis. Specifically, relative to other analysts, we find that OP analysts are more experienced, employed by larger brokerage firms, and more likely to be an All-Star as determined by the Institutional Investor magazine, all indicators of influential analysts.

The results from our study should be of interest to investors and securities regulators who wish to understand the causes of biased earnings forecasts. Our results are also relevant to future researchers who wish to investigate analysts’ forecasting behavior. It is common for researchers to require an analyst to be in the sample for several years. Since less biased analysts do not survive, analyses based on surviving analysts could be biased and should be interpreted with caution.

Our study is not the first to recognize the potential influence of firm management on analysts’ biased earnings forecasts. For example, Francis and Philbrick (1993) argue that analysts issue optimistic earnings forecasts in order to maintain good relations with management (see also Das et al., 1998; Lim, 2001).[3] However, those studies do not examine the benefits of biased forecasts to individual analysts (i.e., improved forecast accuracy and job security) nor simultaneously consider the various earnings forecast biases.

Chen and Matsumoto (2006) study how revisions in stock recommendations affect analysts’ access to management and forecast accuracy. They find that analysts who upgrade a stock experience a significant increase in forecast accuracy relative to analysts who downgrade a stock prior to the passage of regulation FD but not after. They do not study earnings forecast biases or analyst firing.

[3] A recent study by Eames and Glover (2003) raises questions on the robustness of Das et al. (1998).

Hong and Kubik (2003) and Leone and Wu (2002) examine the impact of forecast optimism on analyst turnover (including promotion) but do not consider the other bias measures. More importantly, neither study considers the fear of losing access to management as an explanation for analysts' biased earnings forecasts.

The rest of the paper is organized as follows. Section 2 develops the research hypotheses. Section 3 describes the sample selection and data. Section 4 discusses the variable definitions and presents the regression models. Section 5 reports the test results. We conclude in Section 6.

2. Hypothesis Development

2.1. Benefits from Issuing Biased Earnings Forecasts

Prior research on earnings forecast biases often focus on managerial incentives (see e.g., Matsumoto, 2002; Richardson et al., 2004). Those studies assume that analysts have incentives to issue biased earnings forecasts preferred by managers. We directly test this assumption by demonstrating the benefits individual analysts receive from issuing biased earnings forecasts. Specifically, we examine whether analysts who issue biased earnings forecasts gain better access to management's private information so that their earnings forecast accuracy and job security can be improved relative to other analysts. The possible forms of earnings forecast biases analysts could use to please managers are discussed in section 2.2.

It is well recognized that earnings forecast accuracy is an important determinant of an analyst's reputation, annual compensation, and career success. For example, Mikhail et al. (1999) and Hong et al. (2000) find that analysts whose earnings forecasts are more accurate relative to others are less likely to be fired. The reason forecast accuracy is important to analysts and their brokerage firms is that brokerage firms want analysts who are influential among the buy-side (especially institutional investors) and this influence is directly determined by an analyst's ability to make accurate earnings forecasts. Leone and Wu (2002) document that earnings forecast accuracy is an important determinant of the All-Star analyst ranking in the Institutional Investor magazine's annual survey of buy-side investors. Influential analysts can significantly raise the reputation and influence of their brokerage firms among Wall Street investors and corporate executives, which in turn can bring many tangible and intangible benefits, such as stimulating more trading by their firms' investing clients, helping their firms win more lucrative investment-banking businesses, etc.

Both anecdotal evidence and academic research also suggest that management is an important source of analysts' private information (see Schipper, 1991). One important form of private communication between management and analysts is closed conference calls (before Regulation FD took effect). Bowen et al. (2002) find that conference calls significantly increase analysts' earnings forecast accuracy. Furthermore, Solomon and Frank (2003) report that analysts who issue unfavorable earnings forecasts are often punished in subtle ways by firm management, especially before Regulation FD took effect. Therefore, we hypothesize that analysts have an incentive to use biased earnings forecasts to please management so that they can gain better access to management's private information to improve their earnings forecast accuracy relative to other analysts. This prediction is stated in the following hypothesis:

H1: analysts who issue biased earnings forecasts have more accurate earnings forecasts relative to other analysts.

In addition to suffering a decrease in current earnings forecast accuracy, analysts who do not issue biased earnings forecasts are likely to lose the privileged access to management and their *future* earnings forecast accuracy is expected to deteriorate as a result. Since analysts' forecast accuracy is critical to brokerage firms' reputation and influence, we expect brokerage firms to fire those analysts who do not issue biased earnings forecasts, even after controlling for those analysts' current earnings forecast accuracy. This discussion leads to the following hypothesis:

H2: analysts who issue biased earnings forecasts are less likely to be fired.[4]

Given the hypothesized benefits in H1 and H2, why would not all analysts issue biased earnings forecasts preferred by management? We believe there are several reasons. First, as Hong and Kubik (2003, p. 339) argue, some analysts may not, out of good conscience, be willing to play the biased earnings forecast game with management given what they know. However, good conscience is unobservable and thus cannot be directly tested. Second, even if analysts are interested in playing the biased earnings forecast game, some firm managers may lack incentives. For example, Richardson et al. (2004) find that managers' preference for biased earnings forecasts is stronger for firms whose managers wish to sell a portion of their personal equity holdings in the firm. Thus, if a manager does not plan to sell shares, he should have little incentives to play the biased forecast game, ceteris paribus. Section 5.2.3 reports evidence consistent with this argument. Third, even if both analysts and managers have incentives to play the biased forecast game, it is reasonable to assume that managers prefer to cooperate with analysts who can exert a significant influence on both other analysts and equity investors (referred to as the "bang for the buck" hypothesis).[5] Cooperation with obscure analysts will be less beneficial to managers because these analysts will be less effective in affecting stock investors' expectations. Furthermore, the strategy of giving all analysts who are willing to issue biased forecasts the same private information may not be optimal because it would make no single analyst better off relative to his peers and thus

---

[4] Although an analyst who issues biased forecasts may be able to move up to a more prestigious brokerage firm, we expect this move-up effect to be weaker than the firing effect in H2 because the analyst's current employer will try to offer monetary incentives to retain him. Empirically, we find only weak evidence that analysts who issue biased forecasts are more likely to move up to more prestigious brokerage firms.

[5] This hypothesis has support from both academic research (see e.g., Gintschel and Markov, 2004; Krigman et al., 2001) and anecdotal news reports (see e.g., Smith and Cauley, 1999; Levitt, 1998).

would reduce all analysts' incentives to play the biased forecast game. In section 5.2.6 we provide evidence on the characteristics of the analysts who issue biased forecasts that are consistent with the "bang for the buck" hypothesis.

2.2. Definitions of Earnings Forecast Biases

Although the idea that analysts use biased earnings forecasts to win favor from firm management has been advanced in many studies, the form of the earnings forecast bias analysts are assumed to use to please management varies across studies. Many studies assume that managers prefer optimistic earnings forecasts (see e.g., Francis and Philbrick, 1993) while others assume that managers prefer pessimistic forecasts (see e.g., Matsumoto, 2002). Richardson et al. (2004) reconcile the conflicting assumptions in prior research by analyzing the intertemporal patterns of consensus earnings forecasts. They show that managers prefer initial optimistic consensus earnings forecasts followed by pessimistic consensus earnings forecasts immediately before the earnings announcement.

Richardson et al. (2004) further show that one important reason that managers prefer initial optimism and later pessimism is their desire to sell a portion of their equity holdings in the firm at a higher price. To avoid the perception of illegal insider trading and investor litigations, corporate executives are usually allowed to sell their equity holdings only after the earnings announcement (see Bettis et al., 2000; Roulstone, 2003). In addition, Bartov et al. (2002) find that for firms with similar earnings forecast errors at the beginning of a quarter, firms that can meet or beat analysts' latest earnings forecasts before the earnings announcement enjoy a higher stock return than firms that cannot. Therefore, corporate executives prefer analysts to issue pessimistic earnings forecasts immediately before the earnings announcement and optimistic earnings forecasts immediately after the earnings announcement, both of which lead to higher stock prices.[6]

In addition to different assumptions on the form of the earnings forecast bias preferred by managers, prior research does not differentiate annual versus quarterly earnings forecast biases nor study how individual analysts, if issuing multiple earnings forecasts for the same fiscal period, adjust their forecast biases over the forecast horizon. Because a typical analyst issues both annual and quarterly earnings forecasts, it is important to understand whether analysts issue biased annual or biased quarterly earnings forecasts or both to win favor from management. In this study we consider both annual and quarterly earnings forecasts at the individual analyst level. To our knowledge, we are the first study that examines the intertemporal pattern of <u>individual</u> analysts' annual and quarterly forecast biases.

Although the evidence in Richardson et al. (2004) and our discussion above suggest that analysts should issue OP earnings forecasts to win favor from management, we also investigate the other three earnings forecast biases (i.e., PP, OO, and PO) as well because prior research has argued that managers prefer pure forecast optimism or pure forecast pessimism. By considering the four possible forecast biases simultaneously, we can determine the exact form of the forecast bias preferred by managers. For example, if managers are only interested in meeting or beating analysts' latest earnings forecasts, analysts who issue either OP or PP should have more accurate earnings forecasts and are less likely to be fired. In contrast, if managers prefer the OP bias only, OP analysts should have more accurate earnings forecasts and be less likely fired than other analysts.

3. Data and Sample Selection Procedures

---

6 We refer interested readers to Richardson et al. (2004) for a detailed discussion of managers' preferences for biased earnings forecasts.

Our analyst forecast sample comes from the merged IBES actual/detail file over the period January 1, 1983-June 30, 2000. Our sample starts from 1983 because there are very few earnings forecast observations before 1983 in IBES. The sample ends on June 30, 2000 because Regulation FD became effective on October 23, 2000, which significantly changed the communications between firm management and analysts, and our variables are measured from July 1, year t to June 30, t+1 (see section 4 below for the details). We retain only those analysts that work for a U.S.-based brokerage firm and have non-missing values for the following variables in IBES: annual and quarterly earnings forecasts, actual earnings, earnings announcement date, IBES ticker, analyst code, and broker code. We eliminate late annual (quarterly) earnings announcements by deleting the top one percent of the distribution of the distance between the annual (quarterly) earnings announcement and the fiscal year (quarter) end. In addition, we require each firm to have at least 3 analysts following for the quarterly and annual earnings forecasts separately because some of our regression variables cannot be defined or are unreliable for thinly covered stocks. We obtain similar results if each stock is required to have a minimum of 5 analysts following. For annual earnings forecasts, we further require each analyst to issue at least two one-year ahead annual earnings forecasts between two consecutive annual earnings announcement dates; for quarterly earnings forecasts, we require each analyst to issue at least one one-quarter ahead and one two-quarters ahead quarterly earnings forecast for the same fiscal quarter. Our final annual earnings forecast sample contains a maximum of 228,904 firm-analyst-year observations over the period January 1, 1983-June 30, 2000, representing 32,303 analyst-year observations and 7,871 unique analysts. Our final quarterly earnings forecast sample contains a maximum of 114,075 firm-analyst-year-quarter observations over January 1, 1983-June 30, 2000, representing 15,278 analyst-year observations and 4,359 unique analysts. Note that we do not require each analyst to have both annual and quarterly earnings forecasts for the same fiscal year. The significantly smaller sample size for quarterly earnings forecasts is due to the fact that analysts typically do not issue multiple earnings forecasts for the same fiscal quarter before the quarterly earnings announcement. Note our quarterly forecast sample includes earnings forecasts for all four fiscal quarters.

Data on executive insiders' stock sales and purchases, which are required for some of our analyses, come from First Call/Thomson Financial Insider Research Services Historical Files. The insider trading data are available for only calendar years 1985-2000. Data on brokerage firm classification come from the Securities Data Company over the period 1980-2002.

4. Research Design

4.1. Variable Definitions

Because earnings forecast accuracy is measured at the firm-analyst level, H1 is tested at the firm-analyst level. Similarly, because analyst turnover is defined at the analyst level, H2 is tested at the analyst level. As a sensitivity check, we also test H1 using the average values of the regression variables at the analyst level and obtain similar conclusions. We follow Hong and Kubik (2003) for most of our variable definitions. Below we describe the construction of our regression variables. The role of each variable is discussed in Section 4.2.

Figure 1 depicts the timeline we use to construct our variables for the annual earnings forecasts. Because the majority of our sample firms end their fiscal years on December 31, we define analysts' firing over a one-year period from July 1, year t+1 to June 30, year t+2 (denoted year t+1) to ensure that an analyst's firing is based on his performance in the year immediately before July 1, year t+1 (denoted year t).[7] All the other regression variables are constructed using

[7] The percentage of our sample firms whose fiscal year end falls in December, January, February, and March are 66%, 3.5%, 1.3%, and 6.4%, respectively.

data before July 1, t+1. Hong and Kubik (2003) also use July 1 as the cutoff for their analysis of analyst turnover. Our results are robust to alternative cutoffs (e.g., April 1, or January 1).

$Fire_{i,t+1}$ is equal to 1 if analyst i works for a large brokerage house during the year from July 1, t to June 30, t+1, and moves to a small brokerage house during the year from July 1, t+1 to June 30, t+2 (i.e., demotion), or if analyst i permanently leaves the IBES database during the year from July 1, t+1 to June 30, t+2 (i.e., termination); and zero otherwise. Following Hong and Kubik (2003), a brokerage house is large if it employs at least 25 analysts in year t. Because we are interested in how biased forecasts affect analysts' chance of being fired, analysts who move from a small brokerage firm to a large one (i.e., promotion) or move between equal-status brokerage firms (i.e., parallel moves) are coded zero in $Fire_{i,t+1}$. However, we obtain similar conclusions if parallel moves or promotions are coded one in $Fire_{i,t+1}$.[8] We use $Fire_{i,t+1}$ for both the annual and quarterly earnings forecast analyses. Our definition of $Fire_{i,t+1}$ is consistent with Hong et al. (2000) and Leone and Wu (2002).

An important limitation of $Fire_{i,t+1}$ is that we do not know the real causes of an analyst's job change. We assume that demotion and termination are due to current or expected future poor performance, but it is possible that these analysts left their current employers for better opportunities. However, we show below that $Fire_{i,t+1}$ is negatively associated with current earnings forecast accuracy, suggesting that $Fire_{i,t+1}$ represents a reasonable (though noisy) proxy for the true unobservable incidence of firing.

Variables Related to Annual Earnings Forecasts

[8] The percentages of analysts who experience demotion, termination, promotion and parallel moves in our sample are 1.2, 13.9, 1.4 and 4.9, respectively.

A_{jt} denotes firm j's annual earnings for year t that is announced immediately before July 1, year t+1. A_{jt-1} denotes firm j's annual earnings for year t-1. F_{ijt}^{last} is analyst i's latest forecast of annual earnings A_{jt}, issued in the second half of the period from the earnings announcement date of A_{jt-1} to the earnings announcement date of A_{jt}. F_{ijt}^{first} is analyst i's earliest forecast of annual earnings A_{jt} issued in the first half of the period from the earnings announcement date of A_{jt-1} to the earnings announcement date of A_{jt}.

$OP_{i,t}$ denotes analyst's i's optimism-to-pessimism bias in year t and is defined as follows. First, we define a dummy OP_{ijt} that is equal to 1 if F_{ijt}^{first} is greater than A_{jt} (i.e., initial optimism), and F_{ijt}^{last} is less than or equal to A_{jt} (i.e., later pessimism), and zero otherwise.[9] $OP_{i,t}$ is the average of OP_{ijt} for all the firms covered by analyst i in year t. The other annual earnings forecast biases (i.e., OO_{ijt}, PP_{ijt}, PO_{ijt} at the firm-analyst level and $OO_{i,t}$, $PP_{i,t}$, $PO_{i,t}$ at the analyst level) are defined similarly.

$Accuracy_{i,t}$ is the average accuracy of analyst i's earnings forecasts in year t and is defined following Hong and Kubik (2003). Specifically, we first calculate analyst i's absolute forecast error in year t as $FE_{ijt} = | F_{ijt}^{last} - A_{jt} |$. Second, we rank all analysts based on the absolute forecast errors for firm j in year t (denoted $rank_{ijt}$). The most accurate analyst receives a rank of 1, and the least accurate analyst receives the highest rank. If analysts are equally accurate, we assign those analysts the midpoint of the ranks they take up. Third, we develop a ranking score that adjusts for the difference in analyst coverage across different firms:

$$Accuracy_{ijt} = 100 - \frac{rank_{ijt} - 1}{number\ of\ analysts_{j,t} - 1} \times 100. \tag{1}$$

[9] Inference is similar if the observations whose F_{ijt}^{last} is equal to A_{jt} (6.5% of the sample) are deleted.

Thus, $Accuracy_{ijt}$ ranges from zero to 100. $Accuracy_{i,t}$ is the average of $Accuracy_{ijt}$ for all the firms covered by analyst i in year t, representing the average relative forecast accuracy of analyst i in year t.

An alternative measure of forecast accuracy is the absolute forecast accuracy, defined as the absolute forecast error scaled by lagged stock price. We use $Accuracy_{i,t}$ and $Accuracy_{ijt}$ because they are more consistent with our hypotheses and prior research (e.g., Hong et al., 2000; Jacob et al., 1999; Leone and Wu, 2002; Mikhail et al. 1999). For example, both Mikhail et al. (1999) and Hong et al. (2000) show that it is the relative forecast accuracy rather than the absolute forecast accuracy that determines analyst firing. However, as Hong et al. (2000) acknowledge, the relative accuracy measures could be less reliable for analysts who cover few firms or cover thinly followed firms. In addition, analysts who cover fewer firms may be able to spend more time on each firm and thus produce more accurate earnings forecasts. We control for these effects by including $FirmsCovered_{ijt}$ and $Follow_{ijt}$ in the regression model for H1. $FirmsCovered_{ijt}$ is the number of firms (including firm j) followed by analyst i in year t. $Follow_{ijt}$ is the total number of analysts (including analyst i) who follow firm j in year t.

Consistent with prior research (e.g., O'Brien, 1990; Clement, 1999; Jacob et al., 1999; Mikhail et al., 1999; Hong and Kubik, 2003), relative forecast accuracy is defined using F_{ijt}^{last} rather than F_{ijt}^{first}. We believe using F_{ijt}^{last} to define relative forecast accuracy is preferred for several reasons. First, because management is likely to communicate their private earnings information to favored analysts throughout the year, forecast accuracy defined using F_{ijt}^{last} will more completely reflect the effect of issuing biased earnings forecasts on analysts' access to management. Second, the evidence in Mikhail et al. (1999) suggests that analysts' earnings forecast accuracy before earnings announcements (i.e., $Accuracy_{i,t}$) is important to brokerage firms and their investors. Leone and Wu (2002) also find that $Accuracy_{i,t}$ is a significant determinant of institutional investors' All-Star analyst ranking. Finally, even if analysts obtain more private information from management at the beginning of the year, they may not wish to reveal this private information immediately in F_{ijt}^{first} because doing so will erode their competitive advantage later in the year when they issue F_{ijt}^{last}. Arya et al. (2005) further demonstrate that investors may also prefer this strategy because it reduces other analysts' incentive to herd and thus increases the total information available to investors. In untabulated regression analysis we find forecast accuracy defined using F_{ijt}^{last} is a more important determinant of $Fire_{i,t+1}$ than that defined using F_{ijt}^{first}, suggesting F_{ijt}^{last} is the earnings forecast that analysts care the most.

$Bold_{i,t}$ denotes the average boldness of analyst i's earnings forecasts in year t and is defined similarly to $Accuracy_{i,t}$. First, we calculate the consensus earnings forecast (excluding analyst i) as follows:

$$\bar{F}_{-i,j,t}^{first} = \frac{\sum_{m \neq i} F_{m,j,t}^{first}}{number\ of\ analysts_{j,t} - 1}, \tag{2}$$

where -i is the set of analysts other than analyst i. Second, we calculate analyst i's deviation from the consensus, $deviation_{i,j,t} = | F_{i,j,t}^{first} - \bar{F}_{-i,j,t}^{first} |$. Third, we rank all the analysts who cover firm j in year t based on $deviation_{i,j,t}$. Fourth, we use equation (1) to develop a ranking score (denoted $Bold_{ijt}$) that adjusts for the difference in analyst coverage across firms. Finally, $Bold_{i,t}$ is the average of $Bold_{ijt}$ over all the firms covered by analyst i in year t. Intuitively, $Bold_{i,t}$ captures analyst i's deviation from his peers in earnings forecasts.

$Experience_{i,t}$ is the number of years analyst i appears in the IBES annual earnings forecast database as of year t. $FirmExperience_{ijt}$ is the number of years analyst i follows stock j as of year t. $FirmExperience_{i,t}$ is the average of $FirmExperience_{ijt}$ across all the stocks followed by analyst i in year t. GAP_{ijt} is the distance between the earnings announcement date for A_{jt} and the forecast date for F_{ijt}^{last}. $GAP_{i,t}$ is the average GAP_{ijt} for all the firms covered by analyst i in year t. Because $Accuracy_{ijt}$ is expressed in ranking, we also create a similar ranking variable for $FirmsCovered_{ijt}$, $FirmExperience_{ijt}$ and GAP_{ijt}, denoted $R\_FirmsCovered_{ijt}$, $R\_FirmExperience_{ijt}$, and R_GAP_{ijt}, respectively. Similar to $Accuracy_{i,t}$, $FirmExperience_{i,t}$ and $GAP_{i,t}$ are converted into ranking and denoted $R\_FirmExperience_{i,t}$ and $R\_GAP_{i,t}$, respectively.

Variables Related to Quarterly Earnings Forecasts

Note that the analyst turnover definition ($Fire_{i,t+1}$) is identical for the annual and quarterly forecast analyses. To compute the other regression variables needed for the quarterly earnings forecast analysis, we first identify the quarterly earnings announcements made between the two annual earnings announcement dates for A_{jt-1} and A_{jt} in Figure 1, including the earnings announcement for the last fiscal quarter (i.e., announcement date for A_{jt}). Then, for each quarterly earnings announcement (say fiscal quarter 2 of 1998), we identify all the one-quarter ahead and two-quarters ahead quarterly earnings forecasts that are issued after the announcement of the quarterly earnings two quarters prior (i.e., fiscal quarter 4 of 1997) but before the announcement of the current quarterly earnings announcement (i.e., fiscal quarter 2 of 1998). We do not consider three or more quarters ahead quarterly earnings forecasts because there are very few in IBES. Finally, we retain the first (last) quarterly earnings forecast that is issued in the first (second) half of the period between the announcement of the quarterly earnings two quarters prior (i.e., fiscal quarter 4 of 1997) and the announcement of the current quarterly earnings (i.e., fiscal quarter 2 of 1998).

The quarterly equivalents of $Accuracy_{ijt}$, $Bold_{ijt}$, $FirmsCovered_{ijt}$, $FirmExperience_{ijt}$, GAP_{ijt}, $R\_FirmsCovered_{ijt}$, $R\_FirmExperience_{ijt}$, R_GAP_{ijt}, and $Follow_{ijt}$ are computed for each of the quarterly earnings announcements that fall between the two annual earnings announcement dates for A_{jt-1} and A_{jt} in Figure 1. To obtain the yearly equivalents of $OP_{i,t}$, $OO_{i,t}$, $PP_{i,t}$, $PO_{i,t}$, $Accuracy_{i,t}$, $Bold_{i,t}$, and $Experience_{i,t}$, we first compute the mean of each quarterly equivalent across all quarters in year t for each firm-analyst, followed by the averaging of the mean quarterly equivalent across all firms followed by analyst i in year t.

4.2. Regression Models

We use the following OLS regression model to test H1:

$$Accuracy_{ijt} = \alpha_k + \alpha_t + \alpha_1 Bias_{ijt} + Control\ variables_{ijt} + \varepsilon_{ijt} \tag{3}$$

The model is estimated using annual earnings forecasts at the firm-analyst-year level and quarterly earnings forecasts at the firm-analyst-year-quarter level. Therefore, the subscript ‘t’ in the model refers to either yearly or quarterly observations. α_k and α_t are brokerage firm and year fixed effects, controlling for systematic differences in $Accuracy_{ijt}$ across time and brokerage firms. The control variables are $Bold_{ijt}$, $R\_FirmExperience_{ijt}$, $R\_FirmsCovered_{ijt}$, R_GAP_{ijt}, and $\ln(Follow_{ijt})$. $Bold_{ijt}$ controls for the potential effect of forecast boldness on forecast accuracy because Hong et al. (2000) find that bold but inexperienced analysts are more likely to be fired. $R\_FirmExperience_{ijt}$, $R\_FirmsCovered_{ijt}$, and R_GAP_{ijt} control for the effect of analyst i’s firm-specific forecasting experience, number of firms covered, and forecast timing, respectively, on forecast accuracy. Because the dependent variable is a relative measure, these three variables are also defined on relative terms.[10] Because $Follow_{ijt}$ is identical for all the analysts who follow the same firm j, it is not converted to a ranking variable. We use $\ln(Follow_{ijt})$ to allow for a possible nonlinear effect of $Follow_{ijt}$. $Bias_{ijt}$ refers to OP_{ijt}, OO_{ijt}, PP_{ijt}, or PO_{ijt} for both annual and quarterly earnings forecasts. To avoid multicollinearity, the coefficient on PO_{ijt} is suppressed in model (3). If a forecast bias is used to win favor from management, H1 predicts the coefficient on that forecast bias to be larger than the coefficients on the other forecast biases.

We do not include any firm-specific control variables in regression model (3) because $Accuracy_{ijt}$ is relative forecast accuracy for all analysts covering the same firm and thus automatically controls for firm-specific differences. For example, relative forecast accuracy controls for variations in earnings forecast difficulty across companies and time. As another example, firm size may be a determinant of *absolute* forecast accuracy because large firms tend to have a richer information environment. However, firm size should not have an effect on *relative* forecast accuracy because all analysts who cover the firm face the same information environment. Likewise, regression model (3) does not need to control for management's earnings management incentives or public information disclosures (e.g., quarterly earnings announcements) between the annual earnings announcement dates for A_{jt-1} and A_{jt} because such events are common to all analysts who follow the same firm and thus has been controlled for in $Accuracy_{ijt}$.

Because the definitions of $Accuracy_{ijt}$ and $Bias_{ijt}$ use information in the last earnings forecast, the regression model (3) implicitly assumes that an analyst who receives privileged access

[10] Because R_GAP_{ijt} is an important determinant of forecast accuracy, we also allow the effect of R_GAP_{ijt} to differ for each value of R_GAP_{ijt} and obtain similar inference.

to management's private information before issuing his last earnings forecast can credibly commit to firm management that his last earnings forecast will be biased. This seems a reasonable assumption given the intimate and frequent interactions between firm management and financial analysts.

As argued in section 2.2, $Bias_{ijt}$ is also expected to affect $Accuracy_{ijt+1}$. Unfortunately, such effect is not observable for the analysts who do not issue biased forecasts and thus are fired (see H2).[11] Thus, we do not use $Accuracy_{ijt+1}$ in regression model (3). However, as a sensitivity check, we also report the Heckman (1976) regression result of $Accuracy_{ijt+1}$ on $Bias_{ijt}$ in section 5.2.4.

We use the following logit regression model to test H2:

$$Fire_{i,t+1} = \beta_k + \beta_t + \beta_1 Bias_{i,t} + \beta_2 Accuracy_{i,t} + \beta_3 Bold_{i,t} + \beta_4 \ln(Experience_{i,t}) + \varepsilon_{it} \qquad (4)$$

The model is estimated using annual and quarterly earnings forecasts aggregated at the analyst year level. β_k and β_t are brokerage firm and year fixed effects. $Accuracy_{i,t}$ controls for the effect of past forecast accuracy on $Fire_{i,t+1}$, while $\ln(Experience_{i,t})$ controls for an analyst's tenure in the profession. $Bold_{i,t}$ controls for the effect of forecast boldness on analyst turnover. Hong et al. (2000) find that bold but inexperienced analysts are more likely to leave the analyst profession. $Bias_{i,t}$ refers to $OP_{i,t}$, $OO_{i,t}$, $PP_{i,t}$, or $PO_{i,t}$. Again, to avoid multicollinearity, the coefficient on PO_{ijt} is suppressed in model (4). If a forecast bias is used to win favor with management, H2 predicts the coefficient on that forecast bias to be smaller than the coefficients on the other forecast biases. Note that regression model (4) controls for the current period earnings

[11] For our sample, 20% of the analysts who were terminated (i.e., disappeared from the IBES database) did so only after one year of employment.

forecast accuracy $Accuracy_{i,t}$, thus the coefficient on $Bias_{i,t}$ captures the effect of a forecast bias on the probability of firing above and beyond the current period forecast accuracy.

5. Descriptive Statistics and Regression Results

5.1. Descriptive Statistics

Table 1 reports the descriptive statistics for the variables used in regression models (3) and (4). Panels A and B show the variables used in model (3) for the annual and quarterly earnings forecasts, respectively, while Panels C and D show the variables used in model (4) for the annual and quarterly earnings forecasts, respectively.

The unit of observation in Panel A is a firm-analyst-year. The mean values of OP, OO, PP, and PO indicate that the most common annual earnings forecast bias is OO, followed by PP, OP, and PO. Although it is difficult to assess whether the frequencies of the four biases are normal or abnormal in the absence of a clear benchmark, it is striking to observe that the PO bias is the rarest in the sample. The mean analyst has 4.3 years of stock-specific forecasting experience ($FirmExperience_{ijt}$), follows 25.29 stocks ($FirmsCovered_{ijt}$), and covers stocks with 21.07 analysts following ($Follow_{ijt}$).[12] The mean GAP of 78.89 days suggests that the last annual earnings forecast is on average issued after the 3rd fiscal quarter's earnings announcement date. Panel A also reports the distribution of the ranked variables. The mean of each of those ranked variables is 50 by construction.

[12] The distribution of $FirmsCovered_{ijt}$ at the firm-analyst-year level is distorted because the values of $FirmsCovered_{ijt}$ are identical for all the firms covered by analyst i in year t. The mean (median) of $FirmsCovered_{ijt}$ at the analyst-year level is 13.91 (11). This problem also applies to $FirmsCovered_{ijt}$ in Panel B.

The unit of observation in Panel B is a firm-analyst-year-quarter. Had all analysts who are included in Panel A issued at least two quarterly earnings forecasts for each fiscal quarter, the sample size for Panel B should be four times the size in Panel A (i.e., 228,904*4). The smaller sample size of 114,075 in Panel B reflects the fact that analysts issue either zero or only one quarterly earnings forecast for many fiscal quarters. Despite the significant difference in the sample size between Panel A and Panel B, the frequencies of the four forecast biases in Panel B are close to those in Panel A except that the PP bias has the highest frequency. The mean values of $FirmExperience_{ijt}$, $Follow_{ijt}$, and $FirmsCovered_{ijt}$ are similar to those in Panel A. The mean GAP of 48.67 days suggests that the last quarterly earnings forecast is on average issued in the middle of two consecutive quarterly earnings announcement dates.

The unit of observation in Panel C is an analyst-year. The mean $Fire_{i,t+1}$ indicates that 15% of the analysts are fired over our sample period, a nontrivial percentage. Untabulated analyses further indicate that among the fired analysts in our sample, 20.2% of them are fired in the second year of their career, 22.47% in the third year of their career, 14.59% in the fourth year of their career, and 9.49% in the fifth year of their career. Clearly, the majority of the firing occurs in an analyst's early stage of his career. The distributions of the four forecast biases are similar to those in Panel A. The mean analyst has been in the analyst profession for 5.01 years ($Experience_{i,t}$).

The unit of observation in Panel D is an analyst-year. Due to the sample size difference, the mean $Fire_{i,t+1}$ is slightly smaller than that in Panel C. The distributions of the four forecast biases are similar to those in Panel B. The distribution of $Experience_{i,t}$ is approximately one year higher than that in Panel C.

Table 2 reports the Spearman (top diagonal) and Pearson (bottom diagonal) correlations for the key regression variables in models (3) and (4) using observations at the analyst-year level.

Because the correlations are similar for both Spearman and Pearson, we focus on the Pearson correlations (bottom diagonal) in the following discussion.

$Accuracy_{i,t}^{A}$ is the relative earnings forecast accuracy ($Accuracy_{i,t}$) using annual earnings forecasts while $Accuracy_{i,t}^{Q}$ is the relative earnings forecast accuracy ($Accuracy_{i,t}$) using quarterly earnings forecasts. The other variables in Table 2 are similarly defined. The correlation between $Accuracy_{i,t}$ and $OP_{i,t}$ is significantly positive for both annual and quarterly forecasts, but the correlation between $Accuracy_{i,t}$ and any of the other three biases is either significantly negative or insignificant. These univariate correlations are consistent with the hypothesis that analysts use $OP_{i,t}$ forecasts to gain better access to management's private information. In addition, the significantly positive correlation between $OP_{i,t}^{A}$ and $OP_{i,t}^{Q}$ suggests that analysts often issue both annual and quarterly OP earnings forecasts to please management.

$Fire_{i,t+1}$ is significantly negatively correlated with $OP_{i,t}$ for both annual and quarterly forecasts. Except for the marginally significantly negative correlation between $Fire_{i,t+1}$ and $PP_{i,t}^{A}$, the correlation between $Fire_{i,t+1}$ and any of the other forecast biases is either insignificant or significantly positive. These univariate correlations are consistent with the hypothesis that analysts who issue annual and quarterly OP earnings forecasts are less likely to be fired. This evidence is consistent with the univariate correlations for $Accuracy_{i,t}$.

5.2 Regression Results

5.2.1. H1

Table 3 reports the OLS regression results for H1. Panel A reports the results for annual earnings forecasts while Panel B shows the results for quarterly earnings forecasts. The standard errors are adjusted for heteroskedasticity and correlations for observations of the same stocks using the method of Rogers (1993).

Results for Annual Earnings Forecasts

Column (1) of Panel A shows that relative to PO analysts' forecast accuracy, OP analysts' annual earnings forecasts are more accurate while OO analysts' forecasts are less accurate and PP analysts' forecasts are equally accurate. In addition, the coefficient on OP is significantly larger than those of OO and PP (two-tailed p<0.001). These results are consistent with the hypothesis that analysts use OP forecasts to gain better access to management's private information. The significantly negative coefficient on OO and the insignificant coefficient on PP are inconsistent with the hypothesis that analysts issue consistently optimistic or pessimistic annual earnings forecasts to gain better access to management.[13]

The negative coefficient on $Bold_{ijt}$ suggest that bolder analysts produce less accurate earnings forecasts. The coefficient on $FirmExperience_{ijt}$ is significantly positive, suggesting that experienced analysts produce more accurate forecasts, a finding consistent with Clement (1999). As expected, forecasts issued closer to the earnings announcement date are more accurate. We do not offer any economic interpretation on the coefficients on $R\_FirmsCovered_{ijt}$ and $\ln(Follow_{ijt})$ because they mainly control for the limitations of $Accuracy_{ijt}$ for analysts who follow few firms or thinly covered firms.

[13] An alternative earnings forecast optimism definition used in prior research is defined relative to the consensus earnings forecast of the other analysts who follow the same firm (see e.g., Hong and Kubik, 2003). Including this alternative optimism definition in models (3) and (4) does not alter any of our inferences. In addition, the coefficient on this alternative optimism is significantly negative in model (3) and significantly positive in model (4), suggesting that optimistic analysts produce less accurate earnings forecasts and are more likely to be fired, inconsistent with the hypothesis that analysts use optimistic earnings forecasts to please firm management for more private information.

Because only the coefficient on OP in column (1)'s regression is consistent with H1, column (2) of Panel A reports the regression in column (1) after dropping OO and PP. As expected, the coefficient on OP continues to be significantly positive. The result in column (3) is discussed in section 5.2.3.

Results for Quarterly Earnings Forecasts

Column (1) of Panel B reports the regression coefficients of model (3) for quarterly earnings forecasts. The coefficients on both the control variables and the four forecast biases are consistent with those in column (1) of Panel A. Column (2) of Panel B reports the regression result without OO and PP. As expected, the coefficient on OP remains significantly positive. Overall, the evidence in Panels A and B is consistent with Richardson et al. (2004) who find that managers prefer OP consensus earnings forecasts. The result in column (3) is discussed in section 5.2.3.[14]

5.2.2. H2

Table 4 reports the logit regression results for H2. Panel A reports the results for annual earnings forecasts while Panel B shows the results for quarterly earnings forecasts. Panel C combines the regression variables in Panels A and B into one regression. The standard errors in table 4 are adjusted for heteroskedasticity and correlations for observations of the same brokers using the method of Rogers (1993).

Results for Annual Earnings Forecasts

[14] Including the relative earnings forecast accuracy defined using the initial earnings forecast F_{ijt}^{first} in regression model (3) does not affect the coefficient on OP in Table 3, suggesting that the positive coefficient on OP is not because OP analysts are inherently more accurate than other analysts. In addition, the coefficient on OP is robust to controlling for the ranked signed difference between the reported earnings and an individual analyst's initial or last earnings forecast (defined in the same way as $Accuracy_{ijt}$).

Column (1) of Panel A reports the regression coefficients of model (4) using annual earnings forecasts. Consistent with prior research, more accurate and more experienced analysts are less likely to be fired. The coefficient on $Bold_{i,t}$ is insignificant. The coefficient on OP is significantly negative but the coefficients on OO and PP are insignificant. In addition, the coefficient on OP is significantly larger in magnitude than those on OO and PP (two-tailed p=0.01 or lower). Because model (4) controls for current forecast accuracy, the significant regression coefficient on OP suggests that OP analysts are less likely to be fired presumably because of their improved <u>future</u> earnings forecast accuracy relative to other analysts (see section 5.2.4 for direct evidence). The insignificant coefficients on OO and PP further suggest that consistently issuing optimistic or pessimistic annual earnings forecasts alone is not sufficient to reduce the probability of firing. As a sensitivity check, column (2) of Panel A reports the coefficients of model (4) after dropping OO and PP. Not surprisingly, the coefficient on OP remains significantly negative. The result in column (3) is discussed in section 5.2.3.

Results for Quarterly Earnings Forecasts

Column (1) of Panel B reports the regression coefficients of model (4) for quarterly earnings forecasts. The coefficients on the control variables are consistent with those in column (1) of Panel A. Consistent with the coefficients in Panel A, the coefficients on OP and OO are significantly negative and insignificant, respectively. There is weak evidence at the 10% two-tailed significance level that PP analysts are less likely to be fired relative to the benchmark PO analysts. However, the coefficient on PP is significantly smaller in magnitude than that on OP (two-tailed p=0.05). In addition, as shown in column (2) of Panel B, the effect of OP dominates the other three biases as the coefficient on OP remains significantly negative after the omission of OO and PP in the regression. Overall, the results for the quarterly forecasts are consistent with those for the annual forecasts. The result in column (3) is discussed in section 5.2.3.

Results for Annual and Quarterly Earnings Forecasts Combined

To determine the incremental effect of $OP_{i,t}^{A}$ and $OP_{i,t}^{Q}$ on the probability of firing, Panel C of Table 4 reports the coefficients of model (4) by combining the independent variables in column (1) of Panels A and B. The sample size in this regression is smaller than that in Panel A or Panel B because not all analysts issue both annual and quarterly earnings forecasts for the same fiscal year. The coefficients on the control variables remain in the same directions as those in Panels A and B and significant except for the insignificant coefficient on $Accuracy_{i,t}^{Q}$. Thus, once controlling for the annual earnings forecast accuracy, the quarterly earnings forecast accuracy matters little in the probability of firing. The coefficients on $OP_{i,t}^{A}$ and $OP_{i,t}^{Q}$ are both significantly negative but are not significantly different from each other (two-tailed p=0.59), suggesting that both the annual and quarterly OP biases are associated with the probability of firing.

5.2.3. Further Tests of H1 and H2

Regression models (3) and (4) assume that analysts have incentives to use biased earnings forecasts to please managements of all firms. However, as discussed in section 2.2, the preference for biased earnings forecasts should be stronger for managers who need to sell significant amounts of their personal equity holdings in the firm regularly. Thus, these managers should have a stronger incentive to trade their private information for analysts' biased earnings forecasts. In addition, we also expect the predicted effect of biased forecasts on relative forecast accuracy and the probability of firing to be stronger for firms with difficult-to-forecast earnings. This is because when earnings are easy to predict and thus all analysts' earnings forecasts are already very accurate, having management’s private information will not enable an analyst to significantly improve his relative forecast accuracy. The converse is true when earnings are difficult to predict.

The last column of Table 3 reports the regression results of model (3) allowing the coefficient on OP_{ijt} (annual forecasts in Panel A and quarterly forecasts in Panel B) to vary with the insider trading intensity (denoted $InsiderSell_{ijt}$) and the degree of earnings forecasting difficulty (denoted $Dispersion_{ijt}$). For both the annual and quarterly samples, $InsiderSell_{ijt}$ is a dummy that is equal to 1 if the average net insider selling (expressed in 1982 dollars) by all corporate officers and directors for firm j followed by analyst i during the calendar year immediately before the earnings announcement date for $Accuracy_{ijt}$ is larger than the 75th percentile of our sample. For the annual sample, $InsiderSell_{i,t}$ is the average of $InsiderSell_{ijt}$ over all the firms covered by analyst i in year t. For the quarterly sample, $InsiderSell_{i,t}$ is defined as the mean of $InsiderSell_{ijt}$ across all quarters in year t for each firm-analyst, followed by the averaging of the above mean across all firms covered by analyst i in year t.[15]

Because we wish to capture the ex ante effect of insider selling, $InsiderSell_{i,t}$ is measured before $Accuracy_{i,t}$ and $Fire_{i,t+1}$ (the dependent variables for H1 and H2 respectively). Using insider sales after the measurement of the dependent variables is problematic because insiders tend to sell (buy) after positive (negative) earnings surprises. In addition, insiders should continue to have an incentive to report earnings increases immediately after their stock sales in order to avoid the perception of illegal insider trading. Therefore, $InsiderSell_{i,t}$ should be a reasonable proxy for

[15] Aboody and Kasznik (2000) find that corporate executives manage voluntary disclosures to depress stock prices immediately before new stock option grants. Because new option grants are unavailable for all of our sample firms, they are not included in $InsiderSell_{ijt}$. As a result, our insider selling measure likely understates the true effect of the insider selling incentive.

insiders' ex ante preference for biased earnings forecasts. The correlation between $InsiderSell_{i,t}$ and $InsiderSell_{i,t+1}$ is very high (the Pearson correlation is 62% for our sample).

$Dispersion_{ijt}$ is a dummy that is equal to 1 if the forecast dispersion (defined as the standard deviation of the earnings forecasts scaled by the magnitude of the realized earnings) is greater than the 75th percentile of our sample.[16] For both the annual and quarterly samples, $Dispersion_{ijt}$ is computed using each analyst's first earnings forecast, although results are similar if each analyst's last earnings forecast is used instead. $Dispersion_{i,t}$ is the average of $Dispersion_{ijt}$ over all the firms covered by analyst i in year t and defined similarly to $InsiderSell_{i,t}$.

Note that $InsiderSell_{ijt}$ and $Dispersion_{ijt}$ are not defined as continuous variables because the effects of insider selling and forecast difficulty are likely nonlinear. In addition, continuous measures of $InsiderSell_{i,t}$ and $Dispersion_{i,t}$ could be unduly influenced by a few of the stocks followed by analysts i in year t. Untabulated sensitivity checks indicate that the interaction results for $InsiderSell_{i,t}$ and $Dispersion_{i,t}$ are robust to alternative cutoffs (e.g., 66th, 70th, or 80th percentile), but become insignificant when $InsiderSell_{ijt}$ and $Dispersion_{ijt}$ are defined as continuous variables.

Consistent with our predictions, the coefficients on $OP_{ijt} \times InsiderSell_{ijt}$ and $OP_{ijt} \times Dispersion_{ijt}$ in both Panels A and B of Table 3 are significantly positive with the exception of the positive but insignificant coefficient on $OP_{ijt} \times Dispersion_{ijt}$ in Panel B. The results suggest

[16] Because of zero realized earnings, $Dispersion_{ijt}$ is not defined for 298 firm-analyst-year observations in the annual forecast sample and 462 firm-analyst-year-quarter observations in the quarterly forecast sample. $Dispersion_{ijt}$ is set equal to 1 in those cases.

that the positive effect of issuing OP annual and quarterly earnings forecasts on relative forecast accuracy is stronger for firms with heavy insider sales and hard-to-predict earnings.

The negative coefficients on $InsiderSell_{ijt}$ and $Dispersion_{ijt}$ in Table 3 are expected and consistent with H1 because they reflect the effect of these two variables for only analysts who do not issue OP forecasts. For example, for a low forecast dispersion firm, management's private information should matter less in determining the ranking of the analysts who follow the firm; therefore analysts who do not issue OP forecasts are not going to suffer significantly in forecast accuracy relative to those who issue OP forecasts. In contrast, for a high dispersion firm, management's private information matters more in the ranking and therefore those analysts who do not issue OP forecasts are going to suffer more in forecast accuracy relative to the OP analysts who cover the same firm. Therefore, we should expect non-OP analysts' relative earnings forecast accuracy to be lower for high dispersion firms than for low dispersion firms. A similar reasoning applies to $InsiderSell_{ijt}$. The negative coefficients on $InsiderSell_{ijt}$ and $Dispersion_{ijt}$ do not conflict with our argument in section 4.2 that firm-specific variables should not affect $Accuracy_{ijt}$ when included alone. We have verified that the coefficients on $InsiderSell_{ijt}$ and $Dispersion_{ijt}$ are insignificant when OP_{ijt} , $OP_{ijt} \times InsiderSell_{ijt}$ and $OP_{ijt} \times Dispersion_{ijt}$ are omitted from the interaction model in Table 3.

The last column of Table 4 reports the regression results of model (4) allowing the coefficients on $OP_{i,t}$ to vary with $InsiderSell_{i,t}$ and $Dispersion_{i,t}$. As predicted, the coefficients on $OP_{i,t} \times InsiderSell_{i,t}$ and $OP_{i,t} \times Dispersion_{i,t}$ in Panels A and B of Table 4 are significantly negative except for the insignificant coefficient on $OP_{i,t} \times Dispersion_{i,t}$ in Panel B. These results suggest that the negative effect of issuing annual and quarterly OP forecasts on the probability of firing is stronger for firms with heavy insider sales and hard-to-predict earnings. Overall, the results from the interaction models in Tables 3 and 4 provide further support for our hypotheses.

Because we find little evidence in column (1) of tables 3 and 4 (panels A and B) that OO and PP are associated with improved forecast accuracy and a smaller probability of firing, the interaction models in column (3) of tables 3 and 4 do not allow the coefficients on OO and PP to vary with the insider selling and forecast dispersion variables. As a sensitivity check, we rerun the interaction models in tables 3 and 4 by allowing the coefficients on OO and PP to vary with the insider selling and forecast dispersion variables (results not tabulated). For the annual sample in panel A of table 3, the coefficient on $OP_{ijt} \times InsiderSell_{ijt}$ is larger (i.e., consistent with H1) than the coefficients on $OO_{ijt} \times InsiderSell_{ijt}$ and $PO_{ijt} \times InsiderSell_{ijt}$ but not different from the coefficient on $PP_{ijt} \times InsiderSell_{ijt}$ at the 10% one-tailed level or better; the coefficient on $OP_{ijt} \times Dispersion_{ijt}$ is larger than the coefficient on $PP_{ijt} \times Dispersion_{ijt}$ but not different from the coefficients on $OO_{ijt} \times Disperson_{ijt}$ and $PO_{ijt} \times Dispersion_{ijt}$ at the 10% one-tailed level or better. For the quarterly sample in panel B of table 3, the coefficient on $OP_{ijt} \times InsiderSell_{ijt}$ is significantly larger than the coefficients on $OO_{ijt} \times InsiderSell_{ijt}$, $PP_{ijt} \times InsiderSell_{ijt}$, and $PO_{ijt} \times InsiderSell_{ijt}$ at the 10% one-tailed level or better, but the coefficient on $OP_{ijt} \times Dispersion_{ijt}$ is never significantly larger than any of the other three dispersion interactions at the 10% one-tailed level.

For the annual sample in panel A of table 4, the coefficient on $OP_{i,t} \times InsiderSell_{i,t}$ is significantly smaller (i.e., consistent with H2) than the coefficients on $OO_{i,t} \times InsiderSell_{i,t}$ and $PP_{i,t} \times InsiderSell_{i,t}$ but not different from the coefficient on $PO_{i,t} \times InsiderSell_{i,t}$ at the 10% one-tailed level or better; the coefficient on $OP_{i,t} \times Dispersion_{i,t}$ is smaller than the coefficients on

$OO_{i,t} \times Disperson_{i,t}$ and $PP_{i,t} \times Dispersion_{i,t}$ but not different from the coefficient on $PO_{i,t} \times Dispersion_{i,t}$ at the 10% one-tailed level or better. For the quarterly sample in panel B of table 4, the coefficient on $OP_{i,t} \times InsiderSell_{i,t}$ is significantly smaller than the coefficients on $OO_{i,t} \times InsiderSell_{i,t}$ and $PP_{i,t} \times InsiderSell_{i,t}$ but not different from the coefficient on $PO_{i,t} \times InsiderSell_{i,t}$ at the 10% one-tailed level or better; but the coefficient on $OP_{i,t} \times Dispersion_{i,t}$ is not different from any of the other dispersion interactions at the 10% one-tailed level. Overall, the results from above sensitivity checks are broadly consistent with the reported interaction models in tables 3 and 4 but weaker in significance because of the separation of the control group into three subgroups.[17]

To gauge the economic significance of issuing OP earnings forecasts on analysts' forecast accuracy and job security, we compute the marginal effects of OP for the annual earnings forecast regressions in Panel A of Tables 3 and 4. The coefficient on OP_{ijt} in Panel A, column (2) of Table 3 (6.530) indicates that a one standard deviation increase in OP_{ijt} is associated with an increase in relative forecast accuracy of 2.86 (i.e., 6.530*0.438). For analysts who cover stocks with heavy insider selling and difficult-to-forecast earnings (defined as observations whose values of $InsiderSell_{ijt}$ and $Dispersion_{ijt}$ are equal to one), a one standard deviation increase in OP_{ijt} is associated with an increase in relative forecast accuracy of 3.33 (i.e., [6.079+0.781+0.736]*0.438). As a comparison, a one standard deviation increase in $R\_FirmExperience_{i,t}$ in Panel A, column (2) of Table 3 is associated with an increase in relative forecast accuracy of only 0.41 (i.e., 0.013*31.43).

[17] As a sensitivity check, we also replaced OP in the regressions of columns (2) and (3) of tables 3 and 4 with either OO, PP, or PO. We found no evidence consistent with H1 and H2 for any of those biases.

The coefficient on $OP_{i,t}$ in Panel A, column (2) of Table 4 indicates that a one standard deviation increase in $OP_{i,t}$ is associated with a decrease in the probability of firing by 0.99% evaluated at the mean values of the independent variables. For analysts who cover stocks with heavy insider selling and difficult-to-forecast earnings (defined as observations whose values of $InsiderSell_{i,t}$ and $Dispersion_{i,t}$ exceed the 75th percentile of the sample), a one standard deviation increase in $OP_{i,t}$ is associated with a decrease in the probability of firing by 1.45% evaluated at the mean values of the independent variables. Because the mean unconditional probability of firing is 15% (see Table 1, Panel C), increasing $OP_{i,t}$ by one standard deviation will reduce the probability of firing by 9.7% (i.e., 1.45/15). As a comparison, the coefficient on $Accuracy_{i,t}$ in Panel A, column (2) of Table 4 indicates that a one standard deviation increase in $Accuracy_{i,t}$ is associated with a decrease in the probability of firing by 3.91% evaluated at the mean values of the independent variables. It should be noted that the effect of $Accuracy_{i,t}$ partially reflects the effect of $OP_{i,t}$ because OP analysts also produce more accurate contemporaneous earnings forecasts.

5.2.4. The Effect of Issuing Biased Earnings Forecasts on Future Earnings Forecast Accuracy

As part of the motivation for H2 in section 2.1, we assume that analysts who do not issue biased earnings forecasts will suffer in their future earnings forecast accuracy, even after controlling for current forecast accuracy. We use the following regression model to offer direct evidence on this hypothesis for the annual and quarterly earnings forecasts separately:

$$\begin{aligned} Accuracy_{i,t+1} &= \alpha_k + \alpha_{t+1} + \alpha_1 Bias_{i,t} + \alpha_2 Accuracy_{i,t} + \alpha_3 Bold_{i,t+1} + \alpha_4 \ln(Follow_{i,t+1}) \\ &+ \alpha_5 R\_FirmsCovered_{i,t+1} + \alpha_6 R\_FirmExperience_{i,t+1} + \alpha_7 R\_GAP_{i,t+1} + \varepsilon_{i,t+1} \end{aligned} \tag{5}$$

The above model is similar to model (3) except for the addition of $Accuracy_{i,t}$. In addition, model (5) can only be estimated using the surviving analysts because analysts who do not issue biased earnings forecasts are more likely to be fired. To produce consistent estimates of the regression coefficients of model (5), we use regression model (4) without the year and broker fixed effects to correct for the sample selection bias (see Heckman, 1976). Because regression model (4) is estimated at the analyst year level, the unit of observation for model (5) is also an analyst year. $Bias_{i,t}$ refers to the $OP_{i,t}$ bias and is predicted to be positive. The other variables are defined in section 4.1.

Table 5 reports the regression coefficients of model (5) for annual (Panel A) and quarterly (Panel B) earnings forecasts. The standard errors are adjusted for heteroskedasticity and correlations for observations of the same brokers using the method of Rogers (1993).

For both the annual and quarterly earnings forecasts, the coefficients on the control variables are consistent with those in Table 3 and generally significant. As expected, the coefficient on $Accuracy_{i,t}$ is significantly positive in both panels. The coefficient on $OP_{i,t}$ is significantly positive for the annual earnings forecasts in Panel A but insignificant (though positive) for the quarterly earnings forecasts in Panel B (two-tailed p=0.13). The weaker coefficient on $OP_{i,t}$ in Panel B could be caused by the smaller sample size. Another reason is that not all analysts issue multiple quarterly earnings forecasts for every fiscal quarter (see footnote 1) and thus the values of $OP_{i,t}$ and $Accuracy_{i,t+1}$ could be computed for different mixes of firms, which should weaken the association between $OP_{i,t}$ and $Accuracy_{i,t+1}$. Overall, the results in Table 5 are consistent with the hypothesis that OP analysts produce more accurate future earnings forecasts, even after controlling for the current earnings forecast accuracy. This evidence offers one rationale for why the coefficient on $OP_{i,t}$ in model (4) is negative even after controlling for current forecast accuracy.

5.2.5. Investment Banking Incentive As an Alternative Explanation

Popular press (see e.g., Gasparino, 2002) alleges that analysts use biased earnings forecasts to help their employers win more investment banking businesses. The record settlement between U.S. government regulators and the ten largest securities firms in 2003 directly targets securities firms' alleged abuses of using biased analyst research to win investment-banking business. While several studies (e.g., Michaely and Womack, 1999; Dugar and Nathan, 1995; Lin and McNichols, 1998; Bradshaw et al., 2003) finds evidence supporting the above allegation, a few recent studies (e.g., Cowen et al., 2006; Jacob et al., 2003) find no such evidence.

Because analysts who work for investment banks may have better access to management's private information during the underwriting process of existing clients or during the competition for new clients, our H1 and H2 are potentially consistent with the investment banking incentive. However, such associations are spurious (not causal) because an analyst's primary purpose for issuing biased earnings forecasts is not to obtain management's private information to improve forecast accuracy. Instead, improved forecast accuracy is merely a byproduct of analysts' effort to use biased earnings forecasts to win more investment banking deals.

To determine whether the hypothesized effects of H1-H2 are solely motivated by the investment banking incentive, we rerun regression models (3) and (4) for both annual and quarterly earnings forecasts by allowing the coefficient on OP to vary with $Bookrunner_{i,t}$, a dummy variable that is equal to 1 if a brokerage house served as an equity offering book runner in at least 11 out of the 23 years from 1980 to 2002 (denoted book runner), and 0 if a brokerage house never derived any revenues from investment banking over 1980-2002 (denoted pure brokerage firm). We also tried 15 years and 23 years as cutoffs and obtained similar results. Brokerage firms who served as book runners for fewer than 11 years or only as syndicates over 1980-2002 are excluded from this analysis because the influence of investment banking business is unclear for these firms, although inference is similar if those brokerage firms are combined with the book runners or pure brokerage firms.

If the investment banking incentive is the driver of biased earnings forecasts, the predicted effects of H1 and H2 should not exist for analysts who work for pure brokerage firms. Untabulated regression results find no evidence that the predicted effects of H1 and H2 are stronger for analysts who work for investment banks than for those who work for pure brokerage firms. Thus, the documented results for H1 and H2 cannot be solely explained by the investment banking incentive. However, we cannot rule out the possibility that the predicted effects of H1 and H2 for the investment bank analysts are partially related to the investment banking incentive.

5.2.6. Who Are the OP Analysts?

The results in the previous sections show that analysts who issue OP forecasts produce more accurate earnings forecasts and are less likely to be fired. Thus, a natural question to ask is why not all analysts issue OP forecasts. Section 2.1 offers several plausible explanations. One testable explanation is the "bang for the buck" hypothesis. This hypothesis states that managers will play the biased earnings forecast game only with analysts who can exert a significant influence on investors' expectations. Prior research (see e.g., Jacob et al., 1999; Mikhail et al., 1997; Stickel, 1992) indicates that analysts that are more experienced, from large brokerage houses, and an All-

Star as rated by the Institutional Investor magazine are more influential among investors. Thus, we expect those analysts to be more likely to issue OP forecasts.

Table 6 reports test results consistent with this hypothesis based on the larger annual earnings forecast sample. The unit of observation is an analyst year. Panel A reports the univariate statistics of analyst characteristics by high and low OP using a cutoff of the median OP, while Panel B reports the regression of OP on the multiple analyst characteristics. The regression model also controls for year fixed effects and adjusts the coefficient standard errors for heteroskedasticity and dependence of observations of the same brokerage firms per Rogers (1993). The dependent variable OP is multiplied by 100 in Panel B to increase the precision of the reported regression coefficients. $FirmExperience_{i,t}$ is defined as before. $Brokersize_{i,t}$ is defined as the number of unique analysts that belong to brokerage firm i in year t. $AllStar_{i,t}$ is a dummy variable that is coded one if an analyst is an All-Star as ranked by the Institutional Investors magazine in the prior year, and zero otherwise. Consistent with the hypothesis, Panel A of Table 6 shows that high OP analysts are more experienced, employed by larger brokerage firms, and more likely to be an All-Star. The results from the multiple variable regression in Panel B of Table 6 are consistent with the descriptive statistics in Panel A.

6. Conclusion

Analysts are often alleged to use biased earnings forecasts to please management, but the form of the earnings forecast bias analysts use and the benefits analysts receive from issuing biased forecasts are not clearly identified. We hypothesize that analysts use biased earnings forecasts to gain better access to management's private information to improve their earnings forecast accuracy and job security. Based on prior research, we consider four earnings forecast biases that analysts could use to please firm management (denoted OP, OO, PP, and PO). OP denotes individual analysts whose initial earnings forecasts are optimistic (i.e., forecast is greater than the realized earnings) but whose last earnings forecasts before the earnings announcement are pessimistic (i.e., forecast is no greater than the realized earnings); OO denotes analysts whose initial and last forecasts are both optimistic while PP denotes analysts whose initial and last forecasts are both pessimistic; finally PO denotes analysts whose initial earnings forecasts are pessimistic but whose last forecasts are optimistic. We test our research questions using both annual and quarterly earnings forecasts because individual analysts often issue both annual and quarterly earnings forecasts and thus it is interesting to examine whether the forecast bias analysts use to please management varies across forecast horizon.

We find that analysts who issue both annual and quarterly OP forecasts have more accurate current and future earnings forecasts relative to other analysts and are less likely to be fired by their employers. These effects are stronger for firms with heavy insider sales and hard-to-predict earnings. In addition, we find that those results hold for analysts employed by both investment banks and pure brokerage firms without investment banking business. Taken together, these empirical results are consistent with the hypothesis that analysts use the OP bias to please firm management to gain better access to management's private information. Further analyses indicate that OP analysts are more experienced, employed by larger brokerage firms and more likely to be an All-Star. The characteristics of the OP analysts are consistent with the hypothesis that management is more willing to play the biased earnings forecast game with analysts who have more influence on capital market investors.

Despite the robust and consistent empirical results for H1 and H2, our results should be interpreted with caution because we merely document associations and thus our results could be subject to unknown alternative explanations. In addition, the regression results for H2 should be interpreted with caution because the construct validity of the dependent variable (Firing) cannot be independently verified.

Regulation FD has significantly changed the private communication between firm management and financial analysts. Future research may study how Regulation FD affects analysts' incentives to use biased earnings forecasts to gain better access to management's private information. Although recent research (see e.g., Gintschel and Markov, 2004) shows that Regulation FD significantly reduces the amount of private information analysts receive from firm management, it remains unclear whether the private communication between management and analysts has been completely cut off. For instance, Regulation FD still allows managers to disclose nonmaterial nonpublic information to analysts. As the SEC recognizes, such nonmaterial information could be combined with analysts' own private information to generate material new insights. As a result, firm management may still have substantial leverage in pressing analysts to issue biased earnings forecasts to gain access to their private information.

References

Aboody, D., Kasznik R., 2000. CEO stock option awards and the timing of corporate voluntary disclosures. Journal of Accounting & Economics 29, 73 – 100.

Arya, A., J. Glover, and B. Mittendorf, 2005. Hierarchical Reporting, Aggregation, and Information Cascades, Managerial and Decision Economics, Forthcoming.

Bartov, E., Givoly D., and Hayn, C., 2002. The rewards to meeting or beating earnings expectations. Journal of Accounting&Economics 33, 173 – 204.

Bettis, J., Coles J., Lemmon M., 2000. Corporate policies restricting trading by insiders. Journal of Financial Economics 57, 191 – 220.

Bowen, R., Davis A., Matsumoto D., 2002. Do Conference Calls Affect Analysts' Forecasts? The Accounting Review 77, 285 – 317.

Bradshaw, M., Richardson S., Sloan R., 2003. Pump and dump: an empirical analysis of the relation between corporate financing activities and sell-side analyst research. Working paper.

Chen, S., Matsumoto D., 2006. Upgrades vs downgrades: the impact on access to information and analysts' forecast accuracy. Journal of Accounting Research, forthcoming.

Clement, M., 1999. Anlayst forecast accuracy: do ability, resources, and portfolio complexity matter? Journal of Accounting and Economics 27, 285 – 304.

Cowen, A., Groysberg B., Healy P., 2006. Which types of analyst firms make more optimistic forecasts? Journal of Accounting and Economics 41, 119-146.

Das, S., Levine C., Sivaramakrishnan K., 1998. Earnings predictability and bias in analysts' earnings forecasts. The Accounting Review 73, 277 – 295.

Dugar, A. and Nathan S., 1995. The effects of investment banking relationships on analysts' earnings forecasts and investment recommendations. Contemporary Accounting Research 12, 131 – 161.

Eames, M., Glover S., 2003. Earnings predictability and the direction of analysts' earnings forecast errors. The Accounting Review 78, 707 – 725.

Francis, J., Philbrick D., 1993. Analysts' decisions as products of a multi-task environment. Journal of Accounting Research 31, 216 – 231.

Gasparino, C. 2002. Salomon's Gruberman is now target as lawyers continue analyst probe. The Wall Street Journal, April 12.

Gintschel, A. and Markov, S., 2004. The effectiveness of Regulation FD. Journal of Accounting&Economics, 37(3), 293-314.

Heckman, J. 1976. The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. The Annals of Economic and Social Measurement 5: 475-492.

Hong, H., Kubik J., Solomon A., 2000. Security analysts' career concerns and herding of earnings forecasts. The Rand Journal of Economics 31, 121 – 145.

Hong, H., Kubik J., 2003. Analyzing the analysts: career concerns and biased earnings forecasts. The Journal of Finance 58, 313 – 352.

Jacob, J., Lys T., Neale M., 1999. Expertise in forecasting performance of security analysts. Journal of Accounting and Economics 28, 51 – 82.

Jacob, J., Rock S., Weber D., 2003. Do analysts at independent research firms make better earnings forecasts? Working paper, The University of Colorado.

Levitt, A. 1998. The numbers game. Remarks delivered at the NYU Center for Law and Business, New York, NY.

Krigman, L., W. Shaw, and K. Womack. 2001. Why do firms switch underwriters? Journal of Financial Economics 60, 245-284.

Leone, A., Wu J., 2002. What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts. Working paper, The University of Rochester.

Lim, T., 2001. Rationality and analysts' forecast bias. The Journal of Finance 56, 369 – 386.

Lin, H., M. McNichols. Underwriting relationship, analysts' earnings forecasts and investment recommendations. Journal of Accounting&Economics 25, 101 – 128.

Matsumoto, D., 2002. Management's incentives to avoid negative earnings surprises. The Accounting Review 77, 483-514.

Michaely, R., Womack K., 1999. Conflict of interest and the credibility of underwriter analyst recommendations. The Review of Financial Studies 12, 653 – 686.

Mikhail, M., B. Walther and R. Willis. 1997. Do security analysts improve their performance with experience? Journal of Accounting Research 35, 131-157.

Mikhail, M., Walther B., Willis R., 1999. Does forecast accuracy matter to security analysts? The Accounting Review 74, 185 – 191.

O’Brien, P., 1990. Forecast accuracy of individual analysts in nine industries. Journal of Accounting Research 28, 286 – 304.

Richardson, S., Teoh S., Wysocki P., 2004. The walkdown to beatable analysts’ forecasts: the role of equity issuance and inside trading incentives. Contemporary Accounting Research, forthcoming.

Rogers, W. H. 1993. sg17: Regression standard errors in clustered samples. *Stata TechnicalBulletin* 13: 19-23.

Roulsone, D.T., 2003. The relation between insider-trading restrictions and executive compensation. Journal of Accounting Research 41, 525 – 551.

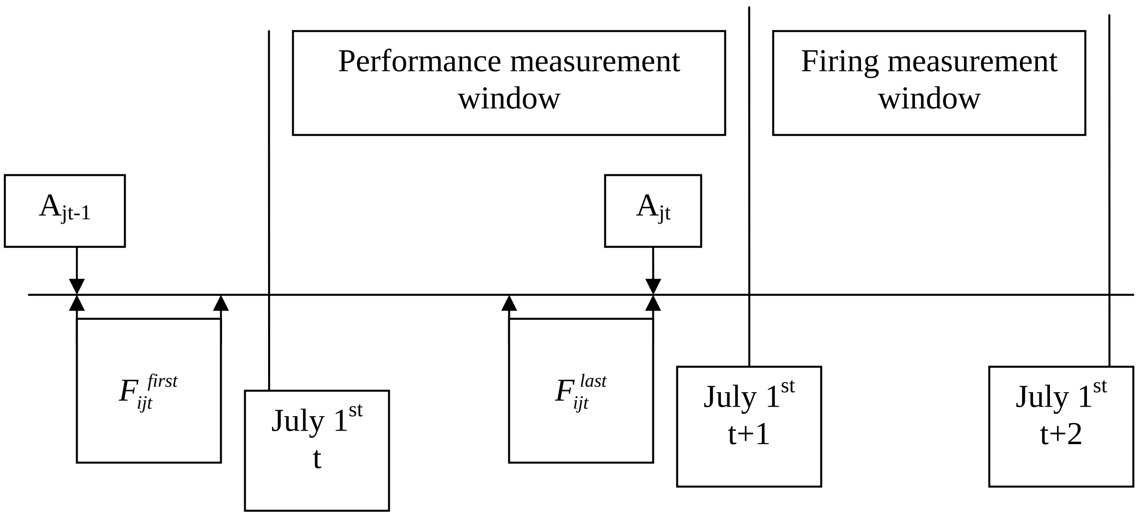
Schipper, K. 1991. Commentary on analysts’ forecasts. Accounting Horizons 5 (4), 105-121.

Smith, R., L. Cauley. 1999. Will upgrade of AT&T stock benefit Salomon? The Wall Street Journal, December 6, 1999, C1.

Solomon, D., Frank R., 2003. SEC takes aim at companies that retaliate against analysts. The Wall Street Journal, June 19, 2003.

Stickel, S. 1992. Reputation and performance among security analysts. Journal of Finance, 1811-1836.

Figure 1. Timeline for the construction of regression variables



Variable definitions:

A_{jt} denotes firm j's annual earnings for year t that is announced immediately before July 1, year t+1;

A_{jt-1} denotes firm j's annual earnings for year t-1;

F_{ijt}^{last} is analyst i's latest forecast of annual earnings A_{jt}, issued in the second half of the period from the earnings announcement date of A_{jt-1} to the earnings announcement date of A_{jt}; and

F_{ijt}^{first} is analyst i's earliest forecast of annual earnings A_{jt} issued in the first half of the period from the earnings announcement date of A_{jt-1} to the earnings announcement date of A_{jt}.

Table 1. Descriptive Statistics over January 1, 1983-July 1, 2000

Panel A. Variables used in model (3) for annual earnings forecasts[a]

variable	*N*	*Mean*	*25%*	*median*	*75%*	*S.D.*
OP_{ijt}	228,904	0.260	0	0	1	0.438
OO_{ijt}	228,904	0.343	0	0	1	0.475
PP_{ijt}	228,904	0.306	0	0	1	0.461
PO_{ijt}	228,904	0.091	0	0	0	0.287
$Accuracy_{ijt}$	228,904	50.00	23.53	50.00	76.19	31.67
$Bold_{ijt}$	228,904	50.00	21.43	50.00	77.78	32.44
$FirmExperience_{ijt}$	228,904	4.30	2	3	6	3.15
$R\_FirmExperience_{ijt}$	228,904	50.00	22.73	50.00	76.67	31.43
$Follow_{ijt}$	228,904	21.07	11	19	29	12.62
$FirmsCovered_{ijt}$	228,904	25.29	14	20	29	22.56
$R\_FirmsCovered_{ijt}$	228,904	50.00	21.15	50.00	78.57	33.13
GAP_{ijt}	228,904	78.89	43	81	104	43.81
R_GAP_{ijt}	228,904	50.00	21.43	50.00	78.57	32.97

Panel B. Variables used in model (3) for quarterly earnings forecasts[b]

variable	*N*	*Mean*	*25%*	*median*	*75%*	*S.D.*
OP_{ijt}	114,075	0.291	0	0	1	0.454
OO_{ijt}	114,075	0.298	0	0	1	0.458
PP_{ijt}	114,075	0.365	0	0	1	0.481
PO_{ijt}	114,075	0.045	0	0	0	0.208
$Accuracy_{ijt}$	114,075	50.00	25.00	50.00	75.00	33.03
$Bold_{ijt}$	114,075	50.00	21.42	50.00	80.00	34.33
$FirmExperience_{ijt}$	114,075	4.89	2	4	7	3.71
$R\_FirmExperience_{ijt}$	114,075	50.00	21.42	50.00	78.57	33.51
$Follow_{ijt}$	114,075	23.96	15	22	32	11.91
$FirmsCovered_{ijt}$	114,075	20.15	13	18	24	11.98
$R\_FirmsCovered_{ijt}$	114,075	50.00	21.00	50.00	80.00	35.39
GAP_{ijt}	114,075	48.67	23	46	76	28.48
R_GAP_{ijt}	114,075	50.00	20.00	50.00	80.00	34.87

Panel C. Variables used in model (4) for annual earnings forecasts[c]

variable	*N*	*Mean*	*25%*	*median*	*75%*	*S.D.*
$Fire_{i,t+1}$	32,303	0.15	0	0	0	0.36
$OP_{i,t}$	32,303	0.25	0.00	0.22	0.38	0.25
$OO_{i,t}$	32,303	0.35	0.13	0.33	0.50	0.29
$PP_{i,t}$	32,303	0.30	0.00	0.25	0.50	0.28
$PO_{i,t}$	32,303	0.09	0.00	0.00	0.13	0.16
$Accuracy_{i,t}$	32,303	49.85	41.33	50.00	58.77	14.70
$Bold_{i,t}$	32,303	50.32	42.09	50.00	58.18	14.18
$Experience_{i,t}$	32,303	5.01	2	4	7	3.76

Panel D: Variables used in model (4) for quarterly earnings forecasts[d]

variable	*N*	*Mean*	*25%*	*median*	*75%*	*S.D.*
$Fire_{i,t+1}$	15,278	0.12	0	0	0	0.32
$OP_{i,t}$	15,278	0.30	0.00	0.25	0.50	0.29
$OO_{i,t}$	15,278	0.32	0.00	0.25	0.50	0.31
$PP_{i,t}$	15,278	0.34	0.00	0.33	0.50	0.31
$PO_{i,t}$	15,278	0.04	0.00	0.00	0.00	0.13
$Accuracy_{i,t}$	15,278	49.65	37.50	50.00	62.50	21.67
$Bold_{i,t}$	15,278	50.28	37.50	50.00	62.50	22.17
$Experience_{i,t}$	15,278	6.22	3	5	9	4.10

[a] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). OP_{ijt} is optimism-to-pessimism forecast bias for analyst i who follows firm j in year t. OO_{ijt} is optimism-to-optimism forecast bias for analyst i who follows firm j in year t. PP_{ijt} is pessimism-to-pessimism forecast bias for analyst i who follows firm j in year t. PO_{ijt} is pessimism-to-optimism forecast bias for analyst i who follows firm j in year t. The four forecast biases are defined using each analyst's first and last annual earnings forecasts over two consecutive annual earnings announcement dates. $Accuracy_{ijt}$ is the standardized earnings forecast accuracy ranking (based on the last earnings forecast) of analyst i relative to other analysts who follow the same firm j in year t. $Bold_{ijt}$ is the standardized ranking of the deviation of analyst i's first annual earnings forecast relative to other analysts' forecasts for the same firm j in year t. $FirmExperience_{ijt}$ is the number of years analyst i follows stock j as of year t. $Follow_{ijt}$ is the total number of analysts (including analyst i) who follow firm j in year t. $FirmsCovered_{ijt}$ is the number of firms (including firm j) followed by analyst i in year t. GAP_{ijt} is the distance in days between the earnings announcement date for A_{jt} and the forecast date for F_{ijt}^{last} for

analyst i in year t. $R\_FirmExperience_{ijt}$, $R\_FirmsCovered_{ijt}$, and R_GAP_{ijt} are the standardized ranking of $FirmExperience_{ijt}$, $FirmsCovered_{ijt}$, and GAP_{ijt} , respectively.

[b] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to any of the quarters that fall within year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). OP_{ijt} is optimism-to-pessimism forecast bias for analyst i who follows firm j in quarter t. OO_{ijt} is optimism-to-optimism forecast bias for analyst i who follows firm j in quarter t. PP_{ijt} is pessimism-to-pessimism forecast bias for analyst i who follows firm j in quarter t. PO_{ijt} is pessimism-to-optimism forecast bias for analyst i who follows firm j in quarter t. The four forecast biases are defined using each analyst's first and last quarterly earnings forecasts issued between the quarterly earnings announcement two quarters prior and the current quarter's earnings announcement. The other variables in Panel B are defined in the same way as the annual definitions in Panel A, using quarterly earnings forecasts.

[c] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $Fire_{i,t+1}$ is equal to one if analyst i is demoted from a large brokerage firm to a small brokerage firm or permanently leaves the profession during the year from July 1, t+1 to June 30, t+2, and zero otherwise. $Experience_{i,t}$ is the number of years analyst i appears in the IBES annual earnings forecast database as of year t. The other variables in Panel C are the average of the respective variables in Panel A across all stocks j followed by analyst i in year t.

[d] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to any of the quarters that fall within year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $Experience_{i,t}$ is defined in Panel C above. The other variables in Panel D are the average of the same variables in Panel B across all stocks j followed by analyst i in year t and are defined as the mean of each quarterly variable across all quarters in year t for each firm-analyst, followed by the averaging of the mean quarterly variable across all firms j followed by analyst i in year t.

Table 2. Correlations for Key Regression Variables over January 1, 1983-July 1, 2000[a]

	$OP_{i,t}^{A}$	$OO_{i,t}^{A}$	$PP_{i,t}^{A}$	$PO_{i,t}^{A}$	$OP_{i,t}^{Q}$	$OO_{i,t}^{Q}$	$PP_{i,t}^{Q}$	$PO_{i,t}^{Q}$	$Accuracy_{i,t}^{A}$	$Accuracy_{i,t}^{Q}$	$Fire_{i,t+1}$
$OP_{i,t}^{A}$		-0.292***	-0.293***	-0.150***	0.257***	-0.011	-0.189***	-0.090***	0.118***	0.030***	-0.107***
$OO_{i,t}^{A}$	-0.378***		-0.548***	-0.125***	-0.044***	0.352***	-0.296***	-0.032***	-0.062***	-0.020**	0.033***
$PP_{i,t}^{A}$	-0.372***	-0.570***		-0.033***	-0.122***	-0.290***	0.386***	0.039***	-0.033***	-0.017**	-0.050***
$PO_{i,t}^{A}$	-0.219***	-0.227***	-0.138***		-0.131***	-0.065***	0.126***	0.140***	-0.014**	0.010	0.016***
$OP_{i,t}^{Q}$	0.297***	-0.049***	-0.112***	-0.130***		-0.419***	-0.440***	-0.170***	0.021**	0.180***	-0.029***
$OO_{i,t}^{Q}$	0.030***	0.350***	-0.295***	-0.037***	-0.336***		-0.547***	-0.154***	-0.001	-0.147***	0.009
$PP_{i,t}^{Q}$	-0.177***	-0.316***	0.402***	0.125***	-0.359***	-0.504***		-0.105***	-0.022**	-0.018**	0.012
$PO_{i,t}^{Q}$	-0.076***	-0.055***	0.071***	0.191***	-0.102***	-0.090***	0.030***		0.006	-0.006	0.016*
$Accuracy_{i,t}^{A}$	0.123***	-0.038***	-0.022***	-0.014**	0.022**	0.001	-0.027**	0.003		0.232***	-0.147***
$Accuracy_{i,t}^{Q}$	0.032***	-0.020**	-0.012	0.011	0.170***	-0.137***	-0.017**	0.007	0.199***		-0.036***
$Fire_{i,t+1}$	-0.064***	0.062***	-0.014*	0.013**	-0.044***	-0.009	-0.001	0.002	-0.137***	-0.037***	

[a] $Accuracy_{i,t}^{A}$ is $Accuracy_{i,t}$ using annual earnings forecasts, while $Accuracy_{i,t}^{Q}$ is $Accuracy_{i,t}$ using quarterly earnings forecasts. See Table 1 for other variable definitions. Spearman correlations are reported in the top diagonal and Pearson correlations are reported in the bottom diagonal. The sample size for the correlations among the annual earnings forecast variables is 32,303; the sample size for the correlations among the quarterly earnings forecast variables is 15,278; the sample size for the correlations across annual and quarterly earnings forecast variables is 14,511. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

Table 3. OLS Regression Results of Analyst Forecast Accuracy (H1)

Panel A. Regression results using annual earnings forecasts [a]

Dependent variable = $Accuracy_{ijt}$	(1)	(2)	(3)
		Coefficient (standard error)	
OP_{ijt}	5.059	6.530	6.079
	(0.296)***	(0.162)***	(0.235)***
OO_{ijt}	-3.106		
	(0.255)***		
PP_{ijt}	-0.105		
	(0.308)		
$Bold_{ijt}$	-0.018	-0.017	-0.017
	(0.002)***	(0.002)***	(0.002)***
$R\_FirmExperience_{ijt}$	0.013	0.013	0.013
	(0.003)***	(0.003)***	(0.003)***
$\ln(Follow_{ijt})$	-0.054	0.058	0.018
	(0.049)	(0.049)	(0.050)
$R\_FirmsCovered_{ijt}$	-0.006	-0.006	-0.006
	(0.003)**	(0.003)**	(0.003)**
R_GAP_{ijt}	-0.108	-0.111	-0.108
	(0.003)***	(0.003)***	(0.003)***
$InsiderSell_{ijt}$			-0.185
			(0.098)*
$OP_{ijt} \times InsiderSell_{ijt}$			0.781
			(0.332)**
$Dispersion_{ijt}$			-0.646
			(0.115)***
$OP_{ijt} \times Dispersion_{ijt}$			0.736
			(0.362)**
Brokerage firm fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
N	228,904	228,904	220,734
R^2	0.038	0.037	0.036

Panel B. Regression results using quarterly earnings forecasts [b]

Dependent variable = $Accuracy_{ijt}$	(1)	(2)	(3)
		Coefficient (standard error)	
OP_{ijt}	8.533	10.740	10.252
	(0.594)***	(0.224)***	(0.291)***
OO_{ijt}	-5.464		
	(0.519)***		
PP_{ijt}	0.125		
	(0.573)		
$Bold_{ijt}$	-0.015	-0.014	-0.014
	(0.003)***	(0.003)***	(0.003)***
$R\_FirmExperience_{ijt}$	0.008	0.008	0.008
	(0.004)**	(0.004)**	(0.004)**
$\ln(Follow_{ijt})$	-0.006	0.009	0.005
	(0.004)	(0.005)**	(0.005)
$R\_FirmsCovered_{ijt}$	-0.009	-0.009	-0.009
	(0.003)**	(0.004)**	(0.004)**
R_GAP_{ijt}	-0.102	-0.105	-0.105
	(0.004)***	(0.004)***	(0.004)***
$InsiderSell_{ijt}$			-0.231
			(0.167)
$OP_{ijt} \times InsiderSell_{ijt}$			1.153
			(0.468)**
$Dispersion_{ijt}$			-1.044
			(0.173)***
$OP_{ijt} \times Dispersion_{ijt}$			0.511
			(0.470)
Brokerage firm fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
N	114,075	114,075	113,000
R^2	0.049	0.044	0.044

[a] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $InsiderSell_{ijt}$ is a dummy that is equal to 1 if the average net

insider selling (expressed in 1982 dollars) by all corporate officers and directors for firm j followed by analyst i during the calendar year immediately before the earnings announcement date for $Accuracy_{ijt}$ is larger than the 75th percentile of our sample. $Dispersion_{ijt}$ is a dummy that is equal to 1 if the forecast dispersion (defined as the standard deviation of the earnings forecasts scaled by the magnitude of the realized earnings) is greater than the 75th percentile of our sample. $Dispersion_{ijt}$ is computed using each analyst's first earnings forecast F_{ijt}^{first} , although results are similar if each analyst's last earnings forecast F_{ijt}^{last} is used instead. See Table 1 for other variable definitions. The standard errors are computed using Rogers' (1993) method, which allows heteroskedasticity and any type of correlation for observations of the same stocks but assumes independence for observations of different stocks. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

[b] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to any of the quarters that fall within year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $InsiderSell_{ijt}$ and $Dispersion_{ijt}$ are defined similarly to Panel A above. See Table 1 for other variable definitions. The standard errors are computed using Rogers' (1993) method, which allows heteroskedasticity and any type of correlation for observations of the same stocks but assumes independence for observations of different stocks. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

Table 4. Logit Regression Results of Analyst Firing (H2)

Panel A. Regression results using annual earnings forecasts [a]

Dependent variable = $Fire_{i,t+1}$	(1)	(2)	(3)
		Coefficient (standard error)	
$Accuracy_{i,t}$	-0.027	-0.028	-0.028
	(0.001)***	(0.001)***	(0.001)***
$OP_{i,t}$	-0.371	-0.366	0.001
	(0.117)***	(0.069)***	(0.158)
$OO_{i,t}$	0.142		
	(0.109)		
$PP_{i,t}$	-0.158		
	(0.126)		
$Bold_{i,t}$	-0.002	-0.002	-0.002
	(0.001)	(0.001)	(0.001)*
$\ln(Experience_{i,t})$	-0.223	-0.222	-0.214
	(0.036)***	(0.036)***	(0.035)***
$InsiderSell_{i,t}$			-0.280
			(0.109)**
$OP_{i,t} \times InsiderSell_{i,t}$			-0.620
			(0.271)**
$Dispersion_{i,t}$			0.404
			(0.093)***
$OP_{i,t} \times Dispersion_{i,t}$			-0.617
			(0.284)**
Brokerage firm fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
N	32,303	32,303	30,650

Panel B. Regression results using quarterly earnings forecasts[b]

Dependent variable = $Fire_{i,t+1}$	(1)	(2)	(3)
		Coefficient (standard error)	
$Accuracy_{i,t}$	-0.004	-0.004	-0.004
	(0.001)***	(0.001)***	(0.001)***
$OP_{i,t}$	-0.527	-0.297	-0.110
	(0.184)***	(0.105)***	(0.154)
$OO_{i,t}$	-0.191		
	(0.173)		
$PP_{i,t}$	-0.292		
	(0.173)*		
$Bold_{i,t}$	0.002	0.002	0.001
	(0.001)	(0.001)	(0.001)
$\ln(Experience_{i,t})$	-0.081	-0.081	-0.073
	(0.038)**	(0.038)**	(0.037)*
$InsiderSell_{i,t}$			0.109
			(0.115)
$OP_{i,t} \times InsiderSell_{i,t}$			-0.738
			(0.392)*
$Dispersion_{i,t}$			-0.049
			(0.116)
$OP_{i,t} \times Dispersion_{i,t}$			-0.011
			(0.263)
Brokerage firm fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
N	15,278	15,278	14,942

Panel C. Regression results using both annual and quarterly earnings forecasts [c]

Dependent variable = $Fire_{i,t+1}$	(1)
	Coefficient (standard error)
$Accuracy_{i,t}^{A}$	-0.036
	(0.003)***
$Accuracy_{i,t}^{Q}$	-0.002
	(0.002)
$OP_{i,t}^{A}$	-0.308
	(0.140)**
$OP_{i,t}^{Q}$	-0.197
	(0.114)*
$Bold_{i,t}^{A}$	-0.003
	(0.002)
$Bold_{i,t}^{Q}$	0.003
	(0.001)**
$\ln(Experience_{i,t})$	-0.121
	(0.041)***
Brokerage firm fixed effects	Yes
Year fixed effects	Yes
N	14,511

[a] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $InsiderSell_{i,t}$ is the average of $InsiderSell_{ijt}$ over all the firms j covered by analyst i in year t. $Dispersion_{i,t}$ is the average of $Dispersion_{ijt}$ over all the firms j covered by analyst i in year t. See Tables 1 and 3 for other variable definitions. The standard errors are computed using Rogers' (1993) method, which allows heteroskedasticity and any type of correlation for observations of the same brokerage houses but assumes independence for observations of different brokerage houses. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

[b] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to any of the quarters that fall within year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $InsiderSell_{i,t}$ and $Dispersion_{i,t}$

are defined as the mean of the same quarterly variable across all quarters in year t for each firm-analyst, followed by the averaging of the mean quarterly variable across all firms j followed by analyst i in year t. See Tables 1 and 3 for other variable definitions. The standard errors are computed using Rogers' (1993) method, which allows heteroskedasticity and any type of correlation for observations of the same brokerage houses but assumes independence for observations of different brokerage houses. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

[c] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $Bold_{i,t}^{A}$ and $Bold_{i,t}^{Q}$ are $Bold_{i,t}$ for annual earnings forecasts and quarterly earnings forecasts, respectively. $OP_{i,t}^{A}$ and $OP_{i,t}^{Q}$ are $OP_{i,t}$ for annual earnings forecasts and quarterly earnings forecasts, respectively. See Tables 1, 2, and 3 for other variable definitions. The standard errors are computed using Rogers' (1993) method, which allows heteroskedasticity and any type of correlation for observations of the same brokerage houses but assumes independence for observations of different brokerage houses. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

Table 5: Heckman Regression Results of Future Earnings Forecast Accuracy

Panel A. Regression results using annual earnings forecasts [a]

Dependent variable = $Accuracy_{i,t+1}$	(1)
	Coefficient (standard error)
$OP_{i,t}$	0.968
	(0.344)***
$Accuracy_{i,t}$	0.068
	(0.008)***
$Bold_{i,t+1}$	-0.036
	(0.009)***
$R\_FirmExperience_{i,t+1}$	-0.001
	(0.005)
$\ln(Follow_{i,t+1})$	0.023
	(0.013)*
$R\_FirmsCovered_{i,t+1}$	-0.001
	(0.003)
$R\_GAP_{i,t+1}$	-0.178
	(0.008)***
Brokerage firm fixed effects	Yes
Year fixed effects	Yes
N	23,289

Panel B. Regression results using quarterly earnings forecasts [b]

Dependent variable = $Accuracy_{i,t+1}$	(1)
	Coefficient (standard error)
$OP_{i,t}$	1.330
	(0.881)
$Accuracy_{i,t}$	0.043
	(0.011)***
$Bold_{i,t+1}$	-0.007
	(0.013)
$R\_FirmExperience_{i,t+1}$	-0.002
	(0.008)
$\ln(Follow_{i,t+1})$	0.032
	(0.026)
$R\_FirmsCovered_{i,t+1}$	-0.002
	(0.007)
$R\_GAP_{i,t+1}$	-0.115
	(0.013)***
Brokerage firm fixed effects	Yes
Year fixed effects	Yes
N	9,737

[a] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $Follow_{i,t+1}$ is the average of $Follow_{ijt+1}$ across all firms j covered by analyst i in year t. $R\_FirmExperience_{i,t}$ and $R\_GAP_{i,t}$ are the averages of $R\_FirmExperience_{ijt}$ and R_GAP_{ijt}, respectively, across all firms j covered by analyst i in year t. See Table 1 for other variable definitions. The standard errors are computed using Rogers' (1993) method, which allows heteroskedasticity and any type of correlation for observations of the same brokerage houses but assumes independence for observations of different brokerage houses. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

[b] The subscript i refers to analyst i; the subscript j refers to stock j; and the subscript t refers to any of the quarters that fall within year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). All the variables in Panel B are the

mean of their quarterly equivalents across all firms j covered by analyst i in year t and are defined as the mean of each quarterly variable across all quarters in year t for each firm-analyst, followed by the averaging of the mean quarterly variable across all firms j followed by analyst i in year t. The standard errors are computed using Rogers' (1993) method, which allows heteroskedasticity and any type of correlation for observations of the same brokerage houses but assumes independence for observations of different brokerage houses. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

Table 6: Characteristics of Analysts Who Issue annual OP Forecasts

Panel A. Descriptive statistics (N=32,303)[a]

Mean (median)[standard Deviation]

Variable	OP>median	OP≤median	P Value from a Ranksum Test of the Difference
$FirmExperience_{i,t}$	3.320 (2.750) [2.082]	2.975 (2.416) [1.949]	<0.001
$Broker\ size_{i,t}$	43.570 (32.000) [42.857]	41.011 (28.000) [41.498]	<0.001
$AllStar_{i,t}$	0.133 (0.000) [0.340]	0.104 (0.000) [0.306]	<0.001

Panel B. Regression of OP on analyst characteristics[b]

Dependent variable = $OP_{i,t}$ *100	(1)
	Coefficient (standard error)
$FirmExperience_{i,t}$	0.225
	(0.087)***
$Broker\ size_{i,t}$	0.013
	(0.003)***
$AllStar_{i,t}$	1.865
	(0.400)***
Year fixed effects	Yes
N	32,303

[a] The subscript i refers to analyst i; and the subscript t refers to year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). $Bro\ker size_{i,t}$ is the number of unique analysts that belong to brokerage firm i in year t. $AllStar_{i,t}$ is coded one if an analyst is an All-Star as determined by the Institutional Investor magazine in year t-1, and zero otherwise. The All-Star data are available for only 1995-2000. See Panel C of Table 1 for other variable definitions.

[b] The subscript i refers to analyst i; and the subscript t refers to year t, defined as the period from July 1, t to July 1, t+1 (see Figure 1). See Panel A above for other variable definitions. The standard errors are computed using Rogers' (1993) method, which allows heteroskedasticity and any type of correlation for observations of the same brokerage houses but assumes independence for observations of different brokerage houses. *, **, *** denote two-tailed significance levels of 10%, 5%, and 1%, respectively.

An Evaluation of Financial Analysts and Naive Methods in Forecasting Long-Term Earnings

Michael Lacina

University of Houston-Clear Lake - School of Business

BuRyung Brian Lee

affiliation not provided to SSRN

Zhaohui Randall Xu

University of Houston, Clear Lake

Advances in Business and Management Forecasting (Vol. 8), Kenneth D. Lawrence, Ronald K. Klimberg (ed.), Emerald Group Publishing Limited, pp.77-101

Abstract:
We evaluate the performance of financial analysts versus naïve models in making long-term earnings forecasts. Long-term earnings forecasts are generally defined as third-, fourth-, and fifth- year earnings forecasts. We find that for the fourth and fifth years, analysts' forecasts are no more accurate than naïve random walk (RW) forecasts or naïve RW with economic growth forecasts. Furthermore, naïve model forecasts contain a large amount of incremental information over analysts' long-term forecasts in explaining future actual earnings. Tests based on subsamples show that the performance of analysts' long-term forecasts declines relative to naïve model forecasts for firms with high past earnings growth and low analyst coverage. Furthermore, a model that combines a naïve benchmark (last year's earnings) with the analyst long-term earnings growth forecast does not perform better than analysts' forecasts or naïve model forecasts. Our findings suggest that analysts' long-term earnings forecasts should be used with caution by researchers and practitioners. Also, when analysts' earnings forecasts are unavailable, naïve model earnings forecasts may be sufficient for measuring long-term earnings expectations.

AN EVALUATION OF FINANCIAL ANALYSTS AND NAÏVE METHODS IN FORECASTING LONG-TERM EARNINGS

Michael Lacina, B. Brian Lee and Randall Zhaohui Xu

ABSTRACT

We evaluate the performance of financial analysts versus naïve models in making long-term earnings forecasts. Long-term earnings forecasts are generally defined as third-, fourth-, and fifth-year earnings forecasts. We find that for the fourth and fifth years, analysts' forecasts are no more accurate than naïve random walk (RW) forecasts or naïve RW with economic growth forecasts. Furthermore, naïve model forecasts contain a large amount of incremental information over analysts' long-term forecasts in explaining future actual earnings. Tests based on subsamples show that the performance of analysts' long-term forecasts declines relative to naïve model forecasts for firms with high past earnings growth and low analyst coverage. Furthermore, a model that combines a naïve benchmark (last year's earnings) with the analyst long-term earnings growth forecast does not perform better than analysts' forecasts or naïve model forecasts. Our findings suggest that analysts' long-term earnings

Advances in Business and Management Forecasting, Volume 8, 77–101

ISSN: 1477-4070/doi:10.1108/S1477-4070(2011)0000008009

forecasts should be used with caution by researchers and practitioners. Also, when analysts' earnings forecasts are unavailable, naïve model earnings forecasts may be sufficient for measuring long-term earnings expectations.

INTRODUCTION

This chapter evaluates the performance of financial analysts versus naïve models in forecasting long-term earnings. Analysts' earnings forecasts are widely used in accounting research as proxy for market expected earnings (Ramnath, Rock, & Shane, 2008; Schipper, 1991). The underlying assumption is that in an informationally efficient market, the capital market should use the best future earnings data available, where the best is defined as the most accurate (Brown, 1993). Indeed, many researchers in recent years have assumed that analysts' forecasts are superior to those of naïve and time series models.[1] However, prior evidence on the superiority of analysts' earnings forecasts over statistical model forecasts mainly originates from studies that focus on a comparison of predictive accuracy for short-term earnings forecasts, typically for the upcoming quarters or the coming year (e.g., Brown, Griffin, Hagerman, & Zmijewski, 1987a, 1987b; Brown, Richardson, & Schwager, 1987; Brown & Rozeff, 1978; Fried & Givoly, 1982; Imhoff & Pare, 1982).

Analysts tend to have a timing advantage over naïve and time series models in predicting short-term earnings due to the information available between the end of the final time period included in the forecast model and the date the analyst makes a forecast. Analysts do not have as much of a timing advantage over naïve and time series methods in making earnings forecasts over longer horizons, which normally extend more than two years from the forecast date. Furthermore, analysts are often evaluated on the accuracy of their short-term forecasts but not of their long-term forecasts (Dechow, Hutton, & Sloan, 2000; Stickel, 1992). This would on average provide analysts with more of an incentive to be accurate in their short-term forecasts than in their long-term forecasts. In fact, Chan, Karceski, and Lakonishok (2003) find that analysts' long-term earnings growth forecasts are overly optimistic and have little predictive power. The questionable predictive ability of analysts' long-term growth forecasts puts doubt on the assumption that analysts' forecasts are the default proxy for market expectations of long-term earnings extending beyond two years. Nevertheless, long-term earnings growth forecasts are widely disseminated by financial analysts. Bradshaw (2004) finds that analysts use their long-term earnings growth forecasts in formulating stock recommendations. Moreover, prior studies plug in up to five years of analysts' earnings forecasts into earnings-based valuation models to infer the implied cost of capital (e.g., Botosan & Plumlee, 2005; Claus & Thomas, 2001; P. Easton, Taylor, Shroff, & Sougiannis, 2002) or assess firms' intrinsic values (e.g., Frankel & Lee, 1998; Sougiannis & Yaekura, 2001).

When earnings forecasts serve as inputs to valuation models, the accuracy of the earnings forecasts directly affects the estimates of cost of capital and intrinsic values. For example, P. Easton and Sommers (2007) find that optimism in analysts' earnings forecasts leads to an upward bias in the estimated cost of capital of about 3%. P. Easton and Monahan (2005) show that cost of capital derived from analysts' earnings forecasts is negatively correlated with realized returns after controlling for proxies for cash flow news and discount rate news. Similarly, prior studies (e.g., Francis, Olsson, & Oswald, 2000; Sougiannis & Yaekura, 2001) find large valuation errors from valuation models that use analysts' forecasts as a proxy for future earnings. Evidence in P. Easton and Monahan (2005) and Sougiannis and Yaekura (2001) suggests that their aforementioned findings are partially due to problems with analyst earnings forecast quality. Therefore, it is important to examine the performance of analysts' forecasts against alternative sources of earnings forecasts such as statistical models. The findings will provide fresh insight into the appropriateness of using analysts' forecasts as the default proxy for expected earnings in academic research.

A number of studies that examine the performance of analysts' long-term earnings forecasts use samples selected based on a transaction that has taken place, which limits the generalizability of their findings.[2] There are exceptions, that is, Cragg and Malkiel (1968) and Rozeff (1983). Cragg and Malkiel (1968) find that analysts' long-term earnings growth forecasts are on the whole no more accurate than naïve forecasts based on past earnings growth. They use analysts' forecasts made in 1962 and 1963 by five brokerage houses for 185 firms. On the contrary, Rozeff (1983) finds that growth rates derived from four- to five-year earnings forecasts from *Value Line* are more accurate than the corresponding growth rates implicit in four expected stock return models. His study uses a sample that includes *Value Line* long-term earnings forecasts made in 1967 (253 firms) and 1972 (348 firms). Given the poor performance of analysts' long-term earnings growth forecasts found in Chan et al.(2003) and the small samples from the 1960s and early 1970s used in Cragg and Malkiel (1968) and Rozeff (1983), it is important to reexamine the performance of analysts' long-term earnings forecasts versus those of naïve models.

We use *I/B/E/S* analyst forecast data to compare analysts' long-term earnings forecasts with those of two naïve models. Whereas the analysts' first year (end of year following last reported annual earnings) and second year earnings forecasts are normally considered short-term forecasts, the third year through fifth-year forecasts are generally considered long term. Analysts' long-term earnings forecasts are either obtained directly on *I/B/E/S* or derived using the analysts' last available explicit earnings forecast with the analysts' long-term earnings growth rate, as is often done in the literature.[3] The two naïve earnings forecast models are a random walk (RW) model and a RW with a drift based on historical inflation and historical real GDP growth (RWGDP).[4] Additionally, some researchers have found that combining analysts' forecasts with naïve benchmarks can improve forecast accuracy (e.g., Cheng, Fan, & So, 2003; Conroy & Harris, 1987; Newbold, Zumwalt, & Kannan, 1987). Therefore, we also examine whether a hybrid model (RWLTG) combining a naïve benchmark, last year's earnings, with the analysts' long-term earnings growth rate forecast can improve long-term earnings forecast accuracy. The performances of the analyst, naïve, and hybrid forecasts are evaluated by examining their accuracy and information content.

The results for short-term forecast horizons show that analysts' earnings forecasts are more accurate than RW and RWGDP forecasts, which is consistent with prior research. However, as the forecast horizon extends beyond the second year, the higher accuracy of analysts' forecasts wanes such that for long-term horizons (especially fourth and fifth years), we cannot conclude whether analysts' forecasts are more accurate than RW or RWGDP forecasts. In some cases, we find evidence that the RWGDP model is more accurate than analysts' forecasts. As far as information content is concerned, a regression analysis shows that analysts' forecasts provide the majority of the information in explaining first- and second-year actual earnings. However, naïve model forecasts provide substantial incremental information over analysts' forecasts in explaining future actual earnings as the forecast horizon is extended beyond the second year.

We perform additional tests of accuracy and information content. First, we run the analyses on sample partitions. The results of these tests show that the performance of analysts' earnings forecasts declines relative to naïve model forecasts for firms with high past earnings growth and low analyst following. Also, when analysts issue explicit (as opposed to growth rate) long-term earnings forecasts, the performance of their forecasts improves relative to naïve model forecasts for only the fifth year in the forecast horizon. However, financial analysts infrequently issue explicit earnings forecasts for the fifth year. Second, we compare earnings forecasts of the hybrid RWLTG model with analysts' forecasts and RWGDP forecasts (the most accurate naïve forecast). We find that the hybrid RWLTG model does not enhance forecast accuracy. Furthermore, the hybrid model forecasts contain less information content in explaining future earnings than RWGDP model forecasts or analysts' forecasts.

Our results convey that academics and practitioners should use analysts' long-term earnings forecasts with caution, especially for firms with high earnings growth. These analyst long-term forecasts appear to be no more accurate than some of the simple, naïve forecasts. Also, much of the information useful in explaining long-term future actual earnings is provided by naïve forecasts as opposed to analysts' forecasts. Our findings imply that the use of naïve forecast models such as RWGDP and RW may be sufficient and easily derived ways of forecasting long-term earnings when analysts' forecasts are unavailable. It is well known that analyst coverage is affected by various factors, and analysts tend to cover firms that are large and profitable (Bhushan, 1989; Hong, Lim, & Stein, 2000). Therefore, using forecasts from naïve models enables researchers to expand the sample to include firms without analyst coverage, thereby reducing the potential sampling bias in research design that limits the generalizability of their findings. This study contributes to the burgeoning stream of research that uses alternative earnings forecasts as a proxy for expected earnings. For example, Allee (2009) and Hou, van Dijk, and Zhang (2010) use earnings forecasts derived from time series models and a cross-sectional model, respectively, to estimate cost of capital.

The chapter proceeds as follows. The second section reviews relevant literature. In the third section, we explain the chapter's methodology. The fourth section discusses the results, including those for the full sample, sample partitions, and the hybrid model. The fifth section contains the conclusions.

LITERATURE REVIEW

Much of the literature that compares analysts' earnings forecasts with naïve or time series forecasts focuses on short-term forecasts. Brown and Rozeff (1978) examine quarterly earnings forecasts ranging from one quarter to five quarters ahead and first (current)-year annual earnings forecasts. They find that *Value Line* analysts' forecasts, on the whole, are more accurate than time series forecasts. Imhoff and Pare (1982) show that analysts' forecasts on average outperform time series forecasts in terms of accuracy when the forecast horizon is four quarters ahead but not when it is three quarters ahead. Fried and Givoly (1982) examine first-year annual earnings forecasts and find that analysts' forecasts are more accurate than forecasts from two time series models. Brown et al. (1987) test analysts' one, two, and three-quarter-ahead forecasts from *Value Line* made one, two, and three months before the end of a quarter and analysts' first- and second-year annual forecasts from *I/B/E/S*. Their findings support the superiority of analysts' forecasts over time series forecasts. Cheng et al. (2003) use *I/B/E/S* analysts' first-year annual forecasts from Hong Kong. For the first 10 months following the previous earnings announcement, both analysts and RW forecasts have information content in explaining actual earnings. However, analysts' forecasts have relatively more information content as the earnings announcement date approaches. Brown et al. (1987a) test quarterly forecasts from one to three quarters ahead and find that the predictive accuracy of analysts' forecasts is superior to that of time series forecasts. They attribute this analyst superiority to two factors: (1) a contemporaneous advantage due to an analyst's ability to make better use of current information and (2) a timing advantage stemming from the acquisition of information by an analyst between the date the naïve forecast is made and the date the analyst forecast is made. However, although timing can be a major advantage for analysts relative to naïve methods for short-term forecasts, this advantage is less likely to have a significant impact on long-term forecasts.

Research that directly examines the performance of analysts' long-term forecasts has been sparse. Cragg and Malkiel (1968) study the accuracy of analysts' five-year earnings growth forecasts from five brokerage houses. They find that analysts' five-year earnings growth forecasts are no more accurate than long-term earnings growth forecasts based on past earnings growth rates or price-to-earnings ratios. On the contrary, analysts' five-year growth forecasts are found to be more accurate than naïve forecasts of no earnings growth. Rozeff (1983) uses four-to-five year earnings growth rates from *Value Line* analysts during 1967 and 1972. These forecasts are found to predict long-term earnings growth better than naïve forecasts from four expected return models. Chan et al. (2003) analyze the growth rates of earnings and sales. They document that analysts' long-term earnings growth forecasts are overly optimistic and have little predictive power for future earnings. A defect of these forecasts is that analysts predict sustained earnings growth rates over a long future time horizon (e.g., three to five years) for a large proportion of firms. On the contrary, the authors show that only 12.2% (2.6%) of their sample firms achieve above median growth in income before extraordinary items for three (five) straight years. Dechow et al. (2000) study analysts' long-term earnings growth forecasts made around the equity offerings and find that the forecasts are systematically optimistic. Bradshaw (2004) documents that analysts use their long-term earnings growth forecasts in generating stock recommendations but that their long-term earnings growth forecasts are *negatively* related to future returns.

METHODOLOGY

Sample Selection

Our sample is from the *I/B/E/S* database. For the month of June for each year from 1988 to 2003, we obtain the median consensus analysts' earnings forecasts for up to five years ahead and the median consensus analysts' forecasted long-term earnings growth rate.[5] *I/B/E/S* recommends the usage of the median (as opposed to mean) long-term earnings growth rate forecast to prevent excessive influence from outliers (Thomson Financial, 2004). We retrieve actual earnings per share (EPS) from *I/B/E/S* through 2007. To allow comparison using similar samples across forecast horizons, we require each firm year to have actual EPS for the upcoming five years.[6] Stock price, which is used as a deflator in some of the analyses, is acquired from the *CRSP* database. We keep only firm years with December fiscal year ends to align the time horizons for analysts' earnings forecasts in our sample. The analysts' earnings forecasts and the actual earnings, which are in per share format, are adjusted for stock dividends and stock splits to coincide with the number of shares outstanding as of the June base month. Furthermore, analysts' forecasts in fully diluted form are adjusted to the basic format. If, for some reason, the firm has yet to release its prior year earnings before the *I/B/E/S* June consensus earnings forecast period, we drop the observation. Our final sample contains 27,081 firm years. There are fewer firm years in the individual analyses due to missing forecasts from analysts and naïve models, missing actual EPS, or missing stock price when applicable.

Analyst and Model Forecasts

The first-year analysts' earnings forecasts are obtained from *I/B/E/S* and designated as year t (first-year) forecasts. For the subsequent four years, year $t+1$ through year $t+4$, explicit analysts' forecasts are obtained from *I/B/E/S*, if available. Explicit forecasts are almost always available for year $t+1$ but are usually unavailable for the long-term horizons, years $t+2$ through $t+4$. If an explicit forecast is not available, we calculate a forecast as follows:

$$\text{ANEPS}_{t+\tau} = \text{ANEPS}_{t+s} \times (1 + \text{LTG})^{\tau-s}$$

where ANEPS_{t+s} is the *I/B/E/S* median consensus analysts' EPS forecast for year $t+s$ (the last year with an explicit EPS forecast), LTG is the median consensus analysts' long-term earnings growth rate forecast on *I/B/E/S*, $\tau = 1, \ldots, 4$, $s = 0, \ldots, 3$, and $\tau > s$.[7] In this chapter, usually the second year's (year $t+1$) explicit EPS forecast is compounded at the long-term earnings growth rate to calculate the analysts' long-term earnings forecast. The compounding of the second year's analysts' earnings forecast with the analysts' long-term earnings growth rate to calculate the subsequent years' analyst earnings forecasts is common in the literature (Claus & Thomas, 2001; P. Easton et al., 2002; Frankel & Lee, 1998; Gebhardt, Lee, & Swaminathan, 2001; Hribar & Jenkins, 2004; and others).

We also produce earnings forecasts using two naïve statistical models, namely, a RW model and a RW with a drift based on past economic growth rate (RWGDP) model. The RW model is specified as follows:

$$\text{RW}_{t+\tau} = \text{EPS}_{t-1}$$

where EPS_{t-1} is last year's actual EPS, and $\tau = 0, \ldots, 4$.

The RWGDP model is specified as follows:

$$\text{RWGDP}_{t+\tau} = \text{EPS}_{t-1}(1 + g)^{\tau+1}$$

where g = historical inflation rate + historical growth in real GDP, and $\tau = 0, \ldots, 4$. The growth rate g is determined using the inflation rate and the growth in real GDP for year $t-1$. The historical inflation rate is retrieved from the Inflationdata.com web site (Capital Professional Services, 2009). The historical growth rate of GDP is based on GDP data at the web site of the U.S. Department of Commerce, Bureau of Economic Analysis (U.S. Department of Commerce, 2009).

We also calculate earnings forecasts using a hybrid (RWLTG) model that combines a RW based on prior year EPS with the analysts' long-term earnings growth forecast. The model is estimated as follows:

$$\text{RWLTG}_{t+\tau} = \text{EPS}_{t-1}(1 + \text{LTG})^{\tau+1}$$

where LTG is the *I/B/E/S* median consensus analysts' long-term earnings growth rate forecast, and $\tau = 0, \ldots, 4$.

An additional issue arises if ANEPS_{t+s} is negative for ANEPS calculations that require analysts' long-term earnings growth forecasts or if EPS_{t-1} is negative for the RWGDP and RWLTG models. First, it is unrealistic to assume that a firm can sustain an increasingly negative EPS over the forecast horizon. Second, positive earnings growth forecasts are meant to convey earnings increases. Therefore, when ANEPS_{t+s} or EPS_{t-1} is negative, we use the negative of the growth rate in formulating the forecast. This implies a reversion toward zero earnings for future periods if the growth rate is positive (most cases). For example, using the RWLTG model as an illustration and assuming that EPS_{t-1} is $-\$1.00$ and LTG is 10%; RWLTG_t is $-\$0.90$, RWLTG_{t+1} is $-\$0.81$, RWLTG_{t+2} is $-\$0.73$, and so on.

Measurement of Forecast Accuracy and Forecast Bias

To compare the forecast accuracy between analysts and naïve models, we calculate forecast error (FE) and relative forecast accuracy (RFA). We use two alternative deflators to calculate FEs. Specifically, we measure FE deflated by price (FE/P) as follows:

$$\frac{\left|\text{EPS}_{t+\tau} - \text{ANEPS}_{t+\tau}\ (\text{or STATEPS}_{t+\tau})\right|}{P_{t-1}} \tag{1}$$

and FE deflated by forecasted EPS (FE/EPS) as follows:

$$\frac{\left|\text{EPS}_{t+\tau} - \text{ANEPS}_{t+\tau}\ (\text{or STATEPS}_{t+\tau})\right|}{\left|\text{ANEPS}_{t+\tau}\ (\text{or STATEPS}_{t+\tau})\right|} \tag{2}$$

where $\text{EPS}_{t+\tau}$ is future actual EPS, $\text{STATEPS}_{t+\tau}$ is the earnings forecast generated by one of the naïve models or the hybrid model discussed above, P_{t-1} is the stock price per share for the end of May, the month previous to the base month, and $\tau = 0, \ldots, 4$.

We also measure the RFA, which directly compares the FE from the analysts' forecast with that from the naïve forecast. RFA deflated by price (RFA/P) is measured as follows:

$$\frac{\left(\left|\text{EPS}_{t+\tau} - \text{ANEPS}_{t+\tau}\right| - \left|\text{EPS}_{t+\tau} - \text{STATEPS}_{t+\tau}\right|\right)}{P_{t-1}}$$

while RFA deflated by EPS (RFA/E) is calculated as follows:

$$\frac{(|\mathrm{EPS}_{t+\tau} - \mathrm{ANEPS}_{t+\tau}| - |\mathrm{EPS}_{t+\tau} - \mathrm{STATEPS}_{t+\tau}|)}{|\mathrm{EPS}_{t+\tau}|}$$

A negative (positive) RFA value implies higher analyst (model) forecast accuracy.

The RFA measure differs from the FE measure. For FE, we calculate the absolute values of earnings FEs of analysts and those of a particular model at the individual observation level and then determine the significance of the difference in means (medians) between the two groups of FEs using a *t*-test (sign test). For RFA, we take the difference in the absolute FEs of analysts and the applicable model at the individual observation level and then measure whether the mean (median) of these differences is significantly different from zero through a *t*-test (sign test). FE and RFA serve as alternative measures of earnings forecast accuracy. The FEs above 1.0 are winsorized at 1.0 and the RFA measures are winsorized at $+1.0$ and -1.0 (Brown et al., 1987a; Fried & Givoly, 1982).

Testing Information Content of Analysts' Forecasts versus Model Forecasts

The above measures of forecast accuracy examine the magnitudes of the deviations of the forecasted earnings from the actual earnings. However, given the earnings forecast with higher accuracy, the earnings forecast with lower accuracy may also contain incrementally useful information in predicting future earnings. For instance, if analysts misestimate the persistence of the prior year's earnings, then a naïve model using the prior year's earnings would likely contain information incremental to that from analysts' forecasts even if analysts' forecasts happen to be more accurate. To explore the information content of analysts' forecasts and model forecasts, we run the following regression using OLS (Cheng et al., 2003; Granger & Newbold, 1973):

$$\frac{\mathrm{EPS}_{t+\tau}}{\mathrm{EPS}_{t-1}} - \frac{\mathrm{STATEPS}_{t+\tau}}{\mathrm{EPS}_{t-1}} = \alpha + \beta\left(\frac{\mathrm{ANEPS}_{t+\tau}}{\mathrm{EPS}_{t-1}} - \frac{\mathrm{STATEPS}_{t+\tau}}{\mathrm{EPS}_{t-1}}\right) + \varepsilon_{t+\tau} \quad (3)$$

where EPS is actual EPS, ANEPS is the analysts' forecast, STATEPS is the earnings forecast from one of the naïve models or the hybrid model, and $\tau = 0, \ldots, 4$. If all information in forecasting future actual earnings is provided by ANEPS, then β will equal one. On the contrary, if all information is provided by STATEPS, then β will equal zero. When information is provided by both ANEPS and STATEPS, $0 < \beta < 1$. It is possible that β could be greater than one or less than zero. In these situations, both forecasts have information content in explaining future earnings but investors put a negative weight on one of the forecasts.

Although Granger and Newbold (1973) hypothesize that the intercept term is zero, we follow Cheng et al. (2003) and include an intercept term to account for any bias in analysts' forecasts. To reduce excessive influence from outliers, we do two procedures. First, we winsorize the dependent variable and the independent variable at $+1.0$ and -1.0. Second, we eliminate outliers based on the guidelines of Belsley, Kuh, and Welsch (1980).

RESULTS

Full Sample

Panel A of Table 1 compares the earnings forecasts made by analysts with those from the RW model. The number of observations is lower for FE/P than FE/EPS due to the requirement of stock price from the *CRSP* database for FE/P.[8] An analysis of FE/P and FE/EPS shows that, in forecasting short-term earnings (years t and $t+1$), analysts' forecasts have significantly lower FEs than the RW model forecasts. For long-term forecasts, the results are mixed based on the FE measures. The median (mean and median) FE/P (FE/EPS) values convey that analysts tend to be more accurate over years $t+2$ through $t+4$. However, the results show that the forecast advantage for analysts steadily declines as the forecast horizon is extended. In fact, mean FE/P is significantly lower for RW forecasts at the 1% level in year $t+4$. An observation of RFA/P and RFA/EPS, which serve as alternative measures of forecast accuracy, confirms analyst superiority over the naïve model for short-term earnings forecasts. On the contrary, for years $t+3$ and $t+4$ (years $t+2$ through $t+4$), the positive mean values of RFA/P (RFA/EPS) signify that RW model forecasts are significantly more accurate at the 1% level. Nevertheless, the median values of RFA/P and RFA/EPS convey that analysts' forecasts are significantly more accurate than RW forecasts for all forecast horizons. Overall, analysts' forecasts outperform the RW model in forecasting short-term earnings. However, the conflicting forecast accuracy results do not support the superiority of either analysts or the RW model in forecasting long-term earnings, especially for years $t+3$ and $t+4$.

We also compute forecast bias, which is measured using Eqs. (1) and (2) except that the numerators are signed values instead of absolute values.

Table 1. Comparison of Forecasts between Analysts and Naïve Models.

		Mean					Median				
		Year t	$t+1$	$t+2$	$t+3$	$t+4$	Year t	$t+1$	$t+2$	$t+3$	$t+4$
Panel A: Analysts' forecasts versus random walk model											
FE/P	Analysts	2.036	3.885	4.941	5.881	7.056	0.408	0.981	1.374	1.816	2.312
	RW	3.198	4.453	4.966	5.615	6.340	0.833	1.376	1.751	2.143	2.478
	Difference	−1.161***	−0.568***	−0.025	0.266	0.716***	−0.426***	−0.395***	−0.378***	−0.327***	−0.166***
	N	12,527	12,248	10,959	10,820	10,782					
FE/EPS	Analysts	26.148	40.089	46.933	50.987	54.754	11.364	24.655	33.846	41.156	48.966
	RW	36.668	45.906	50.229	53.380	55.902	22.857	35.189	42.188	47.945	52.105
	Difference	−10.520***	−5.816***	−3.297***	−2.393***	−1.148***	−11.494***	−10.534***	−8.341***	−6.789***	−3.139***
	N	27,079	26,383	23,127	22,762	22,615					
RFA/P		−1.221***	−0.607***	0.030	0.393***	0.909***	−0.324***	−0.359***	−0.352***	−0.409***	−0.387***
RFA/EPS		−13.093***	−0.867***	6.896***	10.497***	13.693***	−9.756***	−9.155***	−6.500***	−5.438***	−2.166**
Panel B: Analysts' forecasts versus random walk with economic growth model											
FE/P	Analysts	2.036	3.885	4.941	5.881	7.056	0.408	0.981	1.374	1.816	2.312
	RWGDP	3.103	4.356	4.849	5.495	6.200	0.757	1.230	1.531	1.865	2.198
	Difference	−1.067***	−0.470***	0.092	0.386**	0.856***	−0.350***	−0.248***	−0.158***	−0.049	0.114**
	N	12,527	12,248	10,959	10,820	10,782					
FE/EPS	Analysts	26.148	40.089	46.934	50.989	54.756	11.364	24.648	33.849	41.165	48.968
	RWGDP	35.731	44.723	48.856	51.761	54.081	21.152	32.743	39.477	44.618	49.138
	Difference	−9.583***	−4.634***	−1.922***	−0.772**	0.675**	−9.789***	−8.094***	−5.628***	−3.453***	−0.170
	N	27,081	26,384	23,128	22,763	22,616					
RFA/P		−1.119***	−0.481***	0.214***	0.550***	1.098***	−0.210***	−0.183***	−0.111**	−0.081**	0.027
RFA/EPS		−12.702***	−1.315***	6.433***	10.537***	14.671***	−6.695***	−5.032***	−1.938***	−0.045	3.335***

Notes: All values are shown as percentages. FE/P is forecast error deflated by price, specified as $(|\text{EPS}_{t+\tau}-\text{ANEPS}_{t+\tau}$ (or $\text{STATEPS}_{t+\tau})|)/P_{t-1}$, where EPS is actual annual earnings per share, ANEPS is analyst forecasted earnings per share, STATEPS is earnings per share estimated with one of the naive models, and P is stock price per share. FE/EPS is forecast error deflated by earnings per share, specified as $(|\text{EPS}_{t+\tau}-\text{ANEPS}_{t+\tau}$ (or $\text{STATEPS}_{t+\tau})|)/|\text{ANEPS}_{t+\tau}$ (or $\text{STATEPS}_{t+\tau})|$, where EPS is actual annual earnings per share, ANEPS is analyst forecasted earnings per share, and STATEPS is earnings per share estimated with one of the naive models. RFA/P is relative forecast accuracy deflated by price, specified as $(|\text{EPS}_{t+\tau}-\text{ANEPS}_{t+\tau}|-|\text{EPS}_{t+\tau}-\text{STATEPS}_{t+\tau}|)/P_{t-1},$ where EPS is actual annual earnings per share, ANEPS is analyst forecasted earnings per share, STATEPS is earnings per share estimated with one of the naive models, and P is stock price per share. RFA/EPS is relative forecast accuracy deflated by earnings per share, specified as $(|\text{EPS}_{t+\tau}-\text{ANEPS}_{t+\tau}|-|\text{EPS}_{t+\tau}-\text{STATEPS}_{t+\tau}|)/|\text{EPS}_{t+\tau}|$, where EPS is actual annual earnings per share, ANEPS is analyst forecasted earnings per share, and STATEPS is earnings per share estimated with one of the naive models. The measures (FE/P, RFA/P, etc.) are winsorized at −1.0 (if applicable) and +1.0. ***Significance at the 0.01 level (two-tailed). **Significance at the 0.05 level (two-tailed). *Significance at the 0.10 level (two-tailed).

The untabulated statistics show that analysts' earnings forecast bias values indicate analyst optimism, which increases as the forecast horizon is extended. This is consistent with the literature. The RW forecasts convey that they are pessimistically biased, which is not surprising because the assumption with RW forecasts is no growth over prior year's earnings.

Table 1, panel B, compares analysts' earnings forecasts with forecasts from the RWGDP model. Similar to the results in panel A, analysts are superior in forecasting short-term earnings. On the contrary, the findings are mixed with respect to long-term forecasts. An observation of mean FE/P shows that RWGDP long-term forecasts have lower FEs for year $t+3$ (at the 5% significance level) and year $t+4$ (at the 1% significance level). The results for median FE/P convey that analysts' FEs are significantly lower at the 1% level for year $t+2$, there is no significant difference for year $t+3$, and RWGDP model FEs are significantly lower at the 5% level for year $t+4$. The results for mean and median values of FE/EPS convey that analysts are more accurate for years t through $t+3$. However, the findings with respect to mean (median) values of FE/EPS in year $t+4$ indicate lower RWGDP model FEs (no significant difference in FEs). Turning to the alternative measures of forecast accuracy, the positive mean values of RFA/P and RFA/EPS for years $t+2$ through $t+4$ imply that RWGDP long-term forecasts are significantly more accurate at the 1% level. The median values of RFA/P indicate higher accuracy for analysts' forecasts in years $t+2$ and $t+3$ (at the 5% level) and no significant difference in year $t+4$. The median values of RFA/EPS show that while analysts are significantly more accurate at the 1% level in year $t+2$, there is no significant difference in year $t+3$, and the RWGDP model has significantly higher accuracy at the 1% level in year $t+4$. Overall, the results in panel B do not support the conjecture that analysts outperform the RWGDP model in making long-term earnings forecasts. Also, the accuracy of RWGDP model forecasts improves relative to analysts' forecasts as the forecast horizon is extended. The results provide some evidence on the superiority of RWGDP model forecasts over analysts' forecasts for year $t+4$.

The regression results from Eq. (3) with analysts' earnings forecasts and RW earnings forecasts are listed in Table 2, panel A.[9] The parameter β is significantly greater than zero for all forecast periods, indicating that analysts' forecasts have information content in explaining future actual earnings. However, β is also significantly less than one for all forecast horizons, which implies that RW forecasts provide incremental information over analysts' forecasts. The value of β is 0.82 in year t, which conveys that analysts' forecasts for the first year play more of a role in assimilating information about future earnings than do RW model forecasts.

Table 2. Regression Analysis of Information Content of Analysts' Forecasts versus Naïve Model Forecasts.

	α		β	
Forecast Period	Coefficient	*p*-Value	Coefficient	*p*-Value
Panel A: Analysts' forecasts versus random walk model				
t	−0.05	0.00	0.82	0.00
$t+1$	−0.08	0.00	0.64	0.00
$t+2$	−0.05	0.00	0.50	0.00
$t+3$	−0.02	0.00	0.46	0.00
$t+4$	0.00	0.69	0.42	0.00
Panel B: Analysts' forecasts versus. random walk with economic growth model				
t	−0.06	0.00	0.81	0.00
$t+1$	−0.11	0.00	0.64	0.00
$t+2$	−0.12	0.00	0.52	0.00
$t+3$	−0.13	0.00	0.49	0.00
$t+4$	−0.14	0.00	0.46	0.00

Notes:

1. The regression model is as follows:

$$\frac{\mathrm{EPS}_{t+\tau}}{\mathrm{EPS}_{t-1}} - \frac{\mathrm{STATEPS}_{t+\tau}}{\mathrm{EPS}_{t-1}} = \alpha + \beta\left(\frac{\mathrm{ANEPS}_{t+\tau}}{\mathrm{EPS}_{t-1}} - \frac{\mathrm{STATEPS}_{t+\tau}}{\mathrm{EPS}_{t-1}}\right) + \varepsilon_{t+\tau}$$

where EPS is actual annual earnings per share, ANEPS is the analysts' earnings per share forecast, STATEPS is the earnings per share forecast from one of the naïve models (random walk, random walk with economic growth), and $\tau = 0, \ldots, 4$.

2. The dependent and independent variables are winsorized at $+1.0$ and -1.0. Furthermore, outliers are eliminated using the techniques in Belsley et al. (1980).
3. The *p*-values show the significance of the difference from zero.

Nevertheless, the coefficient β steadily decreases as the forecast horizon is extended. Its value is 0.50, 0.46, and 0.42 for years $t+2$, $t+3$, and $t+4$, respectively. The substantially lower coefficients in years $t+2$ through $t+4$ suggest that for longer-term forecasts, much of the information content in explaining future actual earnings originates from the RW model instead of analysts' forecasts. This is likely in part due to (1) less of a timing advantage for analysts in forecasting long-term earnings as opposed to short-term earnings and (2) analysts' high optimism in forecasting long-term earnings.

Table 2, panel B, presents the results from regression Eq. (3) with RWGDP as the naïve model. The results are similar to those in panel A, where RW is the naïve model. The coefficient β in panel B does have a slightly smaller (larger) value than the corresponding coefficient in panel A for year t (years $t+2$ through $t+4$). A two-tailed t-test shows that the difference in coefficients is significant for year t at the 1% level and year $t+2$ at the 5% level.[10] This implies that RWGDP model earnings forecasts contain slightly more (less) information in explaining future earnings that is not in analysts' earnings forecasts than do RW model earnings forecasts for years t (year $t+2$). Furthermore, for years t through $t+4$ in panel B, we find that the coefficient α is significantly less than zero, which is indicative of an optimistic bias in analysts' forecasts.

Sample Partitions and Hybrid Model

Prior research (e.g., Alford & Berger, 1999; Chan et al., 2003) suggests that the performance of financial analysts versus naïve models may be influenced by various attributes. Therefore, we evaluate the performance of analysts' earnings forecasts versus RWGDP model earnings forecasts across different sample partitions. The sample partitions are based on past earnings growth, analyst coverage, and a subsample with only explicit analysts' forecasts. Also, we compare the hybrid model, RWLTG, with the RWGDP model and analysts' forecasts. The objective is to determine whether improvements in accuracy and information content can be achieved by applying the analysts' forecasted long-term earnings growth rate to last year's (year $t-1$) earnings. For brevity, of the naïve models, we analyze only the RWGDP model in these additional tests because it is the most accurate.

Partitioning on Past Earnings Growth

Chan et al. (2003) show that very few firms are able to consistently achieve above-normal earnings growth over five years and the probability of doing so is about equal to pure chance. Furthermore, their findings suggest that financial analysts may incorrectly assume that past above-normal earnings growth will continue well into the future. However, the authors do not explicitly test this conjecture. If analysts often assume that high past earnings growth will continue well into the future, then based on findings in Chan et al. (2003), we would expect analysts' earnings forecasts for high past growth firms to have less accuracy, more bias, and less information content in explaining future actual earnings.

To test whether higher past earnings growth affects the performance of analysts' earnings forecasts relative to naïve forecasts (specifically, the RWGDP forecasts), we partition our sample according to past earnings growth. Past earnings growth is measured as the geometric growth in earnings between year $t-5$ and year $t-1$. It is necessary to mention two limitations of using the past geometric growth rate. First, only sample firms with positive year $t-5$ and positive year $t-1$ earnings can be used. Second, only firms with sufficient earnings histories are included. This may favor analysts' forecasts over RWGDP model forecasts because analysts tend to make more accurate forecasts for firms that are more mature. Firms with earnings growth rates above (below) the median level of 8.63% are designated as high (low) growth firms. This median growth rate is determined before observations are eliminated due to missing future actual earnings.

Table 3, panel A and panel B, presents the results for high and low past earnings growth firms, respectively. There are fewer observations in panel B because the low past growth subsample includes more firms that were in financial trouble, which means more bankruptcies and delistings and fewer observations with five years of future actual earnings. For both high past growth and low past growth firms, the majority of the FE (FE/P and FE/EPS) and RFA (RFA/P and RFA/EPS) values show that analysts are more accurate than the RWGDP model in forecasting short-term (year t and year $t+1$) earnings.

The nature of the findings changes for long-term earnings forecasts, which are the focus of our analysis. A comparison of panels A (high past earnings growth) and B (low past earnings growth) shows that the performance of analysts tends to improve relative to the RWGDP model when the past earnings growth is low. For the high past earnings growth subsample, the mean (median) FE measures FE/P, FE/EPS, RFA/P, and RFA/EPS imply consistently *lower* RWGDP model FEs than analysts' FEs at the 1% level over years $t+3$ and $t+4$ (year $t+4$). However, for low past earnings growth firms, the results are mixed with the mean RFA/EPS measure indicating lower FE for the RWGDP model and the median FE/P, FE/EPS, RFA/P, and RFA/EPS measures indicating lower errors for analysts' forecasts for years $t+2$ through $t+4$. Overall, for firms with high past earnings growth, the results imply a lower level of accuracy for financial analysts' earnings forecasts compared to the naïve RWGDP model forecasts for years $t+3$ and $t+4$. On the contrary, for firms with low past earnings growth, the results are mixed.

Table 3. Comparison of Forecasts between Analysts and Random Walk with Economic Growth Model; Observations Partitioned by Past Earnings Growth.

		Mean					Median				
		Year t	$t+1$	$t+2$	$t+3$	$t+4$	Year t	$t+1$	$t+2$	$t+3$	$t+4$
Panel A: High past earnings growth											
FE/P	Analysts	1.238	2.821	4.024	4.885	6.211	0.267	0.714	1.161	1.535	2.155
	RWGDP	1.936	3.010	3.677	4.165	5.072	0.526	0.926	1.229	1.462	1.808
	Difference	−0.698***	−0.189	0.347*	0.720***	1.139***	−0.259***	−0.212***	−0.068	0.073	0.347***
	N	4,846	4,790	4,523	4,485	4,473					
FE/EPS	Analysts	17.852	32.613	41.495	46.566	51.341	6.937	16.667	25.940	33.215	41.152
	RWGDP	24.978	35.300	40.612	43.836	46.639	13.250	22.188	28.674	33.128	36.779
	Difference	−7.126***	−2.687***	0.883	2.730***	4.702***	−6.313***	−5.521***	−2.734***	0.087	4.373***
	N	8,244	8,130	7,672	7,621	7,600					
RFA/P		−0.766***	−0.163*	0.431***	0.905***	1.433***	−0.183***	−0.169***	−0.054**	0.052	0.306***
RFA/EPS		−10.627***	−1.426***	7.066***	12.654***	18.181***	−5.487***	−4.648***	−0.803	2.867***	8.417***
Panel B: Low past earnings growth											
FE/P	Analysts	1.494	2.801	3.497	4.043	4.798	0.379	0.872	1.160	1.464	1.865
	RWGDP	2.307	3.125	3.479	4.017	4.536	0.706	1.085	1.397	1.725	2.012
	Difference	−0.813***	−0.324**	0.018	0.026	0.262	−0.327***	−0.213***	−0.237***	−0.261***	−0.147**
	N	4,636	4,556	4,175	4,134	4,119					
FE/EPS	Analysts	24.806	36.295	41.197	43.935	46.458	10.345	20.690	26.186	30.751	34.877
	RWGDP	33.659	40.624	44.161	47.236	49.376	20.201	29.240	34.544	39.998	43.479
	Difference	−8.853***	−4.329***	−2.964***	−3.301***	−2.918***	−9.856***	−8.550***	−8.358***	−9.247***	−8.602***
	N	7,667	7,530	6,888	6,834	6,812					
RFA/P		−0.833***	−0.373***	0.068	0.092	0.228**	−0.195***	−0.149***	−0.130***	−0.131***	−0.127***
RFA/EPS		−10.267***	0.511	5.119***	6.500***	7.879***	−5.324***	−3.830***	−2.841***	−2.783***	−2.461***

Notes: All values are shown as percentages. For the observations on the *I/B/E/S* database for June of each year from 1988 to 2007 that have the prior five years of earnings, we find the geometric growth rate in earnings from year $t-5$ to year $t-1$. Panel A (B) presents the results for sample observations with above (below) median prior earnings growth. The forecast measures (FE/P, RFA/P, etc.) are winsorized at −1.0 (if applicable) and +1.0. For variable definitions, see Table 1.***Significance at the 0.01 level (two-tailed).**Significance at the 0.05 level (two-tailed).*Significance at the 0.10 level (two-tailed).

The untabulated bias statistics suggest that for short-term forecasts (years t and $t+1$), analysts' forecasts are less optimistically biased for high past growth firms compared with low past growth firms. However, for longer horizons, analysts' forecasts are more optimistically biased for high past growth firms than low past growth firms, and the difference becomes larger as the forecast horizon is extended. Although financial analysts may often be correct to assume that high past earnings growth will continue over the short term, the bias results imply that analysts may tend to incorrectly assume that high past earnings growth will continue well into the future. This is further supported by the FE (FE/P and FE/EPS) statistics for analysts in Table 3. Although analysts' FEs tend to be lower for high past growth firms in years t and $t+1$, they are clearly higher for high past growth firms in years $t+3$ and $t+4$.[11]

Table 4 summarizes the results from regression Eq. (3) with panel A presenting the results for high past earnings growth firms and panel B displaying the findings for low past earnings growth firms. The coefficient β is higher for high past growth firms for forecast horizons t and $t+1$. However, the situation reverses in years $t+2$ through year $t+4$. The differences are significant at the 1% level for all years except year $t+2$. These results imply that analysts' forecasts have more incremental information content over the RWGDP model in explaining long-term future actual earnings for low past growth firms than for high past growth firms.

Partitioning on Analyst Following

Prior research (Alford & Berger, 1999; Brown, 1997; Coën, Desfleurs, & L'Her, 2009; Lim, 2001; Lys & Soo, 1995) provides evidence that higher analyst following is associated with greater analyst forecast accuracy. Analysts tend to follow firms with information that is more extensive and accurate. This reduces the uncertainty about the firms' prospects and helps analysts to make more accurate earnings forecasts. We partition our sample according to analyst following and examine the performance of analysts' long-term forecasts and the RWGDP model for the sub-samples. Firm years with long-term growth forecasts from more than three (three or fewer) analysts are considered firms with high (low) analyst following.

Untabulated results show that both analysts' forecasts and RWGDP model forecasts are more accurate when there is high analyst following compared with low analyst following. This result is consistent with Previts, Bricker, Robinson, and Young (1994), who show that financial analysts tend to follow firms that smooth earnings. If firms smooth earnings, they

Table 4. Regression Analysis of Information Content of Analysts' Forecasts versus Random Walk with Economic Growth Model; Observations Partitioned by Past Earnings Growth.

	α		β	
Forecast Period	Coefficient	*p*-Value	Coefficient	*p*-Value
Panel A: High past earnings growth				
t	−0.05	0.00	0.99	0.00
$t+1$	−0.12	0.00	0.72	0.00
$t+2$	−0.14	0.00	0.51	0.00
$t+3$	−0.14	0.00	0.42	0.00
$t+4$	−0.17	0.00	0.40	0.00
Panel B: Low past earnings growth				
t	−0.07	0.00	0.81	0.00
$t+1$	−0.10	0.00	0.63	0.00
$t+2$	−0.10	0.00	0.54	0.00
$t+3$	−0.11	0.00	0.55	0.00
$t+4$	−0.13	0.00	0.57	0.00

Notes:

1. For observations on the *I/B/E/S* database for June of each year from 1988 to 2007 that have five prior years of earnings, we find the geometric growth rate in earnings from year $t-5$ to year $t-1$. Panel A (B) presents the results for observations with above (below) median prior earnings growth.
2. The regression model is as follows:

$$\frac{\mathrm{EPS}_{t+\tau}}{\mathrm{EPS}_{t-1}} - \frac{\mathrm{RWGDP}_{t+\tau}}{\mathrm{EPS}_{t-1}} = \alpha + \beta\left(\frac{\mathrm{ANEPS}_{t+\tau}}{\mathrm{EPS}_{t-1}} - \frac{\mathrm{RWGDP}_{t+\tau}}{\mathrm{EPS}_{t-1}}\right) + \varepsilon_{t+\tau}$$

where EPS is actual annual earnings per share, ANEPS is the analysts' earnings per share forecast, RWGDP is the earnings per share forecast from the random walk with economic growth model, and $\tau = 0, \ldots, 4$.

3. The dependent and independent variables are winsorized at $+1.0$ and -1.0. Furthermore, outliers are eliminated using the techniques in Belsley et al. (1980).
4. The *p*-values test the significance of the difference from zero.

are easier to predict by analysts and a RW with a drift model such as RWGDP should be more accurate. Furthermore, for long-term earnings forecasts, the findings on accuracy convey that analysts' forecasts moderately improve relative to RWGDP model forecasts when there is high analyst following. The results from regression Eq. (3) show that the coefficient β is significantly larger at the 1% level for the high analyst following subsample than for the low analyst following subsample for all five years. These results imply that financial analysts' forecasts have more information content in explaining future actual earnings for firms with high analyst coverage.

Explicit Analysts' Forecasts

Due to a scarcity of explicit analysts' long-term earnings forecasts (e.g., fourth-year EPS is expected to be \$2.50), most of the long-term earnings forecasts are calculated through compounding the analysts' second-year earnings forecast with the analysts' long-term earnings growth rate. However, it is possible that the accuracy of analysts' forecasts versus naïve models is different when analysts make explicit forecasts. Therefore, we also run our tests using only explicit forecasts from analysts.

The untabulated results show that the number of explicit forecasts drops precipitously between year $t+1$ and year $t+2$. The FEs (FE/P and FE/EPS) indicate that both analysts' forecasts and RWGDP model forecasts are more accurate for years $t+3$ and $t+4$ for the explicit forecast sample compared with the results for the entire sample noted in Table 1, panel B. This conveys that analysts tend to issue explicit long-term forecasts when earnings are easier to predict. However, the accuracy of analysts' earnings forecasts relative to RWGDP model forecasts for year $t+2$ does not improve when analysts make explicit forecasts. Nonetheless, when analysts make explicit forecasts, there is improvement in the accuracy of analysts' forecasts relative to RWGDP model forecasts for year $t+4$. On the contrary, explicit analysts' for year $t+4$ are scarce. For instance, there are only 1,323 (1,939) year $t+4$ explicit analysts' forecasts available when stock price (EPS) is the deflator. The untabulated regression results are in line with the forecast accuracy results. When analysts make explicit forecasts, the Eq. (3) coefficient β for year $t+2$ ($t+4$) is significantly less (greater) than the corresponding coefficient value in Table 2, panel B, at the 1% level.

Hybrid Model Forecasts

We compare the hybrid model, RWLTG, with the RWGDP model and analysts' earnings forecasts through variations of the previously discussed tests of accuracy and information content. Untabulated results show that combining a naïve model with analysts' long-term earnings growth rate forecasts does not improve forecast accuracy. In matching RWLTG against

RWGDP, median (mean) values indicate that the RWLTG (RWGDP) model is more accurate in forecasting short-term earnings. However, the RWLTG model is inferior to the RWGDP model in long-term earnings forecast accuracy. In addition, the RWLTG model is less accurate than analysts' forecasts in years t and $t+1$. However, the difference in forecast accuracy gets smaller as the forecast horizon is extended. In fact, there is no significant difference in forecast accuracy between the RWLTG model and analysts' forecasts for year $t+4$.

Untabulated regression results using the RWLTG and RWGDP models show that both models have incremental information content in explaining future actual earnings but that the RWGDP model has more information content. Similarly, although both analysts' earnings forecasts and the RWLTG model have incremental information content in explaining future actual earnings, analysts' forecasts have more information content.

CONCLUSIONS

We examine the performance of financial analysts versus naïve models in forecasting long-term earnings. Forecast performance is evaluated through analyzing forecast accuracy and information content. We find that analysts' long-term earnings forecasts (especially for the fourth year and fifth year in the forecast horizon) are often less accurate than forecasts from naïve models. Furthermore, both naïve model earnings forecasts and analysts' long-term earnings forecasts contain information content in predicting long-term earnings. Also, we find that the performance of analysts' forecasts declines relative to naïve model forecasts for subsamples of firms with high past earnings growth and low analyst following. When analysts make explicit earnings forecasts, the performance of analysts' forecasts increases compared to naïve model forecasts for only the fifth year in the forecast horizon. But explicit analysts' forecasts for the fifth year are scarce. Moreover, we test the accuracy and information content of a hybrid model that assumes a RW with a drift based on the analysts' long-term earnings growth rate. We find that this hybrid model is less accurate and has less information content in predicting long-term earnings than the RWGDP model or financial analysts.

Our findings imply that analysts' long-term earnings forecasts should be used with caution by researchers and practitioners as they do not appear to be more accurate than long-term forecasts from naïve models. Furthermore, the naïve models incorporate a large amount of information content useful in explaining future actual earnings that is not in analysts' long-term earnings forecasts. Researchers and practitioners should be especially cautious when using analysts' long-term earnings forecasts for firms with high recent earnings growth. Furthermore, our findings indicate that it may be appropriate to use strong performing naïve models such as the RWGDP model or a pure RW model as a substitute for missing analysts' long-term earnings forecasts in applications such as implementing valuation models.

NOTES

1. Not all naïve forecasts are technically time series forecasts. For example, a pure RW forecast that uses the prior period's earnings as a forecast of future earnings is not a time series forecast because it is not based on a series of time periods. However, time series forecasts are naïve because they are mechanically based on past information. The term ''time series forecast'' is often used loosely in the literature.

2. For example, Dechow et al. (2000) examine the performance of analysts' long-term earnings growth forecasts that pertain to a sample of firms that recently issued equity.

3. The *I/B/E/S* database rarely provides forecast information pertaining to years after the fifth year.

4. The RW model assumes that future annual earnings will equal the most recent prior year's actual earnings.

5. We use June consensus forecasts because we use only December fiscal year-end firms. Thus, as of June, the previous year's financial results are likely to have been released. Also, the focus of this chapter is on long-term forecasts. The forecast month does not have as much of an impact on long-term forecasts as it would on short-term forecasts.

6. This requirement would likely favor analysts because they tend to forecast with more accuracy for firms that are more stable.

7. In defining the variables in this chapter, the firm subscript is suppressed.

8. It is only necessary to show the numbers of observations for the mean values of FE/P and FE/EPS because the numbers of observations are the same in the other related parts of the panel. There is a moderate drop in the number of observations between year $t+1$ and year $t+2$ because only short-term analysts' earnings forecasts are available for some firm years. Also, there is a slight decline in the number of observations over the long-term forecast horizons. As mentioned in the section on Analyst and Model Forecasts, we retrieve explicit EPS forecasts for the long-term horizons, if possible. Some firm years have a per share forecast for one or two long-term forecast period(s) (e.g., years $t+2$ and $t+3$) but not subsequent long-term forecast period(s) (e.g., year $t+4$).

9. In the regression analyses in this chapter, we test for heteroskedasticity using methodology from White (1980) and find that heteroskedasticity is not a problem.

10. We use a two-tailed t-test to conduct statistical comparisons of the values of the coefficient β in panel A with those in panel B for Tables 2 and 4. For the sake of

simplicity, we just discuss the results in the text and do not report the statistical significance in the tables.

11. We also determine analysts' long-term earnings growth rate forecasts for high and low past earnings growth firms. The mean (median) growth rate forecast is 15.37% (14.0%) and 12.55% (11.0%) for high and low past growth firms, respectively. The differences in the means and the medians are significant at the 1% level. Therefore, these findings show that analysts are more optimistic in their long-term earnings growth forecasts for firms with higher past earnings growth.

ACKNOWLEDGMENTS

We thank Jian Cao, Hui Du, Barry Marks, and Haeyoung Shin for their helpful comments and suggestions. Also, we thank participants at the 2010 American Accounting Annual Meeting and the 2010 Southwest Region American Accounting Association Annual Meeting for useful discussions. The second author acknowledges a 2009 summer research grant from the College of Business at Prairie View A&M University.

REFERENCES

Alford, A. W., & Berger, P. G. (1999). A simultaneous equations analysis of forecast accuracy, analyst following, and trading volume. *Journal of Accounting, Auditing & Finance*, *14*(Summer), 219–240.

Allee, K. (2009). *Estimating cost of equity capital with time-series forecasts of earnings*. Working paper. Michigan State University, East Lansing, MI.

Belsley, D., Kuh, E., & Welsch, R. (1980). *Regression diagnostics*. New York, NY: Wiley.

Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting and Economics*, *11*(2–3), 255–274.

Botosan, C., & Plumlee, M. (2005). Assessing alternative proxies for the expected risk premium. *The Accounting Review*, *80*(January), 21–53.

Bradshaw, M. (2004). How do analysts use their earnings forecasts in generating stock recommendations? *The Accounting Review*, *79*(January), 25–50.

Brown, L. (1993). Earnings forecasting research: Its implications for capital markets research. *International Journal of Forecasting*, *9*, 295–320.

Brown, L. (1997). Analyst forecasting errors: Additional evidence. *Financial Analysts Journal*, *53*(November/December), 81–88.

Brown, L., Griffin, P., Hagerman, R., & Zmijewski, M. (1987a). Security analyst superiority relative to time-series models in forecasting quarterly earnings. *Journal of Accounting and Economics*, *9*, 61–87.

Brown, L., Griffin, P., Hagerman, R., & Zmijewski, M. (1987b). An evaluation of alternative proxies for the market's assessment of unexpected earnings. *Journal of Accounting and Economics*, *9*, 159–193.

Brown, L., Richardson, G., & Schwager, S. (1987). An information interpretation of financial analyst superiority in forecasting earnings. *Journal of Accounting Research*, *25*(Spring), 49–67.

Brown, L., & Rozeff, M. (1978). The superiority of analyst forecasts as measures of expectations: Evidence from earnings. *Journal of Finance*, *33*(March), 1–16.

Capital Professional Services. (2009). InflationData.com. Historical US inflation rate 1914-present. Retrieved from http://inflationdata.com/Inflation/Inflation_rate/historicalinflation.aspx

Chan, L., Karceski, J., & Lakonishok, J. (2003). The level and persistence of growth rates. *Journal of Finance*, *58*(April), 643–684.

Cheng, J., Fan, D., & So, R. (2003). On the performance of naïve, analyst and composite earnings forecasts: Evidence from Hong Kong. *Journal of International Financial Management & Accounting*, *14*, 146–165.

Claus, J., & Thomas, J. (2001). Equity Premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stocks. *Journal of Finance*, *56*, 1629–1666.

Coën, A., Desfleurs, A., & L'Her, J. (2009). International evidence on the relative importance of the determinants of earnings forecast accuracy. *Journal of Economics and Business*, *61*(6), 453–471.

Conroy, R., & Harris, R. (1987). Consensus forecasts of corporate earnings: Analysts' forecasts and time series methods. *Management Science*, *33*, 725–738.

Cragg, J., & Malkiel, B. (1968). The consensus and accuracy of some predictions of the growth of corporate earnings. *Journal of Finance*, *23*(March), 67–84.

Dechow, P., Hutton, A., & Sloan, R. (2000). The relation between analysts' forecasts of long-term earnings growth and stock price performance following equity offerings. *Contemporary Accounting Research*, *17*(Spring), 1–32.

Easton, P., & Monahan, S. J. (2005). An evaluation of accounting-based measures of expected returns. *The Accounting Review*, *80*, 501–538.

Easton, P., & Sommers, G. (2007). Effect of analysts' optimism on estimates of the expected rate of return implied by earnings forecasts. *Journal of Accounting Research*, *45*(5), 983–1015.

Easton, P., Taylor, G., Shroff, P., & Sougiannis, T. (2002). Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. *Journal of Accounting Research*, *40*(June), 657–676.

Francis, J., Olsson, P., & Oswald, D. (2000). Comparing the accuracy and explainability of dividend, free cash flow, and abnormal earnings equity value estimates. *Journal of Accounting Research*, *38*, 45–70.

Frankel, R., & Lee, C. (1998). Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics*, *25*, 283–319.

Fried, D., & Givoly, D. (1982). Financial analysts' forecasts of earnings: A better surrogate for market expectations. *Journal of Accounting and Economics*, *4*, 85–107.

Gebhardt, W., Lee, C., & Swaminathan, B. (2001). Toward an implied cost of capital. *Journal of Accounting Research*, *39*(June), 135–176.

Granger, C., & Newbold, P. (1973). Some comments on the evaluation of economic forecasts. *Applied Economics*, *5*, 35–47.

Hong, H., Lim, T., & Stein, J. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance*, *55*, 265–295.

Hou, K., van Dijk, M., & Zhang, Y. (2010). *The implied cost of capital: A new approach*. Working paper. Ohio State University, Erasmus University, and Chinese University of Hong Kong.

Hribar, P., & Jenkins, N. T. (2004). The effect of accounting restatements on earnings revisions and the estimated cost of capital. *Review of Accounting Studies*, *9*, 337–356.

Imhoff, E., & Pare, P. (1982). Analysis and comparison of earnings forecast agents. *Journal of Accounting Research*, *20*(Autumn), 429–439.

Lim, T. (2001). Rationality and analysts' forecast bias. *Journal of Finance*, *56*(1), 369–385.

Lys, T., & Soo, L. G. (1995). Analysts' forecast precision as a response to competition. *Journal of Accounting, Auditing and Finance*, *10*(Fall), 751–763.

Newbold, P., Zumwalt, J., & Kannan, S. (1987). Combining forecasts to improve earnings per share prediction. *International Journal of Forecasting*, *3*, 229–238.

Previts, G., Bricker, R., Robinson, T., & Young, S. (1994). A content analysis of sell-side financial analyst company reports. *Accounting Horizons*, *8*, 55–70.

Ramnath, S., Rock, S. K., & Shane, P. B. (2008). The financial analyst forecasting literature: A taxonomy with suggestions for further research. *International Journal of Forecasting*, *24*, 34–75.

Rozeff, M. (1983). Predicting long-term earnings growth: Comparisons of expected return models, submartingales and value line analysts. *Journal of Forecasting*, *2*, 425–435.

Schipper, K. (1991). Analysts' forecasts. *Accounting Horizons*, *5*, 105–131.

Sougiannis, T., & Yaekura, T. (2001). The accuracy and bias of equity values inferred from analysts' earnings forecasts. *Journal of Accounting, Auditing and Finance*, *16*(Fall), 331–362.

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

Thomson Financial. (2004). Thomson Financial Glossary 2004: A Guide to Understanding Thomson Financial Terms and Conventions for the First Call and I/B/E/S Databases.

U.S. Department of Commerce. (2009). Bureau of Economic Analysis website. Retrieved from http://www.bea.gov/

White, H. (1980). A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica*, *48*, 149–170.

The Size Premium in the Long Run

Ching-Chih Lu*

This Draft: December 25, 2009

Abstract

Contrary to the usual practice of including a size premium in a small firm's cost-of-equity estimation, this paper shows that there should not be such a premium in the long run because firm size is a changing characteristic. By tracking the return performance of firms in the same size group for a longer horizon, I find that the size premium wears off just after two years. This is much shorter than the general assumption used in the cost-of-equity estimation, so the role of the size premium in it should be reconsidered.

Keywords: Cost of Equity Capital, Size Premium, Size Effect, Regime Switching

JEL Classification: G12, G14

*Department of Finance, National Chengchi University, No. 64, Sec.2 Zhinan Rd., Mucha, Taipei 116, Taiwan. E-mail: cclu@nccu.edu.tw.

1 Introduction

In the field of business valuation, practitioners usually include a size premium in a small firm's cost-of-equity estimation to account for a risk source or risk sources that cannot be captured by usual risk factors.[1] That is, on top of the cost of equity a small firm gets from the estimation by the CAPM or other models, it is usually offered an extra premium to compensate for the higher risk it is taking.[2] This paper aims to examine its validity, and the finding suggests that this commonly accepted size premium is not appropriate.

Since Banz (1981) and Reinganum (1981) both demonstrated that small size firms on the New York Stock Exchange usually outperform big firms than what the asset-pricing model of Sharpe (1964), Lintner (1965) and Black (1972) would suggest, the existence of the size effect has come into consideration by standard practice in the finance industry and soon became one of the most exploited concepts in modern finance. This size anomaly leads to an assumption that it might stem from a risk source or risk sources which cannot be explained by the market factor. Berk (1995) explains in theory that market value is inversely correlated with unmeasured risk because investors pay a lower price for a company's stock if it bears a higher risk than its CAPM beta could measure. The seminal works of Fama and French (1993), and Fama and French (1995) also acknowledge another kind of size effect in which

[1] Although there are many ways to define the size of a company, I stick to the most popular criteria, the market value of its equity, to proceed the discussion.

[2] Other than the CAPM, the build-up method and the Fama-French 3-factor model are also popular approaches in business valuation. The build-up method is advocated by the Ibbotson Associates, now a part of Morningstar, Inc., which aims to break down the expected return of a firm into a risk-free rate, a premium for equity risk, a risk premium attributable to this company by the industry it is in, and another risk premium for smaller size if applicable. This size premium is added in practice no matter whether the CAPM model or the build-up method is used. Please see Pratt and Grabowski (2008) Chapter 12 for a thorough discussion. Such a size premium is not required in the Fama-French 3-factor model because size is a risk factor embedded in it already.

small firms usually outperform big firms in realize returns and they use the return differential between small and big stork portfolios (I call it “small stock premium" hereafter for convenience) as a risk factor (also known as *SMB*). If the CAPM holds well, the small stock premium should be proportional to the difference between the CAPM betas of small and big stock portfolios in cross section, and the size premium should not exist. However, empirical evidence shows that the small stock premium is usually much bigger than the CAPM could explain because small firms usually have a significant size premium, which links these two different perspectives of size anomalies together.

Besides serving as a measure of an alternative risk source, the idea of the existence of a small stock premium is often used in forming a trading strategy. Since the commence of the Dimensional Fund Advisors (DFA hereafter) in 1981, the strategy of overweighing small-cap stocks to exploit this small stock premium has been utilized extensively. This same concept is also used to construct ETFs featuring size as an important characteristic. There are currently at least 6 micro-cap and 40 small-cap ETFs trading on the U.S. stock exchanges.[3] The main attraction of these ETFs is to exploit their potentially higher returns over big firms or the market.

With all the acknowledgement from both academics and practitioners, however, there lies an inconsistency between these applications of the size effect. The usage of the SMB factor requires yearly rebalancing of the size portfolios, and a trading strategy related to firm size demands probably even more frequent position adjustments. However, the size premium added to a small firm’s cost-of-equity estimation is based

[3] Size is an important characteristic of these ETFs. However, it may not be the “only" characteristic. For example, the Vanguard Group, a U.S. investment management company, has three ETFs related to small-cap firms. Their exchange ticker symbols are VB, VBR, and VBK, which account for a total of $2.79 billion capital at the end of 2007. VBK is the combination of small-cap and growth stocks, while VBR is a small-cap and value stock ETF.

on the assumption that a firm will carry this extra premium in its discount factor moving forward for an extended period of time. Fama and French (2007) explain that the small stock premium comes from small firms gaining market capitalization and subsequently becoming bigger firms, but a firm's size behaves more like a long-lasting characteristic in the size premium application, which contradicts the empirical evidence. Although we do not know for certain which small firm will move to a bigger size group because of its own success, we do know that firms shift between different size groups in subsequent years after they were first assigned to a certain size rank. The size premium of a firm should be time-varying even if the CAPM beta of the size portfolio is time-invariant, so the cost of equity capital estimation could or should be adjusted accordingly if size has to be taken into consideration.

The existence of the size effect is not always perceived with full faith. This issue has to be addressed first, otherwise the debate of the application of the size premium will become a vain attempt. In the early 1980s when a fierce debate was conducted about the existence and the explanation of the size effect, Roll (1983) and Blume and Stambaugh (1983) both question the empirical importance of this phenomenon because the magnitude of the size effect is too sensitive to the technique used to evaluate the risk-adjusted return. Keim (1983) and Reinganum (1983) show that most of the risk-adjusted abnormal return to small firms occurs in the first two weeks in January, thus makes this effect easily exploited. It was the evaluation and the existence of the size premium being challenged, but the small stock premium was mostly untouched. Fiercer challenges came in the late 1990s, when Booth, Keim, and Ziemba (2000) argue that the January effect is not significantly different from zero in the returns to the DFA 9-10 portfolio over the period 1982-1995,[4] and Horowitz,

[4]The DFA 9-10 portfolio includes stocks with the lowest 20% market capitalization according to NYSE breakpoints.

Loughran, and Savin (2000b) also claim that the size effect ceases to exist after it was made well known because its benefit has already been exploited. Small firms do not have higher returns over big firms from the early 1980s to the mid-to-late 1990s, so the existence of the size effect is in doubt and deserves a thorough examination.

In this paper I will show that the size effect in the traditional definition is still intact given a longer sample period. The disappearance of the size effect in the 1980s and 1990s probably stems from a sample selection bias because the effect re-emerged in the late 1990s. I also examine whether this sample selection anomaly is a recurring scenario with a longer history of stock prices and find that the similar event occurred from the 1940s to 1960s.

However, an analysis of the evolution of the size premium will show that it is inappropriate to attach a fixed amount of premium to the cost of equity of a firm simply because of its current market capitalization. For a small stock portfolio which does not rebalance since the day it was constructed, its annual return and the size premium are all declining over years instead of staying at a relatively stable level. This confirms that a small firm should not be expected to have a higher size premium going forward sheerly because it is small now.

The paper proceeds as follows. Section 2 introduces the data used in this study. All NYSE, AMEX and NASDAQ operating firms are included and they are sorted by their respective market capitalization to form size portfolios. I also examine whether the size effect disappeared during the 1980s and 1990s and discuss its possible impact in this section. Section 3 offers a forward looking perspective of the size effect in response to the assumption of Fama and French (2007) that the small stock premium mainly resulted from firms moving between different size groups. We can also see the evolution of the size premium of the small stock portfolio and find evidence to conclude that a small firm does not always have a larger size premium simply because of its current size. Section 4 provides a method to separate the size premium into different regimes with macroeconomic variables, which shows that it is also very difficult to estimate the size premium with a time-varying estimation. Section 5 offers concluding remarks.

2 Data Description and the Evidence of the Existence of the Size Effect

2.1 Data Description

Monthly stock return data used in this research are collected from the University of Chicago Center for Research in Security Prices (CRSP) database. All NYSE, AMEX and NASDAQ operating firms are included when they are available on the CRSP tape.[5] Unlike Fama and French (1992), this study does not exclude financial firms from the sample because financial leverage is not in discussion. Since the market capitalization of a firm is the only firm characteristic covered in this paper and I also do not incorporate the Compustat database for the book equity data of companies, the number of firms each year is also greater than research considering both size and book-to-market equity characteristics. This choice of sample also prevents the potential survival bias generated by the Compustat database, please see the discussion in Kothari, Shanken, and Sloan (1995). The sample period is from December 1925 to December 2008.

The market portfolio return used in this paper is the CRSP value-weighted return on all NYSE, AMEX, and NASDAQ stocks, and the risk free rate is the total return on 30-day Treasury bill calculated by Ibbotson Associates.

To sort firms into different deciles according to their relative size, I follow the Fama and French (1992, 1993) tradition to use a firm's market equity at the end of June each year as the measure of its size. A firm has to be on the CRSP tape in

[5] American Depository Receipts, closed-end funds, Real Estate Investment Trusts, and companies incorporated outside the U.S. are excluded, which means only firms with CRSP share code 12 or less are included in this research.

June of year t to be included in a size portfolio from July of year t to June of year $t+1$ and years after that.[6] All NYSE listed firms are ranked each year according to their June market value, then these firms are allocated equally into 10 size portfolios on the basis of their relative size, so each portfolio has the same number of NYSE firms. The breakpoints between size portfolios are extracted from these NYSE firms, and AMEX and NASDAQ firms are inserted into these portfolios according to their market capitalization relative to the portfolio breakpoints. The first decile (portfolio 1) contains the smallest firms and the 10th decile (portfolio 10) includes the largest firms. In December 2008, Portfolio 1 has 1,895 firms and portfolio 10 has 158.

2.2 Does the Size Effect Still Exist?

In response to the question raised by Horowitz, Loughran, and Savin (2000b) about whether the size effect still exists, some basic statistics are presented in Table 1 to show that the effect did disappear during the 1980s and the early 1990s, but it was intact in most of the other sample periods. The statistics from the full sample are shown in Panel A. They are consistent with early findings on the size effect: big firms report lower returns than small firms, and the CAPM beta is also negatively related to size. The size premiums in the last row of each panel are calculated as follows:

$$
\begin{aligned}
SP_{i,t} &= R_{i,t} - \left(R_{f,t} + \beta_i (R_{m,t} - R_{f,t})\right), \text{ and} \\
SP_i &= \frac{1}{T}\sum_{t=1}^{T} SP_{i,t} \qquad i = 1, \ldots 10. \qquad (1)
\end{aligned}
$$

[6] Instead of the usual one-year holding period immediately following the size sorting date, I also extend the holding period to longer time spans to see how persistent the size premium is for the same group of firms.

where SP_i represents the average size premium of portfolio i which is shown in the table, $R_{i,t}$ and $R_{m,t}$ are monthly returns on size portfolio i and the market portfolio, respectively. R_f is the risk-free rate. β_i is the CAPM beta estimated by regressing $(R_i - R_f)$ on $(R_m - R_f)$ with the matching sample period. This size premium captures the part of the size portfolio return which cannot be explained by the CAPM. Practitioners usually add it to the cost-of-equity estimation of small-cap firms to compensate for their higher risks. Another way to estimate the size premium is through the estimation of the CAPM alpha. However, I will not adopt this approach because the sample period used by the regression to estimate CAPM coefficients and the one used by the realized return in equation (1) do not always match in this article.

[Insert Table 1 here.]

Panel B displays the statistics of the same variables with the sample period before June 1980, roughly when the size effect was made well known by academia. Although the statistics in the first two panels are not exactly the same, they look very much alike.

Panel C of Table 1 is consistent with the assertion of Horowitz, Loughran, and Savin (2000a) that there is no significant difference between the performance of different size portfolios during the period from 1980 to 1996.[7] The average returns on different size portfolios are no longer negatively related to their market capitalizations. From portfolio 1 to 4, the four smallest size portfolios, the average returns are increasing instead of moving in the opposite direction shown in the early years. The pattern of size premiums is also different from the ones shown in the previous two

[7]This period can be extended to 1998 and the results are still in the similar pattern to what one would get with sample period from 1980 to 1996, so this longer sub-sample period is chosen instead of the one used by Horowitz, Loughran, and Savin (2000a).

panels. For instance, portfolio 1 and 2 did not have the largest size premiums, they had biggest size “discounts" instead.

It is often suggested that pricing anomalies may disappear after they were made known to the public by researchers or financial practitioners if these anomalies were easily exploited. Horowitz, Loughran, and Savin (2000a) show that simply adding $0.125 to the December 31 price of small stocks can easily lower their average January returns from over 8% to -0.37% during the 1982-1997 span. Since Keim (1983) and Reinganum (1983) showed that most of the size premiums to small firms occurred during the first two weeks in January, it is no surprise that the January effect could be totally wiped out just by informed investors flocking into the market to buy small firm stocks in December, and so goes the size premium.

Sixteen years of time is not short, but the recent development shows that the result in Panel C is more likely to be an aberration from the formerly established rule than a new norm. Panel D presents the statistics from the past 10 years and shows that the negative relation between firm size and equity return has been restored, with only a few exceptions from some mid-cap size portfolios. The inconsistency of the mid-cap portfolios probably arises because the sample period is too short to offer a robust pattern between a firm’s size and its return. It has to be noted that the realized equity premium of the U.S. market during these 10 years is slightly below zero, which is significantly lower than the historical standard. This might contribute to the flat security market line, where the beta of size portfolios seems independent of their respective average return.

Another serious threat generated by the data from the 1980s and 1990s is that the return differential between small and big firm size portfolios, also known as SMB in the Fama-French 3-factor model, may have an insignificant or even a negative price of risk. This implies that the SMB factor is either meaningless or has a negative effect on the stock return. We can use a simple cross-sectional regression to show how and why this matters.

[Insert Table 2 here.]

Table 2 displays price-of-risk estimations of the popular Fama-French factors with different sample periods. Following the Fama and MacBeth (1973) procedures, in each sub-sample period I run time-series regressions of each test portfolio return in excess of the risk-free rate ($R^e_{it} = R_{it} - R_{ft}$) on the excess market return ($R^e_{mt} = R_{mt} - R_{ft}$), the returns on the small size portfolios minus the returns on the big size portfolio (SMB), and the differential between the returns on high and low book-to-market equity firms (HML).[8]

$$R^e_{it} = \alpha_i + \beta_i R^e_{mt} + s_i SMB_t + h_i HML_t + \varepsilon_{it} \qquad t = 1, 2, \ldots, T, \forall i. \tag{2}$$

The test portfolios include 5-by-5 portfolios formed on book-to-market equity and size, and 17 industry portfolios.[9] Since there are missing observations in the return series of the portfolio with the highest book-to-market equity and the largest size, it is taken out of the test portfolios. These portfolios are chosen because they cover different aspects of security characteristics.

The next step is to regress the expected returns of test portfolios from each sample period on their respective risk loading estimates from the time-series regression. I

[8]Please refer to Fama and French (1993) for the detailed definition of SMB and HML. Data on these two variables are obtained from Professor Kenneth French's website at Dartmouth University.

[9]All the portfolio data are also acquired from French's website.

take the average return of each portfolio from the corresponding sample period as their return expectation. The cross-sectional regression is:

$$E_T(R_i^e) = \beta_i \lambda_1 + s_i \lambda_2 + h_i \lambda_3 + \alpha_i, \qquad i = 1, 2, \ldots, N. \tag{3}$$

where λ_2 is the price of the risk represented by the size factor SMB. During the period from 1980 to 1998, the price of SMB is insignificantly different from zero and its magnitude is also comparably smaller than it is in the other sub-periods. The number is 0.29 before 1980 and 0.20 after 1998, but it is only 0.07 from July 1980 to June 1998. The other parameters do not change as dramatically over different sub-periods. The price of a risk factor being equal to zero discredits its explanatory power to the cross-sectional variability of returns, and this is exactly the case for the SMB factor from 1980 to 1998.

It may be too early to say that the explanatory power of the SMB factor fully recovers in the post-1996 or the post-1998 period, but it is clear that the zero or slightly negative SMB price during the 1980s and 1990s is not necessary a lasting problem.

2.3 Regime Shifts of the small stock premium

As mentioned earlier, the size premium and the small stock premium are related because the risk-adjusted abnormal return of small firms is an important part of the return differential between small and big stock portfolios. According to Table 1 Panel A, the small stock premium of portfolio 1 is 3.39%, which accounts for half of the return difference between portfolio 1 and 10. Since the size premium is highly dependent on the asset pricing model and the sample period it is using, I will focus on the possible structural change or regime shift of the small stock premium in this section first.

Although the differential between the returns on size portfolio 1 and portfolio 10 is different from the definition of the SMB factor in the Fama and French 3-factor model, I will borrow this acronym to represent the small stock premium for the following discussion. Motivated by the earlier discussion of the disappearance of the small stock premium in the 1980s and 1990s and the reappearance in the following years, I believe that there may exist structural changes or regime shifts of the expected mean of SMB. Panel A of Figure 1 exhibits the annual return differential between portfolio 1 and portfolio 10, in which we see annual SMB alternates between high and low values but certain persistency exists. From 1984 to 1998, the supposedly positive SMB is negative in most years except in 1988 and 1991 to 1993. The sample average of the equity risk premium during these 15 years is 10.53%, which is well above the historical average. Big firms performed exceptionally well while small firms did not during this period, so the disappearance of SMB should certainly came from the size premium, or lack thereof.

[Insert Figure 1 here.]

Assuming that the expected mean and variance of SMB can be expressed by a two state Markov-switching model, so the state variable S_t, which governs the regime shift, takes a value of 1 or 2. When $S_t = 1$, the expected mean of SMB_t is in the state of a low value, while $S_t = 2$ represents the state when the expected mean of SMB_t is high.

$$y_t = \mu_k + \sigma_k \varepsilon_t \qquad \varepsilon_t \sim N(0,1). \tag{4}$$

where y_t represents SMB_t, μ_k and σ_k are state-dependent mean and standard deviation of SMB_t. k=1 or 2, which identifies the state SMB_t is in at time t.

The state variable S_t is assumed to follow a 2-state first-order Markov process with fixed transition probabilities as follows:

$$\begin{aligned} p &= \Pr(S_t = 1|S_{t-1} = 1) \\ 1-p &= \Pr(S_t = 2|S_{t-1} = 1) \\ q &= \Pr(S_t = 2|S_{t-1} = 2) \\ 1-q &= \Pr(S_t = 1|S_{t-1} = 2) \end{aligned} \tag{5}$$

The mean and variance of SMB are determined by the current state, and the state variable S_t is not dependent on the past information beyond one period.

SMB_t under each state is assumed to follow the normal distribution and the parameters of the distribution function are only contingent on the state k, so

$$f(y_t|S_t = k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(\frac{-(y_t - \mu_k)^2}{2\sigma_k^2}\right) \tag{6}$$

for $k = 1,2$. The log-likelihood function is

$$\ln \mathscr{L}(y_1, y_2, \ldots, y_T; \theta) = \sum_{t=1}^{T} \ln[\Pr(S_t = 1) f(y_t|S_t = 1) + \Pr(S_t = 2) f(y_t|S_t = 2)] \tag{7}$$

and the regime probability $\Pr(S_t = k)$ can be estimated with the following recursive representation proposed by Gray (1996):

$$\Pr(S_t = 1) \quad = \quad (1-q)\left[\frac{f(y_{t-1}|S_{t-1} = 2)\Pr(S_{t-1} = 2)}{f(y_{t-1}|S_{t-1} = 1)\Pr(S_{t-1} = 1) + f(y_{t-1}|S_{t-1} = 2)\Pr(S_{t-1} = 2)}\right]$$

$$+p\left[\frac{f(y_{t-1}|S_{t-1}=1)\Pr(S_{t-1}=1)}{f(y_{t-1}|S_{t-1}=1)\Pr(S_{t-1}=1)+f(y_{t-1}|S_{t-1}=2)\Pr(S_{t-1}=2)}\right] \quad (8)$$

where the lowercase p and q are the transition probabilities defined in equation (5) and $\Pr(S_t = 2) = 1 - \Pr(S_t = 1)$.

Table 3 presents the estimation results of the above Markov-switching model along with an unconditional normal distribution model as its comparison. The sample period is from July 1940 to December 2008 instead of starting from July 1926 because it has to be trimmed short in the following sections to accommodate the portfolio positions with longer holding periods. According to the log-likelihood values, AIC, and BIC statistics of these two models, the Markov-switching model fits the sample better than the model with the assumption that SMB follows an unconditional normal distribution. The expected mean of the low SMB state is insignificantly different from zero, which explains why SMB can disappear over an extended period. The average annualized returns under two different states are -2.67% and 44.97%.

[Insert Table 3 here.]

Panel B of Figure 1 displays the smoothed probability in state 2 (high SMB state). Table 3 also shows the transition probabilities p and q, which are 0.9579 and 0.8090, respectively. These results imply that the low SMB regime is more persistent than the high SMB regime. On average the high SMB regime lasts for 5.2 months, and the low SMB regime keeps at the same state for 23.8 months. If the true data generating process of SMB follows the description of this Markov-switching model, it is no surprise that the small stock premium could disappear over a long period during the 1980s and most of the 1990s then resurfaces in recent years.

From Figure 1 we can also see that SMB is persistently low from 1946 to 1963, which indicates that the experience from the 1980s and 90s indeed has a predecessor. Repeat the same exercise done in Table 1 for this period, we can find that portfolio 1 has an average size premium at -1.77% per annum, while portfolio 10 has a slightly positive 0.42% average size premium. The average of SMB from 1946 to 1963 is -0.74%, which mostly stems from the low size premium of small stocks instead of the difference between their respective CAPM projections.[10] These results show that the temporary disappearance of the size effect is a recurring event. However, when we look at a longer time span, the small stock premium could still hold true at least on average.

[10]CAPM beta is still negatively related to firm size during this period, but the slope of the security market line calculated with returns on size portfolios and their respective betas is smaller than it is calculated with the full sample.

3 Size as a Genetic Code or a Short-Lived Characteristic?

If the size premium ceases to exist like Horowitz, Loughran, and Savin (2000b) assert, or its magnitude has no relation to firm size, there is no need to give a "premium" to a small firm when estimating its cost of equity capital. In fact, given what we see in Panel C of Table 1 we might have to give small-cap firms a discount if the negative size premium of portfolio 1 remains. The data from the last 10 years seem to restore the order of the size premium and the necessity to add it to small firms, but I will show in this section that it still remains to be proved whether a small-cap firm should require this size premium in its cost-of-equity estimation.

3.1 Design of the $t+j$ Portfolio

Fama and French (2007) find that the return differential between small and big firms is mainly driven by small-cap firms moving up the size rank to become large-cap firms. This perspective changes the assumption of the size premium a small firm should get in the long run. The logic is simple: a small firm becomes a big firm because its market capitalization increases faster than its peer, which usually results from its fast growing price. However, small firms cannot keep the higher average return of old once they become big firms, otherwise the small stock premium will turn into a big stock premium. Although this is mainly an explanation of the small stock premium instead of the size premium, the discussion in the previous section shows that these two premiums are related.

Since the Fama-French size portfolios are constructed in each June and are held for a whole year until they are rebalanced in June next year, their finding implies that some firms are likely to switch to different size groups sooner than a year, especially for the small firms to become big firms. The usual practice of the size premium estimation is to calculate it with annually rebalanced size portfolios,[11] then we add this number to a firm's cost of equity for the following years to discount its future cash flows to the present value. We know this is probably a proper assessment of the discount factor for the first year, but is it still proper if an originally small firm becomes a big firm from the second year on and does not warrant such a premium hereafter?

To investigate whether the size premium is changing over time and how it evolves, I design the following $t+j$ size portfolio approach. In the traditional size portfolio formation, securities are assigned to each portfolio in June and the portfolios are held from July to June next year under a buy-and-hold strategy. In the $t+j$ size portfolio approach I also choose to sort securities in June of each year t, but instead of holding the portfolios for the following year, I also look at the monthly returns for an one-year holding period from July of year $t+j$-1 to June of year $t+j$, where $j = 2, \ldots, 15$.[12] All the firms are identified and tracked by their CRSP permanent number. If a firm goes bankrupt or is merged by another firm in the following years, then it is taken out of the portfolio once it is off the CRSP tape. Otherwise it keeps in the same $t+j$ size portfolio as assigned in the initial sorting date no matter how big or how small its market capitalization becomes.

[11]For getting the size premium estimation, some practitioners rebalance the size portfolios more frequently. For example, Ibbotson Associates sorts and assigns all eligible companies to different size portfolios with the closing price and shares outstanding data for the last trading day of March, June, September and December instead of June each year.

[12]This approach reduces to the traditional size portfolio formation when $j = 1$.

For example, the firms in t+2 portfolios from July 1989 to June 1990 were sorted and assigned to different size portfolios in June 1988; the same composition of firms is used in t+1 portfolios from July 1988 to June 1989, which are 12 months immediately after the sorting date. The t+3 portfolios in July 1990 also consist of the same firms, except for those were delisted during the first two years. There is also another set of t+2 portfolios from July 1988 to June 1989, each consists firms sorted by their June 1987 size. We can string together all the t+2 portfolios to see how firms perform a year after its original sorting date for a whole year. The same process is done for all $t+j$ size portfolios. This approach allows us to follow the average performance of firms j years after they were assigned to a specific size group.

If a firm's size behaves as a characteristic and this attribute follows the firm for an extended period of time, return patterns among different $t+j$ size portfolios should not change much for different j. On the other hand, if a small firm deserves a lower size premium after it becomes a bigger firm, the size premium in the following years will decrease accordingly. By tracking the historical performance of firms sorted by size, we can get a better idea on how the size premium of a firm behaves and whether it is a good indicator of an extra risk source.

3.2 Size Premium is Changing Over Time

Practitioners usually consider a fixed size premium for a firm for subsequent years, which implies that either firms will not migrate to other size groups, or they will still demand the same size premium even after they switch to different size groups. To make a valid comparison between different $t+j$ portfolios, I change the starting date of all portfolios from July 1926 to July 1940 to accommodate the t+15 portfolios, which have companies being sorted in June 1926 but will not report the first return observation until July 1940.[13]

Table 4 presents the average size premiums of different $t{+}j$ size portfolios in reference to the respective CAPM projected returns on the traditional size portfolios. The "traditional" size portfolio means that firms are sorted and assigned to different size portfolios according to their June market capitalization, and the portfolios are held from July of the same year to June next year. The definition of the average size premium of a $t{+}j$ size portfolio is

$$
\begin{aligned}
SP_{i,t}^{t+j} &= R_{i,t}^{t+j} - \big(R_{f,t} + \beta_i (R_{m,t} - R_{f,t})\big), \text{ and} \\
SP_{i}^{t+j} &= \frac{1}{T}\sum_{t=1}^{T} SP_{i,t}^{t+j}, \qquad (9)
\end{aligned}
$$

where $R_{i,t}^{t+j}$ represents the time t return on the $t{+}j$ portfolio of firms in the ith size group, and β_i is the same as in equation (1).

[Insert Table 4 here.]

The first decile size portfolio, which contains firms with the lowest market capitalizations among all listed firms on the sorting date, usually has a large and significant CAPM alpha and a beta too low to project the realized return. Table 1 shows that portfolio 1 has a size premium of 3.39% per annum with the sample period from July 1926 to December 2008. The corresponding number in Table 4 is the average size premium of the $t{+}1$ portfolio for portfolio 1. Although the benchmark is still calculated with the same beta, it drops to 1.49% because the sample period here does not start until July 1940. The difference reflects a large historical size premium for the

[13] The security return data on CRSP tape start from December 1925, so June 1926 becomes the first available sorting date.

small firms from 1926 to 1940. The premiums change a lot with different sample periods, but the pattern is nevertheless revealing. The smallest firms still get a bigger size premium, while the biggest firms even get a size discount.

If firms are supposed to be awarded a fixed size premium for years, we should see the numbers in Table 4 remain stable over different $t+j$ portfolios within each size group. The result is apparently contrary to this hypothesis. The size premium of portfolio 1 drops dramatically two years after the initial sorting date and becomes insignificantly different from zero in the third year. After that the small firms get a discount and such a discount gradually becomes significantly different from zero. On the other hand, portfolio 10 sees its size premium going up from the negative value in the first two years to a positive but insignificant number for the most part of the following eight years. Most of the size portfolios have a declining size premium after the sorting date except for portfolio 10, which reflects the fact that returns on different size portfolios tend to converge to the same number over years. Table 5 shows that the difference in average returns on different size portfolios gradually becomes insignificant as sorting dates pass by.

[Insert Table 5 here.]

If history can be any guide to the future performance, we are likely to overestimate the cost of equity capital of small firms and under-estimate the cost of equity of big firms by the current treatment of the size premium.

3.3 Robustness Check

We have seen in Table 1 that the historical averages of both the size premium and the small stock premium are sensitive to the choice of the sample period, but the pattern remains unchanged if given a long enough horizon. Here I will verify that the findings in this section are not sensitive to different breakpoints of size groups.

Fama and French (2007) divide firms into two groups in terms of size to explain the cause of the Fama-French *SMB* factor, so I also divide all the acting firms into two groups according to the NYSE median market-cap breakpoint in each June.

For better examining the relation between firm size and the corresponding return performance, I also rank firms according to their size each June and form three portfolios with firms of their size in the bottom 30%, middle 40%, and top 30% (S-30%, M-40% and B-30% hereafter) by the NYSE market-cap breakpoints.

The size premiums calculated with new breakpoints are displayed in Table 6. The big size portfolios (Big or B-30%) all have very small and insignificant size premiums like the size premium of portfolio 10 reported in Table 4. Please be noted that I still use the traditional size portfolio approach (it is equivalent to the t+1 portfolio here) with the new breakpoints and the sample period from 1926 to 2008 to estimate CAPM betas. The size premiums of "Small" and "S-30%" size portfolios are significant through t+1 to t+4 or t+5 portfolios, respectively, and they are also declining as j goes up. Ten or seven years after the initial sorting dates, these two small size portfolios even have a discount. These characteristics are all consistent with the pattern shown in portfolio 1 in Table 4.

[Insert Table 6 here.]

Comparing Table 6 to Table 4, it is apparent that the size premium for small stocks in the traditional sense does exist no matter how many size groups the stocks are divided into, but it fades out gradually if the same composition of firms is held longer than a year.[14]

If a group of firms have the same stream of expected future cash flows, it is possible that the firm with a higher risk is going to be priced lower. Such a firm may end up having a higher return because it is more likely to have a higher dividend yield. However, small firms do not only gather higher returns through higher dividend yields, they usually have higher capital appreciation rates too. Fama and French (2007) explain that migration of stocks across size groups is the cause of the small stock premium.[15] Once a small firm's market capitalization increases and it is qualified as a big firm, a size premium should not apply anymore. According to Table 4 and 6, small firms did have higher size premiums when they were first assigned to the small size portfolio, but this effect does not persist. A firm which belongs to portfolio 1 sees its size premium turns into a discount after a few years if it is still expected to be compensated as a small stock. It is probably reasonable for a small firm to get a larger discount factor than the CAPM suggests because it bears higher risks than the model can explain for the time being, but the usual practice could very likely over-compensate the risks a small firm is bearing.

If the size effect has to be considered in the cost-of-equity estimation, we should search for the root of this short-lived premium and identify the risk source it represents. This is just as important as how much it is, if not more important.

[14]The small stock premium fades away until it is barely noticeable. However, the size premium for small stocks sometimes becomes a size discount if the same composition of stocks is held for a few years.

[15]In their article Fama and French use "size premium" to refer to the fact that small-cap firms have higher returns than big-cap firms without risk adjustment, which is equivalent to the "small stock premium" used in this paper. As shown earlier that these two premiums are related.

4 Size Premium under Different Economic Situations

Section 3 shows that a small firm can have a higher size premium only in the short run. Over a longer time span, a firm's size and even its sensitivity to risk are all subject to change, and its size premium changes accordingly.[16] In light of these results, I propose not to include a fixed size premium in the long-term cost-of-equity estimation. However, the size premium, no matter how short-lived it is, still appears to exist in the first few years for small firms. Take the popular discounted cash flow method as an example, the first few years matter the most if given a steady stream of future cash flows. By excluding the size premium from the cost-of-equity estimation, one might argue that we are also likely to understate the risk a small firm is taking.

The simplest way to resolve this conundrum seems to apply a time-varying cost of equity by adding different size premiums to the estimation according to the results in Table 4. The short-term size effect is thus accounted for, and the long-term size premium is also no longer permanent. However, Table 4 only displays the standard deviation of the average of the size premium, the variation of the annual size premium per se is much larger. If the size premium swings between high and low levels like the two-regime small stock premium model shown in section 2.3, adding an average size premium into the short-term cost-of-equity estimation may not help the matter. We could easily over-estimate the cost of equity of small firms in one period and suppress their value, while under-estimate the cost of equity in another period

[16]CAPM betas of all size groups are monotonically decreasing from $t+1$ through $t+15$ portfolios. These results are not shown in the tables, but they are available upon request. In this paper I use the traditional size portfolios with the full sample (July 1926 to December 2008) to estimate CAPM betas to get a consistent benchmark in all cases but ones in Table 1.

and bring the price to an un-deserving high level. In this section I will examine the likelihood of this scenario.

The concept of connecting financial distress to firm size has been discussed in the asset pricing literature to explain the anomalous cross-sectional pattern of stock returns. Queen and Roll (1987) find that a firm's unfavorable mortality rate is a decreasing function of its size, and Campbell, Hilscher, and Szilagyi (2008) further show that size has a negative relation with the excess return between safe and distress stocks. I will examine from a different angle to see whether economic distress has an effect on the size premiums.

I divide the sample period into several two-regime scenarios according to different macroeconomic variables related to distress and calculate the size effect under each regime. There are two reasons for this experiment: the first is that only the systematic risk should be taken into account when pricing a firm or an asset. If small firms are supposed to be awarded a higher premium sheerly because of their failure risk, then we should be able to distinguish different patterns of their size premium under different economic situations. Second, in light of the success of a simple Markov-switching model used on the small stock premium in section 2, it is natural to try a two-regime model on the size premium as well. However, the estimation of the size premium is highly contingent on the choice of the asset pricing model and the sample period, so I do not investigate the possible regime shifts of the size premium directly. Instead, I will try to explore the relation between the size premium and three different candidates of macroeconomic variables. If the size premium is at least partly driven by systematic risk sources, its magnitude should vary as the economic environment changes.

4.1 Identifying the States of Economy

The first state variable is an indicator variable which identifies the economic status during a business cycle: a dummy variable which equals 1 for months in the expansion period and 0 for months in the contraction period.[17] When in distress, smaller firms usually get hit harder because they have thinner cushion in common equity and their ability to raise capital via new debts, bank loans, or even government bailouts is also poorer than big firms. On the other hand, small firms which survive the storm can often see a sudden boom in their stock returns, as were evidenced by their bigger beta.[18] Whether the bigger volatility in the stock return for the small stock portfolio can translate to separate size premiums is the focus of the investigation. According to NBER's Business Cycle Dating Committee, there are 14 business cycles since 1926 to date with the shortest contraction period being 6 months and the shortest expansion period being 24 months.

The second indicator is the market trend, which is similar to the idea of the business cycle. I distinguish the bull and bear markets by a Markov-switching model on the CRSP value-weighted market portfolio return with the similar procedure laid

[17]NBER's Business Cycle Dating Committee publishes the U.S. business cycle peak and trough months on the NBER website. Their latest announcement on 12/01/2008 declares that the previous expansion period peaked in December 2007 and a recession soon followed. The conclusion of the current recession has not yet been determined as the writing of this paper. I assume all of year 2008 fell into the contraction period to make the sample period consistent with other state variables.

[18]Fama and French (1993) point out that small firms do not participate in the economic boom of the middle and late 1980s for an unknown reason. This finding is consistent with the argument of the disappearance of the size effect in the 1980s and 1990s. Indeed, the small stock premium was -10.4% per annum from December 1982 to July 1990, the expansion period right after the longest recession since the Great Depression. However, small firms greatly outperform big firms during the economic booms after the Great Depression or the recession caused by 1973 oil crisis, with average small stock premiums at 55.9% and 23.1%, respectively. It is probably premature to judge the experience in the 1980s as a new norm or just an anomaly. Nonetheless, the magnitude of SMB during the expansion periods in the middle 1930s and the late 1980s could counter the argument raised by Fama and French (1993).

out in section 2.3.[19] Regime 1 represents the state of the bear market with a lower mean return and higher volatility; regime 2 indicates the bull market with a higher mean return and lower volatility. An indicator variable is used to represent the bull market with its value being equal to 1 when the regime 2 smoothed inference of the month is greater than 0.5, and 0 otherwise. The reason to use a dummy to identify the market trend instead of the realized market return is to filter out noise. When we apply the size premium on the cost of equity capital estimation, we look for the long-term performance instead of the short-term disturbance. Looking too much into the day-to-day or month-to-month performance will mix up true trend and noise. For instance, even during the huge market downturn in the Great Depression, when the Dow Jones Industrial Average (DJIA) dropped from then historical high of 381.17 on 9/3/1929 to the following lowest point of 41.22 on 7/8/1932, we can still see the market posted double digit gains on return during the process. In February and June 1931, the monthly returns derived from the DJIA were 12.40% and 16.90%, respectively. These were great rallies even in any bull market, but they still cannot stop the free fall of the stock market and the investment environment would not be changed simply because of a sudden spark of life. Since the cost of equity capital and the size premium are all about the long term prospect of the firm, it is more fitting to examine the general market trend in this simple fashion.

The third indicator is the credit spread between AAA and BAA corporate bond rates. The data are obtained from the Federal Reserve Bank of St. Louis website. Although we cannot link a firm's size directly to its credit rating, large firms usually get better ratings and lower borrowing rates.[20] When there is abundant credit

[19] There is no consensus on the definition of bear or bull markets other than a general description. Here I adopt the market trend definition of the model 1 in Chen (2009).

[20] According to the summary statistics provided by Altman and Rijken (2004), firm's credit rating is negatively related to the market value of equity. I also compare the average market values between

floating in the market, the credit spread tends to narrow down because banks and funds compete against each other for an investment opportunity without thinking too much about the risk. This process will eventually drive the spread down. On the other hand, the credit spread increases when the credit market is in a dire condition and investors take default risks more seriously. Every banker will think twice before lending money out. When the credit spread is high, it is more likely that small firms endure a higher borrowing cost than big firms, therefore their failure risk induced by the poorer credit rating is also higher. I continue to apply the same technique previously used in the market trend indicator to separate the credit spreads into two different states, and then convert the smoothed inference into a dummy variable using the 0.50 threshold.

The transition probabilities of staying in the same state for the Markov-switching model of the market trend are 0.892 (bear market) and 0.963 (bull market); they are 0.987 (low credit spread) and 0.974 (high credit spread) for the credit spread. The common feature of these macroeconomic variables is that the states defined by them are all very persistent, so we can link these variables with the shift of the size premium over a longer span instead of the month-by-month movement. Once the state variable of the market trend shifts to the bull market state, it would stay put for 27 months on average, and a credit spread dummy remains in the state of a lower mean value for 78 months.

[Insert Figure 2 here.]

firms with investment grade ratings and with non-investment grade ratings over the past 15 years. The average size of firms with better credit is 9 to 10 times bigger than the size of poorer rating firms. The sample includes all firms in the Compustat database from 1994 to 2008.

Figure 2 illustrates three different dummy variables on the right-hand side and their original data on the left.[21] It has to be noted that these state variables are all asymmetrical. We see expansion periods more often than contraction periods, longer bull markets than bear markets, and more days with low credit spreads than days with high ones. Over the total 822 observations, there are 698 months identified as in the expansion period, 646 months in the bull market, and 552 months in the low credit spread regime.

4.2 The Size Premium under Different Economic Environments

These state variables do not highly coincide with each other, but they are all capable of separating the size premium of small stocks under different states. I also use the $t+j$ portfolio approach to see whether these states can identify the size effect of stocks over the long run. Table 7 and 8 present the size premiums of the first and the 10th size portfolios under different economic situations.

[Insert Table 7 here.]

[Insert Table 8 here.]

The first column of Table 7 or 8 shows the same average size premiums as the corresponding column in Table 4. Through the second column to the last, the average size premiums under different states of the same macroeconomic variable are paired with each other. The second and third columns are the average size premiums in the expansion or contraction state identified by the business cycle dummy; the fourth and fifth columns show the averages during bull or bear markets from the market

[21] I use the GDP growth rate for the business cycle dummy as its "original data". However, it is well known that the Business Cycle Dating Committee of the NBER does not determine the peaks and troughs by the GDP data alone.

trend dummy; and the last two columns are average size premiums in the high or low state of the credit spread dummy.

The last row of each table shows the number of observations in a specific state. These three dummy variables post asymmetric states as earlier mentioned, but the credit spread dummy is significantly different from the others because the state brings the higher average returns has a lot less observations than the state brings the higher return for the other two dummy variables.[22]

Small stocks usually have a high and significant size premium, and this premium is even more pronounced in the expansion period or the high credit spread period, and interestingly, during the bear market. Portfolio 1 has a positive premium for most of the $t+j$ portfolios during the market downturn because the market trend dummy successfully identifies the low return period of the market, which in turn drives the benchmark even lower than the drop of the realized return on small stocks. The time series dynamics of the size premium revealed by the $t+j$ portfolio approach present a different scenario for the business cycle dummy. It is indecisive whether a small firm has a greater size premium during the expansion or contraction period.

Table 8 displays the size premium, or more precisely, the size discount of portfolio 10. Large firms usually can be explained well by the CAPM or other asset pricing models, so the common practice does not require a size premium on them. Even under different states, the size premiums are still small in magnitude comparing to the corresponding statistics of portfolio 1. If we focus on the first few $t+j$ portfolios, the business cycle does not seem to play an important role. The average size premi-

[22]The state generates the higher average return does not necessarily have the higher size premium. The latter also depends on the sensitivity to the market risk and the market return under this “unfavorable" state.

ums under different regimes of the market trends or credit spreads are much more different, but they are still not as pronounced as their counterparts in portfolio 1.

A one-sided t test on unequal sized variables is also applied here to compare the difference between average size premiums under different economic states. The size premiums in Table 7 and 8 are shown in **boldface fonts** if the difference is significant at the 10 percent level. We cannot reject the null hypothesis that none of the size premium pairs of portfolio 1 or 10 are significantly different during different periods of business cycles. The same test for different market trends shows the similar result for the first nine years for portfolio 1 and the first two years for portfolio 10. The state variable derived from the credit spread data is the most successful of all. The difference of the average size premiums of $t+j$ portfolios is significant at 10 percent level for most of the cases for portfolio 1, and it is also significant for the first 6 years for portfolio 10.

The size premium a small firm should demand for bearing higher risks is limited only in the first few years and its magnitude is difficult to predict. The empirical results imply that we should be very careful to identify the risks a firm is bearing instead of taking it only by the firm's current size. If there are other systematic risks which is related to size, we should reconsider whether that is the cause of a firm being riskier than the others and assign the specific risk premium to it accordingly.

5 Conclusion

This study verifies the existence of the size effect of annually rebalanced size portfolios with a longer sample period, but suggests not to include the size premium in the cost-of-equity estimation of small firms because this effect is only short-lived.

The assertion of the disappearance of the size effect in the 1980s and 90s was just a result of sample selection. Similar events of temporary disappearance of the size effect from different periods were found but they have never been proved permanent. Suffice it to say that the size effect did not simply disappear because it was revealed by academics and exploited by practitioners. It is shown in section 2 that the small stock premium can be better captured by a two-state Markov-switching model rather than the usual stationary normal distribution assumption. This empirical evidence is consistent with the story of the temporary disappearance of the size effect in the 1980s and 1990s.

Using the $t+j$ portfolio approach designed for this study, I demonstrate that the small stock premium declines if we hold the size portfolio longer than the usual one-year holding period rule. This can be considered as evidence of Fama and French (2007)'s finding that the size premium stems from small firms moving up the size rank to become big firms. Since firms move between size groups, the size premium should not be considered as a constant and it has to reflect the new size group they are currently in. The popular perception of a fixed size premium used by practitioners in the cost-of-equity estimation is obviously mistaken. I track the size premiums of different size portfolios for the subsequent 15 years after their formation date and find that most of the premiums converge toward zero, so firms should not be awarded a size premium for a long-term estimation.

If the size premium of a firm is estimated with the assumption that a firm moves from one size group to another all the time, it should be time-varying as well. The average size premium of portfolio 1, which includes all NYSE, NASDAQ and AMEX firms with market capitalization less than the first decile market-cap breakpoint of all NYSE listed firms, is 1.49% for the first year after its creation for the past 68 years. The same composition of firms still merit an average of 1.02% premium in the following year, but it declines rapidly after that. Adding a fixed size premium according to a firm's current size could very well overstate the relation between a firm's size and the risk it is bearing.

Certain macroeconomic variables can help us to distinguish the possible regimes of the size premium. These variables include the business cycle, the market trend, and the credit spread. However, the decision to distinguish the size premium of a firm under the assumption of one specific state is very difficult to make given how highly volatile the monthly size premium is. Adding a naive size premium to a firm's cost of equity capital estimation still potentially introduces more errors no matter this size premium is fixed or time-varying.

References

Altman, Edward I., and Herbert A. Rijken, 2004, How Rating Agencies Achieve Rating Stability, *Journal of Banking and Finance* 28 11, 2679–2714.

Banz, Rolf W., 1981, The Relationship between Return and Market Value of Common Stocks, *Journal of Financial Economics* 9, 3–18.

Berk, Jonathan B., 1995, A Critique of Size-Related Anomalies, *Review of Financial Studies* 8, 275–86.

Black, Fischer, 1972, Capital Market Equilibrium with Restricted Borrowing, *Journal of Business* 45, 444–455.

Blume, Marshall E., and Robert F. Stambaugh, 1983, Biases in Computed Returns: An Application to the Size Effect, *Journal of Financial Economics* 12 3, 387–404.

Booth, David G., Donald B. Keim, and William T. Ziemba, 2000, Is There Still a January Effect?, in *Security Market Imperfections in Worldwide Equity Markets* (Cambridge University Press, Publications of the Newton Institute. Cambridge; New York and Melbourne).

Campbell, John Y., Jens Hilscher, and Jan Szilagyi, 2008, In Search of Distress Risk, *Journal of Finance* 63 6, 2899–2939.

Chen, Shiu-Sheng, 2009, Predicting the Bear Stock Market: Macroeconomic Variables as Leading Indicators, *Journal of Banking and Finance* 33 2, 211–23.

Cochrane, John H., 2005, *Asset pricing*. (Princeton University Press Revised Edition. Princeton and Oxford).

Fama, Eugene F., and Kenneth R. French, 1992, The Cross-Section of Expected Stock Returns, *Journal of Finance* 47, 427–65.

Fama, Eugene F., and Kenneth R. French, 1993, Common Risk Factors in the Returns on Stock and Bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 1995, Size and Book-to-Market Factors in Earnings and Returns, *Journal of Finance* 50, 131–55.

Fama, Eugene F., and Kenneth R. French, 2007, Migration, *Financial Analysts Journal* 63, 48–58.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, Return, and Equilibrium: Empirical Tests, *Journal of Political Economy* 81, 607–636.

Gray, Stephen F., 1996, Modeling the Conditional Distribution of Interest Rates as a Regime-Switching Process, *Journal of Financial Economics* 42, 27–62.

Horowitz, Joel L., Tim Loughran, and N. E. Savin, 2000a, The Disappearing Size Effect, *Research in Economics* 54 1, 83–100.

Horowitz, Joel L., Tim Loughran, and N. E. Savin, 2000b, Three Analyses of the Firm Size Premium, *Journal of Empirical Finance* 7 2, 143–53.

Keim, Donald B., 1983, Size-Related Anomalies and Stock Return Seasonality: Further Empirical Evidence, *Journal of Financial Economics* 12 1, 13–32.

Kothari, S. P., Jay Shanken, and Richard G. Sloan, 1995, Another Look at the Cross-Section of Expected Stock Returns, *Journal of Finance* 50, 185–224.

Lintner, John, 1965, The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *The Review of Economics and Statistics* 47, 13–37.

Pratt, Shannon P., and Roger J. Grabowski, 2008, *Cost of Capital: Applications and Examples*. (Wiley) 3 edn.

Queen, Maggie, and Richard Roll, 1987, Firm Mortality: Using Market Indicators to Predict Survival, *Financial Analysts Journal* 43, 9.

Reinganum, Marc R., 1981, Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings' Yields and Market Values, *Journal of Financial Economics* 9 1, 19–46.

Reinganum, Marc R., 1983, The Anomalous Stock Market Behavior of Small Firms in January: Empirical Tests for Tax-Loss Selling Effects, *Journal of Financial Economics* 12 1, 89–104.

Roll, Richard, 1983, On Computing Mean Returns and the Small Firm Premium, *Journal of Financial Economics* 12 3, 371–86.

Sharpe, William F., 1964, Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *The Journal of Finance* 19, 425–442.

Figure 1: The return difference between the first and the 10th decile size portfolios and the smoothed probability of the high small stock premium regime. Panel A shows the annual portfolio return difference between small and big stocks. It is apparent that big firms outperform small firms most of the time from the mid-1980s to late 1990s. This account for the “disappearance" of the size effect in that time span. Similar situation also happened in the 1950s and late 1960s to early 1970s. The smoothed inference of the high SMB regime is shown in Panel B.

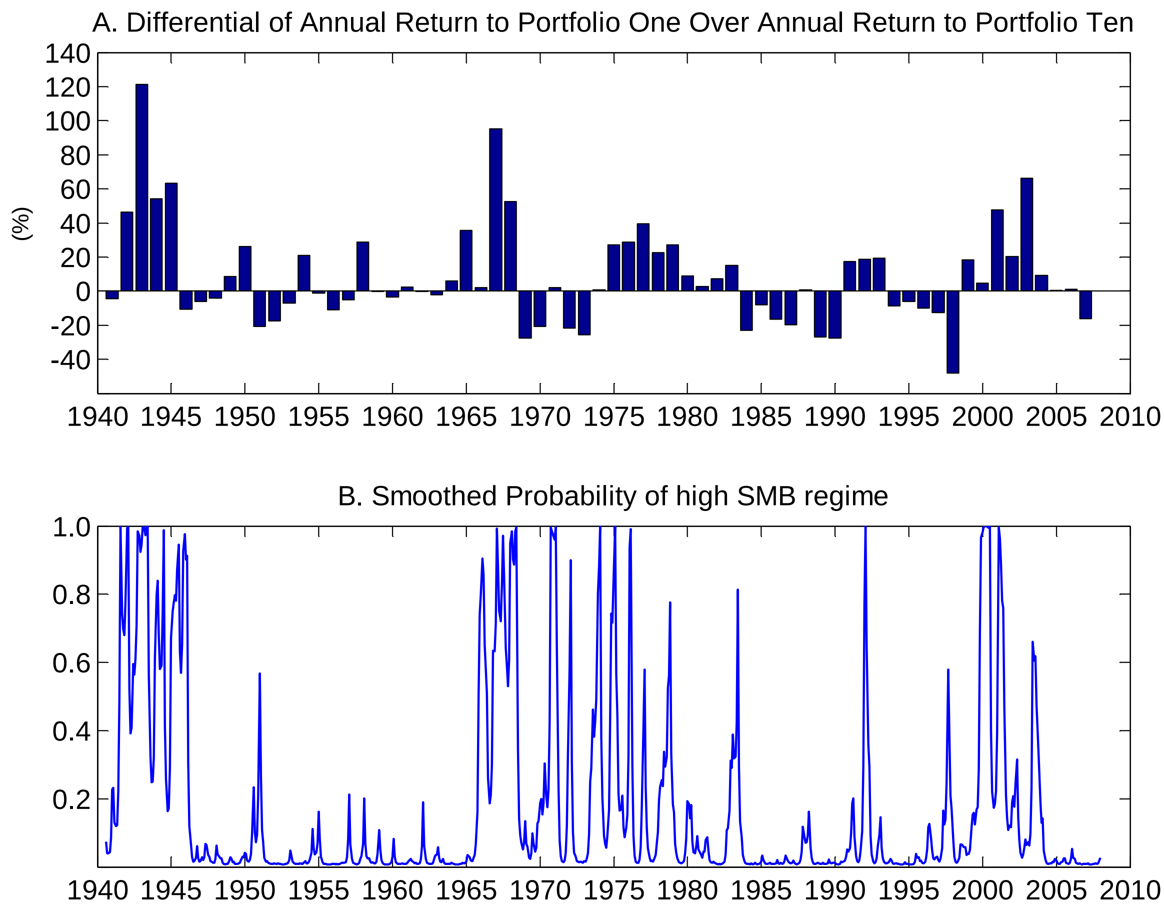


Figure 2: Three different dummy variables indicates three different economic environments. The first row includes the GDP growth rate of the U.S. and the business cycle dummy. The second row presents the CRSP monthly return and the market trend dummy variable derived from the smoothed probability of the bull market regime. The third row contains the credit spread and the high credit spread dummy also generated from the smoothed inference of a two-state Markov-switching model.

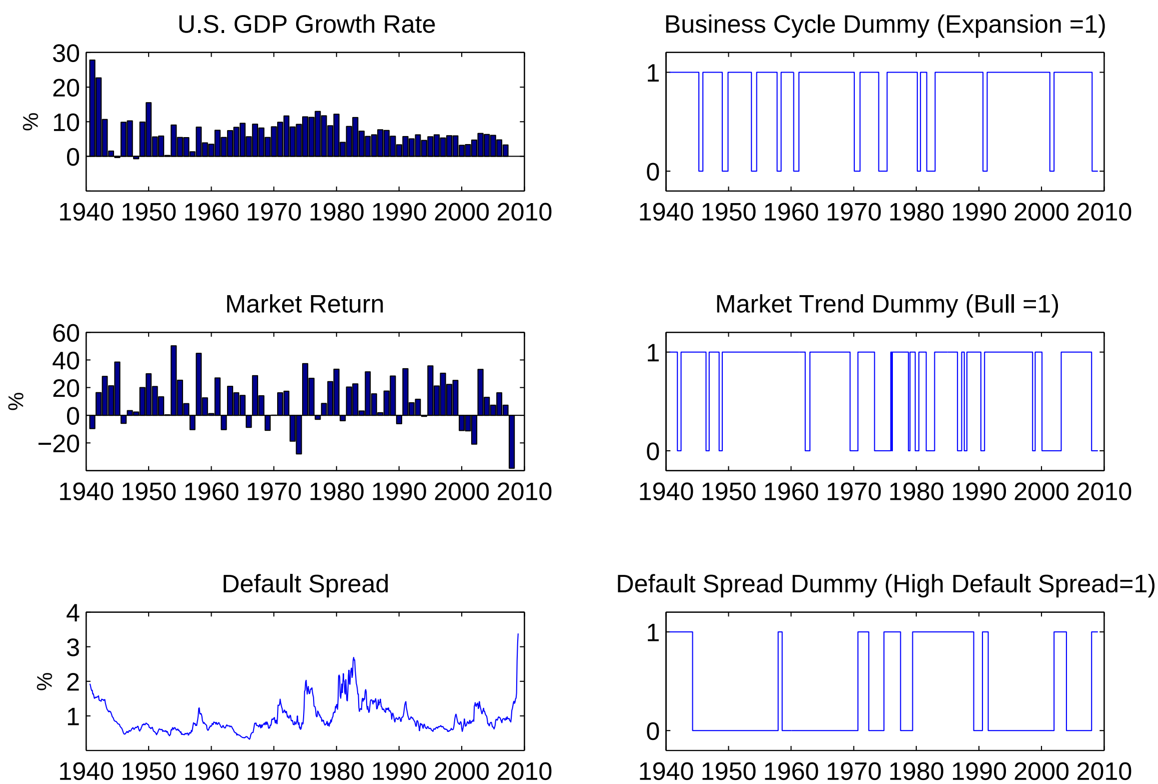


Table 1: **Returns on Size Portfolios and Size Premiums in Reference to CAPM**

Panel A. Full Sample (1926.7 to 2008.12)

	1 (Small)	2	3	4	5	6	7	8	9	10 (Big)
Mean Return	17.36	14.79	14.52	14.37	13.68	13.22	12.75	12.16	11.66	10.14
Standard Dev.	35.46	30.86	28.39	26.58	25.08	23.68	22.77	21.82	20.24	17.80
β	1.46	1.40	1.34	1.27	1.25	1.20	1.16	1.13	1.05	0.93
Size Premium	3.39	1.21	1.37	1.70	1.21	1.08	0.85	0.53	0.54	-0.10

Panel B. 1926.7 to 1980.6

	1 (Small)	2	3	4	5	6	7	8	9	10 (Big)
Mean Return	20.44	16.19	15.61	15.23	14.14	13.84	12.58	12.22	11.45	9.70
Standard Deviation	41.17	34.89	31.96	29.55	27.82	26.30	25.13	23.80	22.12	19.04
CAPM β	1.60	1.48	1.41	1.32	1.29	1.24	1.19	1.14	1.07	0.93
Size Premium	5.14	1.79	1.80	2.11	1.30	1.38	0.50	0.54	0.33	-0.29

Panel C. 1980.7 to 1998.6

	1 (Small)	2	3	4	5	6	7	8	9	10 (Big)
Mean Return	12.93	14.50	15.96	16.52	17.23	16.96	17.16	15.94	16.84	17.40
Standard Dev.	17.63	17.89	17.77	17.66	17.16	16.24	16.09	15.58	15.32	14.32
β	0.95	1.07	1.10	1.10	1.09	1.05	1.08	1.04	1.04	0.96
Size Premium	-2.99	-2.61	-1.40	-0.90	-0.08	0.01	-0.03	-0.93	0.01	1.31

Panel D. 1998.7 to 2008.12

	1 (Small)	2	3	4	5	6	7	8	9	10 (Big)
Mean Return	9.14	8.05	6.48	6.26	5.23	3.61	6.03	5.36	3.87	-0.03
Standard Dev.	25.11	26.08	23.24	22.94	21.33	19.83	19.57	20.24	17.13	16.10
β	1.06	1.21	1.15	1.13	1.13	1.08	1.08	1.14	0.98	0.92
Size Premium	7.47	6.59	4.95	4.68	3.66	1.97	4.38	3.80	2.07	-1.92

All securities in NYSE, AMEX and NASDAQ are sorted at the end of June of each year t and are assigned to ten different size portfolios according to NYSE breakpoints. The size portfolios are constructed with securities in each size group with their respective market cap as weights and are held from July of year t through June of year $t+1$.

β's are estimated with regression of monthly portfolio returns in excess of the Ibbotson Associates risk free rate on the CRSP value-weighted market returns in excess of the same risk free rate.

The size premium is calculated by subtracting the product of the CAPM beta and the equity premium from the size portfolio returns in excess of the risk free rate. All the equity risk premiums in different panels are estimated from their respective sample periods.

Returns, standard deviations and size premiums are all annualized and in percentage points.

Table 2: **Prices of Fama-French Risk Factors**

	1926.7-2007.12	1926.7-1980.6	1980.7-1998.6	1998.7-2007.12
$R_m - R_f$	0.64 (0.17)	0.70 (0.23)	0.84 (0.29)	-0.04 (0.44)
SMB	0.24 (0.11)	0.29 (0.14)	-0.04 (0.17)	0.47 (0.37)
HML	0.38 (0.12)	0.41 (0.15)	0.41 (0.18)	0.24 (0.35)

I calculate the price of risk of the Fama-French (1993) three factors with Fama and MacBeth (1973)'s two-pass regression approach. These data are retrieved from Professor French's website at Dartmouth. Test portfolios are obtained from 25 portfolios formed on size and book-to-market equity and 17 industry portfolios. Since there exist missing values in one of the 25 size/BM portfolio, it is taken out of the portfolio set. The returns on the remaining 41 test portfolios are named as R_{it}, $i = 1, 2, \ldots, N, N = 41$.
First we find beta estimates from the time-series regressions,

$$R^e_{it} = \alpha_i + \beta_i R^e_{mt} + s_i SMB_t + h_i HML_t + \varepsilon_{it} \qquad t = 1, 2, \ldots, T, \forall i.$$

where $R^e_{it} = R_{it} - R_{ft}$ and $R^e_{mt} = R_{mt} - R_{ft}$.
Then estimate the factor risk premiums λ from a cross-sectional regression,

$$E_T(R^e_i) = \beta_i \lambda_1 + s_i \lambda_2 + h_i \lambda_3 + \alpha_i, \qquad i = 1, 2, \ldots, N.$$

Since the pricing errors α_i are likely to be correlated, we follow Cochrane (2005)'s suggestion to run a GLS cross-sectional regression and the estimations of the price of risk are

$$\begin{aligned} \hat{\lambda} &= (\beta \Sigma^{-1} \beta)^{-1} \beta \Sigma^{-1} E_T(R^e), \text{and} \\ \sigma^2(\hat{\lambda}) &= \frac{1}{T}\left[(\beta \Sigma_f^{-1} \beta)^{-1} + \Sigma_f\right] \end{aligned}$$

where β is an N-by-3 matrix with $[\beta_i \; s_i \; h_i]$ in each row, $\lambda = [\lambda_1 \; \lambda_2 \; \lambda_3]$, f is a T-by-3 matrix of the risk factors, R^e_{mt}, SMB, HML.
The sample period is broken down like in Table 1. The parameter estimates in each subperiod use only observations from that subperiod. Standard deviations of λ estimates are reported in parentheses.
The insignificance of parameters in the subperiod from July 1996 to December 2007 probably results from sample selection and short sample period. The most interesting finding is on λ_2, the price of the risk factor SMB. During the sample period from July 1980 to June 1996, the price of this factor is not only insignificant but also much smaller in its value.

Table 3: **Regime Switching Model of the return difference between the 1st and 10th decile Size Portfolios**

	Regime Switching Model Parameter	Regime Switching Model Standard Deviation		Unconditional Normal Dist Parameter	Unconditional Normal Dist Standard Deviation
μ_1	-0.002436	0.00189	μ	0.004590	0.001825
μ_2	0.036465	0.01184			
σ_1^2	0.001263	0.00013	σ^2	0.052284	0.000136
σ_2^2	0.008167	0.00179			
p	0.9579	0.01991			
q	0.8090	0.11592			
Log-Likelihood Value	1367.73901			1257.87773	
AIC	-2723.47802			-2511.75546	
BIC	-2695.20758			-2502.33198	

Table 4: **Size Premium of *t+j* Decile Size Portfolio**

	Small	2	3	4	5	6	7	8	9	Big
t+1	1.49	0.57	0.94	1.26	0.87	0.48	1.02	0.48	0.50	-0.19
	(0.56)	(0.42)	(0.34)	(0.31)	(0.26)	(0.22)	(0.18)	(0.16)	(0.12)	(0.11)
t+2	1.02	1.70	1.63	1.50	1.16	0.53	0.36	0.84	0.36	-0.14
	(0.52)	(0.40)	(0.33)	(0.29)	(0.25)	(0.21)	(0.18)	(0.15)	(0.13)	(0.11)
t+3	-0.67	1.33	1.51	0.77	1.46	0.47	0.34	0.52	0.17	0.03
	(0.48)	(0.39)	(0.32)	(0.29)	(0.25)	(0.22)	(0.18)	(0.15)	(0.13)	(0.12)
t+4	-1.60	1.96	0.79	1.69	0.82	-0.04	0.59	0.37	0.40	0.10
	(0.45)	(0.37)	(0.32)	(0.29)	(0.25)	(0.22)	(0.18)	(0.16)	(0.12)	(0.12)
t+5	-0.83	1.42	1.26	0.58	-0.44	0.73	0.88	0.53	0.27	0.10
	(0.44)	(0.37)	(0.31)	(0.27)	(0.24)	(0.20)	(0.19)	(0.15)	(0.12)	(0.12)
t+6	-0.18	0.43	0.91	0.38	0.29	0.90	0.49	0.77	0.18	0.14
	(0.44)	(0.36)	(0.30)	(0.27)	(0.23)	(0.21)	(0.19)	(0.14)	(0.13)	(0.12)
t+7	-1.57	0.51	0.43	0.27	0.66	0.89	-0.78	0.12	0.50	0.29
	(0.43)	(0.35)	(0.30)	(0.26)	(0.24)	(0.21)	(0.17)	(0.15)	(0.14)	(0.12)
t+8	-1.31	-0.54	0.86	0.99	0.19	0.12	0.34	0.27	0.64	0.11
	(0.42)	(0.33)	(0.30)	(0.25)	(0.23)	(0.20)	(0.18)	(0.14)	(0.13)	(0.13)
t+9	-1.38	-0.46	0.43	-0.02	0.98	0.01	1.27	-0.42	0.47	0.16
	(0.39)	(0.32)	(0.30)	(0.26)	(0.24)	(0.21)	(0.20)	(0.17)	(0.14)	(0.13)
t+10	-1.61	-0.72	-0.65	1.22	-0.08	0.33	-1.02	-0.26	0.76	0.20
	(0.38)	(0.31)	(0.30)	(0.25)	(0.23)	(0.21)	(0.20)	(0.19)	(0.13)	(0.14)
t+11	-1.30	-0.62	-0.76	0.05	0.12	0.18	-0.36	0.56	-0.12	0.31
	(0.39)	(0.31)	(0.28)	(0.26)	(0.24)	(0.20)	(0.21)	(0.17)	(0.13)	(0.14)
t+12	-1.62	-1.60	-0.83	1.11	0.12	0.37	0.14	-0.21	-0.17	0.33
	(0.39)	(0.30)	(0.30)	(0.26)	(0.23)	(0.21)	(0.20)	(0.16)	(0.14)	(0.14)
t+13	-1.40	-2.30	-0.20	0.72	0.36	-0.04	-0.62	-0.51	-0.26	0.35
	(0.38)	(0.31)	(0.30)	(0.26)	(0.25)	(0.21)	(0.19)	(0.18)	(0.15)	(0.14)
t+14	-2.64	-1.08	-1.22	0.90	-0.45	-1.08	-0.91	-0.84	-0.26	0.42
	(0.38)	(0.31)	(0.31)	(0.27)	(0.25)	(0.22)	(0.21)	(0.19)	(0.15)	(0.15)
t+15	-3.14	-0.86	-1.50	-0.01	-1.02	-1.29	-0.83	-0.81	-1.21	0.68
	(0.39)	(0.31)	(0.30)	(0.26)	(0.24)	(0.24)	(0.23)	(0.20)	(0.16)	(0.15)

Standard deviations of mean returns (or return differential in the last column) are in the parentheses.

CAPM betas used in this table are estimated with full sample period (July 1926 to December 2008) instead of the trimmed sample period (July 1940 to December 2008) for the $t+j$ portfolios. The size premium of the $t+1$ portfolios here and the size premium of the Panel A of Table 1 should be the same if given the same length of sample.

Table 5: **Average Returns on *t+j* Decile Size Portfolio and Decile 1- Decile 10 Return Difference**

	Small	2	3	4	5	6	7	8	9	Big	1-10
t+1	16.17	14.85	14.78	14.61	14.02	13.29	13.58	12.76	12.27	10.68	5.49
	(0.81)	(0.74)	(0.69)	(0.67)	(0.63)	(0.60)	(0.59)	(0.57)	(0.53)	(0.49)	(0.63)
t+2	15.71	15.98	15.47	14.84	14.30	13.33	12.92	13.13	12.13	10.73	4.97
	(0.80)	(0.74)	(0.69)	(0.67)	(0.63)	(0.60)	(0.60)	(0.57)	(0.54)	(0.48)	(0.60)
t+3	14.01	15.61	15.35	14.12	14.61	13.27	12.89	12.81	11.94	10.90	3.12
	(0.79)	(0.75)	(0.69)	(0.66)	(0.63)	(0.62)	(0.59)	(0.57)	(0.53)	(0.48)	(0.58)
t+4	13.08	16.23	14.64	15.03	13.97	12.77	13.14	12.66	12.17	10.97	2.12
	(0.78)	(0.73)	(0.69)	(0.66)	(0.65)	(0.61)	(0.59)	(0.55)	(0.53)	(0.48)	(0.56)
t+5	13.85	15.69	15.10	13.93	12.71	13.53	13.43	12.81	12.04	10.97	2.88
	(0.78)	(0.73)	(0.70)	(0.66)	(0.64)	(0.60)	(0.58)	(0.56)	(0.53)	(0.47)	(0.55)
t+6	14.50	14.71	14.76	13.72	13.44	13.71	13.04	13.06	11.95	11.01	3.49
	(0.78)	(0.74)	(0.69)	(0.65)	(0.62)	(0.60)	(0.59)	(0.56)	(0.53)	(0.47)	(0.55)
t+7	13.12	14.79	14.27	13.61	13.80	13.70	11.77	12.41	12.27	11.15	1.96
	(0.79)	(0.73)	(0.68)	(0.63)	(0.63)	(0.60)	(0.59)	(0.56)	(0.53)	(0.47)	(0.56)
t+8	13.38	13.73	14.70	14.34	13.34	12.92	12.89	12.55	12.41	10.98	2.40
	(0.78)	(0.72)	(0.68)	(0.64)	(0.63)	(0.61)	(0.58)	(0.55)	(0.52)	(0.47)	(0.55)
t+9	13.30	13.82	14.27	13.33	14.13	12.82	13.82	11.86	12.24	11.03	2.27
	(0.76)	(0.70)	(0.69)	(0.64)	(0.63)	(0.60)	(0.59)	(0.55)	(0.53)	(0.46)	(0.51)
t+10	13.08	13.56	13.20	14.57	13.07	13.13	11.54	12.03	12.53	11.07	2.00
	(0.75)	(0.69)	(0.69)	(0.64)	(0.63)	(0.59)	(0.59)	(0.55)	(0.53)	(0.46)	(0.50)
t+11	13.38	13.65	13.09	13.40	13.27	12.99	12.19	12.85	11.65	11.18	2.20
	(0.74)	(0.70)	(0.68)	(0.63)	(0.63)	(0.58)	(0.58)	(0.54)	(0.53)	(0.46)	(0.49)
t+12	13.06	12.68	13.02	14.46	13.27	13.18	12.69	12.08	11.60	11.20	1.87
	(0.74)	(0.68)	(0.69)	(0.63)	(0.63)	(0.59)	(0.56)	(0.55)	(0.53)	(0.46)	(0.50)
t+13	13.28	11.97	13.65	14.07	13.51	12.77	11.93	11.78	11.51	11.21	2.07
	(0.74)	(0.68)	(0.69)	(0.62)	(0.61)	(0.59)	(0.58)	(0.54)	(0.53)	(0.46)	(0.49)
t+14	12.04	13.19	12.62	14.25	12.70	11.72	11.65	11.45	11.51	11.28	0.76
	(0.73)	(0.67)	(0.67)	(0.62)	(0.62)	(0.59)	(0.59)	(0.55)	(0.52)	(0.46)	(0.48)
t+15	11.54	13.42	12.34	13.34	12.12	11.52	11.72	11.48	10.56	11.55	-0.01
	(0.74)	(0.66)	(0.66)	(0.63)	(0.60)	(0.59)	(0.58)	(0.53)	(0.52)	(0.46)	(0.50)

Standard deviations of mean returns (or return differential in the last column) are in the parentheses.

Table 6: **Robustness Check: Size Premium of Different Size Portfolios in Reference to CAPM Projected Return**

	Small	Big	S-30%	M-40%	B-30%
t+1	0.96	0.02	0.91	0.91	-0.05
	(0.32)	(0.05)	(0.40)	(0.21)	(0.06)
t+2	1.51	0.05	1.60	0.77	0.02
	(0.31)	(0.05)	(0.38)	(0.20)	(0.07)
t+3	1.09	0.11	0.94	0.70	0.08
	(0.30)	(0.06)	(0.36)	(0.19)	(0.08)
t+4	0.99	0.14	0.72	0.65	0.13
	(0.28)	(0.07)	(0.35)	(0.18)	(0.08)
t+5	0.44	0.20	0.95	0.46	0.15
	(0.26)	(0.07)	(0.34)	(0.17)	(0.08)
t+6	0.30	0.23	0.49	0.52	0.21
	(0.25)	(0.07)	(0.32)	(0.17)	(0.09)
t+7	0.03	0.24	-0.10	0.07	0.28
	(0.24)	(0.07)	(0.30)	(0.17)	(0.09)
t+8	0.17	0.20	-0.25	0.37	0.19
	(0.23)	(0.08)	(0.30)	(0.16)	(0.09)
t+9	0.10	0.21	-0.31	0.52	0.15
	(0.23)	(0.09)	(0.29)	(0.16)	(0.10)
t+10	-0.22	0.17	-1.05	-0.14	0.26
	(0.22)	(0.09)	(0.27)	(0.16)	(0.10)
t+11	-0.35	0.22	-1.04	-0.30	0.24
	(0.21)	(0.09)	(0.26)	(0.16)	(0.10)
t+12	-0.28	0.21	-1.30	0.23	0.18
	(0.21)	(0.10)	(0.27)	(0.16)	(0.11)
t+13	-0.28	0.13	-1.16	-0.02	0.16
	(0.21)	(0.10)	(0.26)	(0.16)	(0.11)
t+14	-0.50	0.07	-1.52	-0.55	0.21
	(0.21)	(0.11)	(0.26)	(0.16)	(0.12)
t+15	-0.97	0.10	-1.68	-0.87	0.22
	(0.20)	(0.12)	(0.26)	(0.17)	(0.12)

Standard deviations of mean returns (or return differential in the last column) are in the parentheses.

Table 7: **Average Size Premium of Portfolio 1 under Different Economic Environments**

	Total	Expansion	Contraction	Bull Mkt	Bear Mkt	High CS	Low CS
t+1	1.49	2.07	-1.78	0.65	4.57	**5.45**	**-0.45**
	(0.56)	(0.61)	(1.42)	(0.57)	(1.57)	**(1.15)**	**(0.62)**
t+2	1.02	1.36	-0.86	0.15	4.24	**4.57**	**-0.71**
	(0.52)	(0.56)	(1.35)	(0.53)	(1.47)	**(1.01)**	**(0.60)**
t+3	-0.67	-0.70	-0.47	-1.08	0.84	2.17	-2.06
	(0.48)	(0.52)	(1.30)	(0.50)	(1.32)	(0.90)	(0.57)
t+4	-1.60	-1.51	-2.09	-2.13	0.35	**2.62**	**-3.67**
	(0.45)	(0.48)	(1.30)	(0.47)	(1.23)	**(0.83)**	**(0.54)**
t+5	-0.83	-0.82	-0.87	-1.33	1.02	**3.34**	**-2.87**
	(0.44)	(0.48)	(1.19)	(0.45)	(1.24)	**(0.79)**	**(0.53)**
t+6	-0.18	-0.23	0.06	-0.72	1.80	**3.18**	**-1.83**
	(0.44)	(0.47)	(1.17)	(0.45)	(1.21)	**(0.75)**	**(0.54)**
t+7	-1.57	-1.67	-0.97	-1.26	-2.70	**2.56**	**-3.59**
	(0.43)	(0.46)	(1.16)	(0.43)	(1.24)	**(0.72)**	**(0.53)**
t+8	-1.31	-1.27	-1.51	-1.30	-1.32	**1.60**	**-2.73**
	(0.42)	(0.44)	(1.28)	(0.43)	(1.14)	**(0.72)**	**(0.51)**
t+9	-1.38	-1.25	-2.12	-1.93	0.64	**3.54**	**-3.79**
	(0.39)	(0.42)	(1.13)	(0.42)	(1.01)	**(0.68)**	**(0.48)**
t+10	-1.61	-1.47	-2.36	**-2.99**	**3.48**	**2.38**	**-3.56**
	(0.38)	(0.40)	(1.13)	**(0.40)**	**(1.03)**	**(0.65)**	**(0.47)**
t+11	-1.30	-1.21	-1.83	**-2.64**	**3.61**	**1.22**	**-2.54**
	(0.39)	(0.41)	(1.17)	**(0.40)**	**(1.03)**	**(0.65)**	**(0.48)**
t+12	-1.62	-1.80	-0.61	**-2.60**	**1.97**	**1.23**	**-3.01**
	(0.39)	(0.41)	(1.13)	**(0.41)**	**(1.06)**	**(0.69)**	**(0.47)**
t+13	-1.40	-1.22	-2.42	-2.20	1.55	0.35	-2.25
	(0.38)	(0.40)	(1.16)	(0.40)	(1.03)	(0.68)	(0.47)
t+14	-2.64	-2.33	-4.37	-3.39	0.11	**0.33**	**-4.09**
	(0.38)	(0.40)	(1.12)	(0.39)	(1.04)	**(0.67)**	**(0.46)**
t+15	-3.14	-3.20	-2.82	**-4.41**	**1.53**	**1.30**	**-5.32**
	(0.39)	(0.42)	(1.12)	**(0.39)**	**(1.12)**	**(0.74)**	**(0.45)**
Number of Observations	822	698	124	646	176	270	552

The standard deviation of the average size premium is in the parenthesis.

The first column shows the average size premium of the first decile size portfolio, which is the same as the first column of Table 4.

The number of observations in each state is in the last row of the table. The second and third columns are the expansion and contraction states; the fourth and fifth columns are the bull and bear market states; and the last two columns are the high and low credit spread states.

The size premiums are shown in **boldface fonts** if the difference is significant at the 10 percent level using a one-sided t test.

Table 8: **Average Size Premium of Portfolio 10 under Different Economic Environments**

	Total	Expansion	Contraction	Bull Mkt	Bear Mkt	High CS	Low CS
t+1	-0.19	-0.17	-0.27	-0.29	0.21	**-1.10**	**0.26**
	(0.11)	(0.12)	(0.29)	(0.11)	(0.32)	**(0.20)**	**(0.13)**
t+2	-0.14	-0.14	-0.12	-0.39	0.80	**-1.10**	**0.34**
	(0.11)	(0.12)	(0.29)	(0.11)	(0.34)	**(0.20)**	**(0.13)**
t+3	0.03	0.03	0.05	**-0.34**	**1.38**	**-0.87**	**0.47**
	(0.12)	(0.12)	(0.30)	**(0.11)**	**(0.35)**	**(0.20)**	**(0.14)**
t+4	0.10	0.04	0.43	**-0.33**	**1.66**	**-0.63**	**0.45**
	(0.12)	(0.13)	(0.31)	**(0.11)**	**(0.35)**	**(0.21)**	**(0.14)**
t+5	0.10	-0.03	0.85	**-0.42**	**2.02**	**-0.73**	**0.51**
	(0.12)	(0.13)	(0.32)	**(0.11)**	**(0.36)**	**(0.21)**	**(0.14)**
t+6	0.14	0.00	0.95	**-0.43**	**2.22**	**-0.59**	**0.50**
	(0.12)	(0.13)	(0.33)	**(0.11)**	**(0.38)**	**(0.21)**	**(0.15)**
t+7	0.29	0.11	1.29	**-0.37**	**2.68**	-0.29	0.57
	(0.12)	(0.13)	(0.34)	**(0.12)**	**(0.39)**	(0.22)	(0.15)
t+8	0.11	-0.08	1.17	**-0.49**	**2.30**	-0.55	0.43
	(0.13)	(0.14)	(0.33)	**(0.12)**	**(0.42)**	(0.22)	(0.16)
t+9	0.16	0.01	1.03	**-0.52**	**2.67**	-0.60	0.54
	(0.13)	(0.14)	(0.32)	**(0.12)**	**(0.44)**	(0.21)	(0.17)
t+10	0.20	0.03	1.16	**-0.45**	**2.60**	-0.51	0.55
	(0.14)	(0.15)	(0.34)	**(0.12)**	**(0.46)**	(0.22)	(0.17)
t+11	0.31	0.12	1.37	**-0.45**	**3.10**	-0.38	0.65
	(0.14)	(0.16)	(0.36)	**(0.12)**	**(0.49)**	(0.22)	(0.18)
t+12	0.33	0.20	1.08	**-0.43**	**3.11**	-0.37	0.67
	(0.14)	(0.16)	(0.37)	**(0.13)**	**(0.49)**	(0.23)	(0.18)
t+13	0.35	0.18	1.27	**-0.42**	**3.15**	-0.25	0.64
	(0.14)	(0.16)	(0.39)	**(0.13)**	**(0.48)**	(0.24)	(0.18)
t+14	0.42	0.21	1.55	**-0.28**	**2.96**	-0.14	0.68
	(0.15)	(0.16)	(0.38)	**(0.13)**	**(0.51)**	(0.24)	(0.19)
t+15	0.68	0.49	1.76	**-0.13**	**3.67**	-0.03	1.03
	(0.15)	(0.16)	(0.39)	**(0.13)**	**(0.53)**	(0.24)	(0.19)
Number of Observations	822	698	124	646	176	270	552

The standard deviation of the average size premium is in the parenthesis.

The first column shows the average size premium of the 10th decile size portfolio, which is the same as the last column of Table 4.

Column 2 to column 7 use the same dummy variables to separate different states as the corresponding columns in Table 7.

The size premiums are shown in **boldface fonts** if the difference is significant at the 10 percent level using a one-sided *t* test.

How useful are historical data for forecasting the long-run equity premium?

John M. Maheu and Thomas H. McCurdy*

March 2006

Abstract

We provide an optimal approach to forecasting the long-run (unconditional) equity premium in the presence of structural breaks. This forecasting procedure determines in real time how useful historical data are in updating our prior belief about the distribution of market excess returns. The value of historical data has varied considerably, implying that ignoring structural breaks or using a rolling window is not optimal. We obtain realistic out-of-sample forecasts for the entire 1885-2003 period; the forecast at the end of the sample is 4.02 for the structural break model and 5.10 for a no-break model. The results are robust to a wide-range of distributional assumptions about excess returns.

*Maheu (jmaheu@chass.utoronto.ca), Department of Economics, University of Toronto; McCurdy (tmccurdy@rotman.utoronto.ca), Joseph L. Rotman School of Management, University of Toronto, and Associated Fellow, CIRANO. We thank David Goldreich, Mark Kamstra, Lisa Kramer, Jan Mahrt-Smith, Kevin Wang, as well as seminar participants of the $(EC)^2$ conference Istanbul 2005 and the Bank of Canada, for helpful comments. Lois Chan provided excellent research assistance. We are also grateful to the SSHRC for financial support.

1 Introduction

An important topic in finance is the forecast of the return premium on a well diversified portfolio of equity relative to a riskfree asset. Accurate forecasts of this market equity premium are required for capital budgeting, investment, and pricing decisions.

There is an extensive literature that seeks to explain the long-run equity premium. Most of this literature takes as given simple point estimates of the premium obtained as the sample average from a long series of excess return data.[1] In addition, many forecasters, including those using dynamic models with many predictors, report the sample average of excess returns as a benchmark.[2]

The use of a sample average as a forecast of the long-run equity premium assumes that excess returns are stationary and that the process governing them does not undergo structural breaks. Once we allow for structural breaks, it is not clear whether or not historical data are useful for forecasting the equity premium. For instance, including data prior to a structural break may result in a biased forecast. The purpose of this paper is to investigate the value of data in updating our beliefs about the long-run equity premium, and to provide forecasts of the premium while allowing for structural breaks.

We focus on the unconditional distribution of excess market returns and define the long-run premium as the mean of that distribution.[3] Investment and capital budgeting decisions often span many years. With this investment horizon, the long-run equity premium is the relevant measure. Jacquier, Kane, and Marcus (2005) discuss the importance of accurate premium estimates for long-orizon portfolio choice. In addition, by focusing on the long-run premium, as opposed to short-run dynamic models of the premium, we may be less susceptible to model misspecification. That is, the existence of a long-run value of the premium is consistent with different underlying models of risk.

Nevertheless, even for the unconditional distribution of excess returns, misspecified models may provide evidence of structural breaks when the underlying data generating process (DGP) is in fact stable. For example, suppose one assumed a Normal distribution for excess returns when in fact the DGP has fat tails. In this case, realizations in the tail of the maintained Normal distribution could be mistakenly interpreted in real time

[1] For example, Table 1 in a recent survey by Mehra and Prescott (2003) lists four estimates of the unconditional premium using sample averages of data from 1802-1998, 1871-1999, 1889-2000, and 1926-2000.

[2] Derrig and Orr (2004) survey a wide range of both academic and practitioner data-based estimates of the equity premium. There are many asset pricing models that have been used to estimate this premium, building on the three-factor model of Fama and French (1992) or the arbitrage pricing theory of Ross (1976). Another approach uses earnings or dividend growth to model the equity premium, for example, Donaldson, Kamstra, and Kramer (2004) and Fama and French (2002). Estimates of the equity premium in the presence of regimes changes include Mayfield (2004) and Turner, Startz, and Nelson (1989). Recent examples of premium forecasts include Campbell and Thompson (2004), and Goyal and Welch (2004).

[3] In this paper we view the full data set as being potentially partitioned into sequences of data generated from different stationary models. Therefore, within each partition there is a well defined unconditional premium.

as evidence of a structural break. To minimize this potential problem, we use a very flexible model to forecast the long-run premium. In particular, our maintained model is a mixture-of-Normals which can capture skewness and excess kurtosis, both of which are well known features of returns. For robustness, we compare our results to the nested Normal distribution case to see if the more general distribution affects our inference about structural change.[4]

The Bayesian approach to prediction integrates out parameter uncertainty. For example, see Barberis (2000), and Kandel and Stambaugh (1996). Important papers by Pastor and Stambaugh (2001) and Kim, Morley, and Nelson (2005) provide smoothed historical estimates of the equity premium in the presence of structural breaks using a dynamic risk-return model.[5] These papers are based on the structural break model of Chib (1998) which provides estimates conditional on a maintained number of breaks in-sample.

A primary objective of our paper is to stress the learning aspect that would occur in real time and its implications for decision making. That is, we investigate how the evidence for structural breaks changes over time and assess the effects on real time forecasts of ignoring this information. Therefore, our forecasts of the premium also incorporate time-varying model uncertainty. Our approach provides period-by-period out-of-sample forecasts of the premium, incorporating the probability of structural breaks in the past data as well as the possibility of breaks in the future. A by-product of our approach is that it generates an estimate of the number of historical observations that are useful at each point in time for forecasting the long-run premium.

In addition, our maintained model of excess returns, which is subject to structural breaks, can capture heteroskedasticity, asymmetry and fat tails. These are features that may be important for forecasts of the equity premium as well as for identifying structural breaks. As noted above, this allows us to assess the impact of outliers on structural break identification.

Intuitively, if a structural break occurred in the past we would want to adjust our use of the old data in our estimation procedure since those data could bias our estimates and forecasts. This might suggest a rolling window estimator that only uses a portion of the available data. However, such an approach will not be optimal. Indeed, some combination of the data that follow a perceived break, and the (biased) data that preceded it may be a better approach.

To formally deal with this issue, we use the methodology of Maheu and Gordon (2005) and assume that structural breaks are exogenous, unpredictable events that result in a change in the parameter vector associated with the maintained model (in this case a mixture-of-Normals model of excess returns). The structural break model is constructed from a series of submodels. Each submodel has an identical parameterization for excess

[4]A second reason to take the maintained specification of excess returns seriously is that our Bayesian approach provides exact finite sample inference only if the model is well specified.

[5]Additional work on structural breaks in finance include Andreou and Ghysels (2002) and Pettenuzzo and Timmermann (2004).

returns but the parameter is estimated with a different history of data. Each of the submodels assume that once a break occurs, past data are not useful in learning about the new parameter value, only future data can be used to update beliefs. Submodels are differentiated by when they start and the data they use. New submodels are continually introduced through time to allow for multiple structural breaks, and for a potential break out-of-sample.

Since structural breaks can never be identified with certainty, Bayesian model averaging provides a predictive distribution, which accounts for past and future structural breaks, by integrating over each of the possible submodels weighted by their probabilities.[6] Therefore new submodels, which are based on recent shorter histories of data, only receive significant weights once their predictive performance warrants it. The model average optimally combines the past (potentially biased) data from before the estimated break point, which will tend to have less uncertainty about the premium due to sample length, with the less precise (but unbiased) estimates based on the more recent post-break data. Note that this implies that, in the presence of structural breaks, there does not exist an optimal rolling window estimator.

This approach provides a method to combine submodels estimated over different histories of data. After estimation we can estimate the average number of useful observations at any point in time. In addition, submodel uncertainty is accounted for in the analysis. For example, we show that there is considerable uncertainty as to the number of past observations to use in forecasting the premium toward the end of our sample.

The empirical results provide strong support for structural breaks. In particular, our evidence for structural breaks points towards at least 2 major breaks (1929 and 1940), and possibly a more recent structural break in the late 1990s. Note that these breaks are detected in real time and are not the result of a full-sample analysis. For example, using only data up to 1929:11, there is strong evidence (probability .94) that the most recent structural break occurred at 1929:6.

Ignoring structural breaks results at times in substantially different premium forecasts, as well as overconfidence in those estimates. When a structural break occurs there is a decrease in the precision of the premium estimate which improves as we learn about the new premium level. Uncertainty about the premium comes from two sources: submodel uncertainty and parameter uncertainty. For example, the uncertainty after the break in 1929 is mainly due to parameter uncertainty whereas the uncertainty in the late 1990s is from both submodel and parameter uncertainty. Differences between premium forecasts which account for structural breaks and those which do not, can be important for many applications. For example, we show that neglecting structural breaks has important implications for a pension fund manager who must finance future liabilities.

Due to the presence of asymmetry and fat tails in excess returns, we favor inference from our structural break model using a mixture-of-Normals submodel with two components. This model produces kurtosis values well above 3 and negative skewness throughout our sample of data. Our statistical measures clearly favor this specification.

[6] Other examples of Bayesian model averaging include Avramov (2002), and Cremers (2002).

Interestingly, the premium forecasts (predictive mean) are quantitatively similar to the structural break model with a single-component submodel. Where they differ is in the shape of the predictive distribution of the premium. In general the two-component model indicates that the predictive distribution of the premium is more disperse. This higher uncertainty associated with the equity premium will be important for investment decisions.

There is another important difference between the alternative parameterizations of the submodel. As we learn about the distribution governing excess returns, sometimes we infer a break that is later revised to be an outlier and not a structural break. The richer specification of the two-component submodel is more robust to these *false breaks*. One reason for this is that the two-component model is characterized by a high and low variance state. This allows for heteroskedasticity in excess returns. Therefore, outliers can occur and not be evidence of a break in the distribution of excess returns.

In summary, this paper makes several contributions to the prediction of the equity premium. First, we show that historical data are useful in updating our prior beliefs regarding the equity premium. In the presence of structural breaks, we provide an optimal approach to estimating and forecasting the long-run equity premium using historical data on excess returns. Our structural change model produces realistic forecasts of the premium over the entire 1885-2003 sample. The paper also illustrates the importance of submodel uncertainty and the value of modeling higher-order moments of excess returns when inferring structural breaks and predicting the equity premium. Ignoring structural breaks leads to substantially different premium forecasts as well as overconfidence in the estimates.

The paper is organized as follows. The next section describes the data sources. Section 3 provides an overview of alternative ways to use historical data in order to forecast the equity premium. Included are a case in which all data are used, a fixed-length rolling window of data, and the proposed optimal use of data when structural breaks are taken into account. Section 4 introduces a flexible mixture-of-Normals model for excess returns as our submodel parameterization. Section 5 reviews Bayesian estimation techniques for the mixture model of excess returns. The proposed method for optimal use of data for estimation and forecasting in the presence of structural breaks is outlined in Section 6. Results are reported in Section 7 using data from 1885 to 2003. Conclusions are found in Section 8.

2 Data

The equity data are monthly returns, including dividend distributions, on a well diversified market portfolio. The monthly equity returns for 1885:2 to 1925:12 were obtained from Bill Schwert; details of the data construction can be found in Schwert (1990). Monthly equity returns from 1926:1 to 2003:12 are from the Center for Research in Security Prices (CRSP) value-weighted portfolio, which includes securities on the New York stock exchange, American stock exchange and the NASDAQ. The returns were converted to continuously compounded monthly returns by taking the natural logarithm of the gross monthly return.

Data on the risk-free rate from 1885:2 to 1925:12 were obtained from annual interest rates supplied by Jeremy Siegel. Siegel (1992) describes the construction of the data in detail. Those annual interest rates were converted to monthly continuously compounded rates. Interest rates from 1926:1 to 2003:12 are from the U.S. 3 month T-bill rates supplied by the Fama-Bliss riskfree rate file provided by CRSP.

Finally, the monthly excess return, r_t, is defined as the monthly continuously compounded portfolio return minus the monthly riskfree rate. It is scaled to an annual excess return by multiplying by 12.

Figure 1 displays a time series plot of the annualized monthly excess returns while Table 1 reports summary statistics for excess returns. Both the skewness and kurtosis estimates suggest significant deviations from the Normal distribution.

3 Forecasting the Equity Premium

We define the long-run equity premium as the expected value of excess returns on a well diversified value-weighted portfolio of securities. In this paper we are concerned with methods of forecasting the long-run equity premium from a series of historical data. If there were no structural breaks, and excess returns were stationary, it would be optimal to use all available data. However, in the presence of breaks, our forecast of the premium, and our uncertainty about that forecast, could be very misleading if our modeling/forecasting does not take account of those structural breaks.

To focus on this issue, consider 3 alternative forecasts of the equity premium γ:

$\hat{\gamma}_{ALL,t-1}$ which is based on all available data up to time $t-1$;

$\hat{\gamma}_{W,t-1}$ which is based on a fixed-length rolling window of past data; and

$\hat{\gamma}_{B,t-1}$ uses historical data optimally given the possibility of structural breaks.

The first ignores any structural breaks. Using the average of the entire sample of excess returns is a common example of this approach. The second forecast recognizes that the distribution of excess returns may have undergone a structural break. The method therefore uses a rolling window of historical data for estimation. This has the advantage of dropping past data which may bias the estimate, but with the possible disadvantage of dropping too many data points, resulting in a reduction in the accuracy of the premium estimate. In addition, the second estimator is implicitly assuming that structural breaks are reoccurring by using a fixed window of data at each point in time. The final approach provides optimal use of past data in forecasting the premium. For this estimate, the number of useful data will vary over time and depend on our inference concerning structural breaks.

Section 4 describes our maintained mixture-of-Normals model of excess returns, which is subject to structural breaks. To model the value of historical data for our forecasts of the equity premium, it is natural to use Bayesian methods which stress the learning aspect of statistical inference. That is, how do our beliefs regarding the premium change after observing a set of realizations of excess returns? Section 5 outlines Bayesian estimation of the single-component and the mixture-of-Normals model of excess returns. Once structural breaks are allowed, the usefulness of historical data will be dependent on how recently a break has occurred. Given assumptions about the form of structural breaks, Section 6 provides a methodology to optimally use historical data in this setting. This provides the details of the out-of-sample estimate of $\hat{\gamma}_{B,t-1}$ with comparisons to $\hat{\gamma}_{ALL,t-1}$ and $\hat{\gamma}_{W,t-1}$.

4 Mixture-of-Normals Model for Excess Returns

Financial returns are well known to display skewness and kurtosis and our inference about the market premium may be sensitive to these characteristics of the shape of the distribution. Our maintained model of excess returns is a discrete mixture-of-Normals. Discrete mixtures are a very flexible method to capture various degrees of asymmetry and tail thickness. Indeed a sufficient number of components can approximate arbitrary distributions (Roeder and Wasserman (1997)). A k-component mixture model of returns can be represented as

$$r_t = \begin{cases} N(\mu_1, \sigma_1^2) & \text{with probability } \pi_1 \\ \vdots & \vdots \\ N(\mu_k, \sigma_k^2) & \text{with probability } \pi_k, \end{cases} \tag{4.1}$$

with $\sum_{j=1}^{k} \pi_j = 1$. It will be convenient to denote each mean and variance as μ_j, and σ_j^2, with $j \in \{1, 2, ..., k\}$. Data from this specification are generated as: first a component j is chosen according to the probabilities $\pi_1, ..., \pi_k$; then a return is generated from $N(\mu_j, \sigma_j^2)$. In other words, returns will display heteroskedasticity. Often a two-component specification is sufficient to capture the features of returns. Figure 2 displays examples of excess return distributions that can be obtained from only two components. Relative to the Normal distribution, the distributions exhibit fat-tails, skewness and combinations of skewness and fat-tails

Since our focus is on the moments of excess returns, in particular the mean, it will be useful to consider the implied moments of excess returns as a function of the model parameters. The relationships between the uncentered moments and the model parameters for a k-component model are:

$$\gamma = E r_t = \sum_{i=1}^{k} \mu_i \pi_i, \tag{4.2}$$

in which γ is defined as the equity premium; and

$$\gamma_2^{'} = Er_t^2 = \sum_{i=1}^{k}(\mu_i^2 + \sigma_i^2)\pi_i \tag{4.3}$$

$$\gamma_3^{'} = Er_t^3 = \sum_{i=1}^{k}(\mu_i^3 + 3\mu_i\sigma_i^2)\pi_i \tag{4.4}$$

$$\gamma_4^{'} = Er_t^4 = \sum_{i=1}^{k}(\mu_i^4 + 6\mu_i^2\sigma_i^2 + 3\sigma_i^4)\pi_i. \tag{4.5}$$

for the higher-order moments of returns. The higher-order centered moments $\gamma_j = E[(r_t - E(r_t))^j]$, $j = 2, 3, 4$, are then

$$\gamma_2 = \gamma_2^{'} - (\gamma)^2 \tag{4.6}$$

$$\gamma_3 = \gamma_3^{'} - 3\gamma\gamma_2^{'} + 2(\gamma)^3 \tag{4.7}$$

$$\gamma_4 = \gamma_4^{'} - 4\gamma\gamma_3^{'} + 6(\gamma)^2\gamma_2^{'} - 3(\gamma)^4. \tag{4.8}$$

As a special case, a one-component model allows for Normally distributed returns. As shown above, only two components are needed to produce skewness and excess kurtosis. If $\mu_1 = \cdots = \mu_k = 0$ and at least one variance parameter differs from the others the resulting density will have excess kurtosis but not asymmetry. To produce asymmetry and hence skewness we need $\mu_i \neq \mu_j$ for some $i \neq j$. Section 5 discusses a Bayesian approach to estimation of this model.

5 Bayesian Estimation

In the next two subsections we review Bayesian estimation methods for the mixture-of-Normals model. An important special case is when there is a single component $k = 1$ which we discuss first.

5.1 Gaussian Case, $k = 1$

When there is only one component our model for excess returns reduces to a Normal distribution with mean μ, variance σ^2, and likelihood function,[7]

$$p(r|\mu, \sigma^2) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(r_i - \mu)^2\right) \tag{5.1}$$

where $r = [r_1, ..., r_T]^T$. In the last section, this model is included as a special case when $\pi_1 = 1$.

[7]For the one-component case we drop the component subscript on the model parameters.

Bayesian methods require specification of a prior distribution over the parameters μ and σ^2. Given the independent priors $\mu \sim N(b, B)I_{\mu>0}$, and $\sigma^2 \sim IG(v/2, s/2)$,[8] Bayes rule gives the posterior distribution of μ and σ^2 as

$$p(\mu, \sigma^2|r) \propto p(r|\mu, \sigma^2)p(\mu)p(\sigma^2) \tag{5.2}$$

where $p(\mu)$ and $p(\sigma^2)$ denote the probability density functions of the priors. Note that the indicator function $I_{\mu>0}$ is 1 when $\mu > 0$ is true and otherwise 0. This restriction enforces a positive equity premium.

Our object of interest is the long-run equity premium γ defined as the mean of the excess returns distribution. Although closed form solutions for the posterior distribution are not available, we can use Gibbs sampling to simulate from the posterior and estimate quantities of interest. The Gibbs sampler iterates sampling from the following conditional distributions which forms a Markov chain.

1. sample $\mu \sim p(\mu|\sigma^2, r)$
2. sample $\sigma^2 \sim p(\sigma^2|\mu, r)$

These steps are repeated many times and an initial set of the draws are discarded to minimize startup conditions and ensure the remaining sequence of the draws is from the converged chain.[9] After obtaining a set of N draws $\{\mu^{(i)}, (\sigma^2)^{(i)}\}_{i=1}^N$ from the posterior, we can estimate moments using sample averages. For example, the posterior mean of γ, which is an estimate of the equity premium conditional on this model and data, can be estimated as

$$E[\mu|r_T] \approx \frac{1}{N}\sum_{i=1}^{N} \mu^{(i)}. \tag{5.3}$$

To measure the dispersion of the posterior distribution of the equity premium we could compute the posterior standard deviation of γ in an analogous fashion, using sample averages obtained from the Gibbs sampler in $\sqrt{E[\mu^2|r] - E[\mu|r]^2}$. Alternatively, we could summarize the marginal distribution of the equity premium with a histogram or kernel density estimate.

This simple model which assumes excess returns follow a Gaussian distribution cannot account for the asymmetry and fat tails found in return data. Modeling these features of returns may be important to our inference about the premium. The next section provides details on estimation for models with two or more components which can capture the higher-order moments of excess returns.

---

[8] Where $IG(,)$ denotes the inverse gamma distribution. See Bernardo and Smith (2000).

[9] See Chib (2001), Geweke (1997), Robert and Casella (1999) for background information on Markov chain Monte Carlo methods of which Gibbs sampling is a special case. See Johannes and Polson (2005) for a survey of financial applications.

5.2 Mixture Case, $k > 1$

In the case of $k > 1$ mixture-of-Normals the likelihood of excess returns is

$$p(r|\mu, \sigma^2, \pi) = \prod_{t=1}^{T} \sum_{j=1}^{k} \pi_j \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{1}{2\sigma_j^2}(r_t - \mu_j)^2 \right) \tag{5.4}$$

where $\mu = [\mu_1, ..., \mu_k]'$, $\sigma^2 = [\sigma_1^2, ..., \sigma_k^2]'$, and $\pi = [\pi_1, ..., \pi_k]$. Bayesian estimation of mixtures has been extensively discussed in the literature and our approach closely follows Diebolt and Robert (1994). We choose conditionally conjugate prior distributions which facilitate our Gibbs sampling approach. The independent priors are $\mu_i \sim N(b_i, B_{ii})$, $\sigma_i^2 \sim IG(v_i/2, s_i/2)$, and $\pi \sim \mathcal{D}(\alpha_1, ..., \alpha_k)$, where the latter is the Dirichlet distribution. We continue to impose a positive equity premium by giving zero support to any parameter configuration that violates $\gamma > 0$.

Discrete mixture models can be viewed as a simpler model if an indicator variable z_t records which observations come from component j. Our approach to Bayesian estimation of this model begins with the specification of a prior distribution and the augmentation of the parameter vector by the additional indicator $z_t = [0 \cdots 1 \cdots 0]$ which is a row vector of zeros with a single 1 in the position j if r_t is drawn from component j. Let Z be the matrix that stacks the rows of z_t, $t = 1, ..., T$.

With the full data r_t, z_t the data density becomes

$$p(r|\mu, \sigma^2, \pi, Z) = \prod_{t=1}^{T} \sum_{j=1}^{k} z_{t,j} \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{1}{2\sigma_j^2}(r_t - \mu_j)^2 \right). \tag{5.5}$$

Bayes theorem now gives the posterior distributions as

$$\begin{aligned} p(\mu, \sigma^2, \pi, Z|r) &\propto p(r|\mu, \sigma^2, \pi, Z)p(\mu, \sigma^2, \pi, Z) && (5.6)\\ &\propto p(r|\mu, \sigma^2, \pi, Z)p(Z|\mu, \sigma^2, \pi)p(\mu, \sigma^2, \pi). && (5.7) \end{aligned}$$

The posterior distribution has an unknown form, however, we can generate a sequence of draws from this density using Gibbs sampling. Just as in the $k = 1$ case, we sample from a set of conditional distributions and collect a large number of draws. From this set of draws we can obtain simulation consistent estimates of posterior moments. The Gibbs sampling routine repeats the following steps for posterior simulation.

1. sample $\mu \sim p(\mu|\sigma^2, \pi, Z, r)$
2. sample $\sigma_i^2 \sim p(\sigma_i^2|\mu, \pi, Z, r)$ $i = 1, ..., k$
3. sample $\pi \sim p(\pi|\mu, \sigma^2, Z, r)$
4. sample $z_t \sim p(z_t|\mu, \sigma^2, \pi, r)$, $t = 1, ..., T$.

Step 1–4 are repeated many times and an initial set of the draws are discarded to minimize startup conditions and ensure the remaining sequence of the draws is from the converged chain.

Below we detail each of the Gibbs sampling steps. Conditional on z_t we can recast the model as

$$r_t = z_t\mu + u_t, \quad u_t \sim N(0, z_t\sigma^2) \tag{5.8}$$

To jointly sample from the conditional distribution of μ using Gibbs sampling results for the linear regression model, we transform to a homoskedastic model as in

$$y_t = x_t\mu + v_t, \quad v_t \sim N(0, 1) \tag{5.9}$$

with $y_t = r_t/\sqrt{z_t\sigma^2}$, $x_t = z_t/\sqrt{z_t\sigma^2}$. Now the conditional posterior of μ is multivariate normal and a draw is obtained as

$$\begin{aligned} \mu &\sim N(M, V^{-1}) && (5.10)\\ M &= V^{-1}(X^T y + B^{-1}b) && (5.11)\\ V &= X^T X + B^{-1}. && (5.12) \end{aligned}$$

where $b = [b_1 \cdots b_k]^T$, B is a matrix of zeros with diagonal terms B_{ii}, y_t is a row of the vector y, and x_t is a row vector of the matrix X. The conditional posterior of σ_j^2 is,

$$\sigma_j^2 \sim IG\left(\frac{v_j + T_j}{2}, \frac{\sum_{t=1}^T (r_t - \mu_j)^2 z_{t,j} + s_j}{2}\right), \quad j = 1, ..., k. \tag{5.13}$$

where $T_j = \sum_{t=1}^T z_{t,j}$. Only the observations attributed to component j are used to update the variance σ_j^2.

With the conjugate prior for π, we sample the component probabilities as,

$$\pi \sim \mathcal{D}(\alpha_1 + T_1, ..., \alpha_k + T_k). \tag{5.14}$$

Finally, to sample $z_{t,i}$, note that,

$$p(z_{t,i}|r, \mu, \sigma, \pi) \propto \pi_j \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2\sigma_i^2}(r_t - \mu_i)^2\right), \quad i = 1, ..., k, \tag{5.15}$$

which implies that they can be sampled as a Multinomial distribution for $t = 1, ..., T$.

It is well known that in mixture models the parameters are not identified. For example, switching all states Z and the associated parameters gives the same likelihood value. Identification can be imposed through prior restrictions. However, in our application, interest centers on the moments of the return distribution and not the underlying mixture parameters. The moments of returns are identified. If for example, we switch all the parameters of component 1 and 2 we still have the same premium value $\gamma = \sum_{i=1}^k \mu_i\pi_i$.

Therefore, we do not impose identification of the component parameters but instead compute the mean, variance, skewness and kurtosis using (4.3)-(4.8) after each iteration of the Gibb sampler. It is these posterior quantities that our analysis focuses on. In the empirical work, we found the Markov chain governing these moments to mix very efficiently. As such, 5000 Gibbs iterations, after a suitable burnin period provide accurate estimates.

5.3 Model Comparison

Finally, the Bayesian approach allows for the comparison and ranking of models by Bayes factors or posterior odds. Both of these require calculation of the marginal likelihood. This is defined as

$$p(r|M_i) = \int p(r|\mu, \sigma^2, \pi, M_i)p(\mu, \sigma^2, \pi|M_i)d\mu d\sigma^2 d\pi \tag{5.16}$$

where M_i indexes a particular model. For the class of models considered in this paper we can calculate an estimate of this marginal likelihood using output from the posterior simulator. The Bayes factor for model M_0 versus model M_1 is defined as $BF_{01} = p(r|M_0)/p(r|M_1)$. A Bayes factor greater than one is evidence that the data favor M_0. Kass and Raftery (1995) summarize the support for M_0 from the Bayes factor as: 1 to 3 not worth more than a bare mention, 3 to 20 positive, 20 to 150 strong, and greater than 150 as very strong.

6 Optimal Use of the Data

6.1 Accounting for Structural Breaks

In this section we outline a method to deal with potential structural breaks. Intuitively, if a structural break occurred in the past we would want to adjust our use of the old data in our estimation procedure since those data can bias our estimates and forecasts. To formally deal with this, we follow the methodology of Maheu and Gordon (2005) and assume that structural breaks are exogenous unpredictable events that result in a change in the parameter vector associated with the maintained model, in this case a mixture-of-Normals model of excess returns.

The structural break model is constructed from a series of identical parameterizations (mixture-of-Normals, k fixed) that we label *submodels*. What differentiates the submodels is the history of data that is used to form the posterior density of the parameter vector θ. As a result, θ will have a different posterior density for each submodel, and a different predictive density for excess returns. Each of the individual submodels assume that once a break occurs, past data are not useful in learning about the new parameter value, only future data can be used to update beliefs. Structural breaks are identified by the probability distribution on submodels. Since breaks are permitted out-of-sample, new submodels are continually introduced through time. As more data arrives, the posterior density of the submodel parameter is updated from its prior. This allows for an increasing number of structural breaks through time.

Submodels are differentiated by when they start and the number of data points they use. Since structural breaks can never be identified with certainty, Bayesian model averaging provides a predictive distribution, which accounts for past and future structural breaks, by integrating over each of the possible submodels weighted by their probabilities. New submodels only receive significant weights once their predictive performance warrants it. The model average optimally combines the past (potentially biased) data from before the estimated break point, which will tend to have less uncertainty about the premium due to sample length, with the less precise (but unbiased) estimates based on the more recent post-break data. This approach provides a method to combine submodels estimated over different histories of data, and assess how many historical observations should be used to estimate the premium at any point in time.

To begin, define the information set $I_{a,b} = \{r_a, ..., r_b\}$, $a \leq b$, with $I_{a,b} = \{\emptyset\}$, for $a > b$, and for convenience let $I_t = I_{1,t}$. Let M_i be a submodel that assumes a structural break occurs at time i.[10] As we have mentioned, under our assumptions the data $r_1, ..., r_{i-1}$ are not informative about the submodel parameter due to the structural break, while the subsequent data $r_i, ..., r_{t-1}$ are informative. If θ denotes the parameter vector, then $p(r_t|\theta, I_{i,t-1}, M_i)$ is the conditional data density for submodel M_i, given θ, and the information set $I_{i,t-1}$. Now consider the situation where we have the data I_{t-1} and we want to consider forecasting out-of-sample r_t. A first step is to construct the posterior density for each of the possible submodels. If $p(\theta|M_i)$ is the prior distribution for the parameter vector θ of submodel M_i, then the posterior density of θ for submodel M_i based on $I_{i,t-1}$ has the form,

$$p(\theta|I_{i,t-1}, M_i) \propto \begin{cases} p(r_i, ..., r_{t-1}|\theta, M_i)p(\theta|M_i) & i < t \\ p(\theta|M_i) & i = t, \end{cases} \tag{6.1}$$

$i = 1, ..., t$. In the first case, only data after the assumed break at time $i-1$ are used. For $i = t$ past data are not useful at all since a break is assumed to occur at time t, and therefore the posterior becomes the prior. Thus, at time $t-1$ we have a set of submodels $\{M_i\}_{i=1}^t$, which use different numbers of data points to produce predictive densities for r_t.[11] For instance, given $\{r_1, ..., r_{t-1}\}$, M_1 assumes no breaks in the sample and uses all the data $r_1, ..., r_{t-1}$ for estimation and prediction; M_2 assumes a break at $t = 2$ and uses $r_2, ..., r_{t-1}$; ....; M_{t-1}, assumes a break at $t-1$ and uses r_{t-1}; and finally M_t assumes a break at t and uses no data. Thus M_t assumes a break occurs out-of-sample, in which case, past data is not useful. In the usual way the predictive density for submodel M_i is

[10] The exception to this is the first submodel of the sample M_1 for which there is no prior data.

[11] In our application, submodels are differentiated only by the assumption of when a break occurred. In addition to this, it is possible to allow for different families of submodels. However, there may not be a common interpretation of θ among different specifications.

formed by integrating out the parameter uncertainty,

$$p(r_t|I_{i,t-1}, M_i) = \int p(r_t|I_{i,t-1}, \theta, M_i)p(\theta|I_{i,t-1}, M_i)d\theta, \quad i = 1, ..., t. \tag{6.2}$$

For M_t the posterior is the prior under our assumptions.

Up to this stage we have said nothing about how to combine these submodels. First note that the usual Bayesian methods of model comparison and combination are based on the marginal likelihood of a common set of data. This cannot be used to compare the submodels $\{M_i\}_{i=1}^t$, since they are based on different histories of data. Therefore we require a new method to combine the submodels. In keeping with our interpretation of a structural break, we assume the occurrence of past structural breaks does not indicate anything about the occurrence of future structural breaks.[12] As such, we only have a subjective prior on the likelihood of a break.[13]

Consistent with this, the financial analyst places a subjective prior $0 \le \lambda_t \le 1$, $t = 1, ..., T$ that a structural break occurs at time t. A value of $\lambda_t = 0$ assumes no break at time t, and therefore submodel M_t is not introduced. This now provides a mechanism to combine the submodels.

To develop some intuition, we consider the construction of the structural break model for the purpose of forecasting, starting from a position of no data at $t = 0$. If we wish to forecast r_1, all we have is a prior on θ. We can obtain the predictive density using (6.2) which gives $p(r_1|I_0) = p(r_1|I_0, M_1)$ and, after observing r_1, we have $P(M_1|I_1) = 1$. Now allow for a break at $t = 2$, with $\lambda_2 \neq 0$, the predictive density is the mixture

$$p(r_2|I_1) = p(r_2|I_{1,1}, M_1)p(M_1|I_1)(1 - \lambda_2) + p(r_2|I_{2,1}, M_2)\lambda_2.$$

The first term is the predictive density using all data times the probability of no break. The second term is the predictive density derived from the prior assuming a break, times the probability of a break.[14] After observing r_2 we can update submodel probabilities,

$$\begin{aligned} P(M_1|I_2) &= \frac{p(r_2|I_{1,1}, M_1)p(M_1|I_{1,1})(1 - \lambda_2)}{p(r_2|I_1)} \\ P(M_2|I_2) &= \frac{p(r_2|I_{2,1}, M_2)\lambda_2}{p(r_2|I_1)}. \end{aligned}$$

Now we require a predictive distribution for r_3 given past information. Again, allowing for a break at time $t = 3$, $\lambda_3 \neq 0$, the predictive density is formed as

$$p(r_3|I_2) = [p(r_3|I_{1,2}, M_1)p(M_1|I_2) + p(r_3|I_{2,2}, M_2)p(M_2|I_2)]\,(1 - \lambda_3) + p(r_3|I_{3,2}, M_3)\lambda_3.$$

In words, this is (predictive density assuming no break at $t = 3$)×(probability of no break at $t = 3$) + (predictive density assuming a break at $t = 3$)×(probability of a

[12] If we assumed past breaks told us something about future breaks, then λ_t could be estimated as a function of past data. We do not pursue this extension in this paper.

[13] Non-sample information may be important in forming the prior on breaks.

[14] Recall that in the second density $I_{2,1} = \{\emptyset\}$.

break at $t = 3$). Once again $p(r_3|I_{3,2}, M_3)$ is derived from the prior. The updated submodel probabilities are

$$P(M_1|I_3) = \frac{p(r_3|I_{1,2}, M_1)p(M_1|I_2)(1-\lambda_3)}{p(r_3|I_2)} \quad (6.3)$$

$$P(M_2|I_3) = \frac{p(r_3|I_{2,2}, M_2)p(M_2|I_2)(1-\lambda_3)}{p(r_3|I_2)} \quad (6.4)$$

$$P(M_3|I_3) = \frac{p(r_3|I_{3,2}, M_3)\lambda_3}{p(r_3|I_2)}. \quad (6.5)$$

In this fashion we sequentially build up the predictive distribution of the break model. As a further example of our model averaging structure, consider Figure 3 which displays a set of submodels available at $t = 10$, where the horizontal lines indicate the data used in forming the posterior. The forecasts from each of these submodels, which use different data, are combined (the vertical line) using the model probabilities. M_{11} represents the prior in the event of a structural break at $t = 11$. If there has been a structural break at say $t = 5$, then as new data arrive, M_5 will receive more weight as we learn about the regime change.

Intuitively, the posterior and predictive density of recent submodels after a break will change quickly as new data arrives and once their predictions warrent it they receive larger weights in the model average. Conversely, old submodels will only change slowly when a structural break occurs. Their predictions will still be dominated by the longer and older data prior to the structural break.

Given this discussion, and a prior on breaks, the general predictive density for r_t can be computed as the model average

$$p(r_t|I_{t-1}) = \left[\sum_{i=1}^{t-1} p(r_t|I_{i,t-1}, M_i)p(M_i|I_{t-1})\right](1-\lambda_t) + p(r_t|I_{t,t-1}, M_t)\lambda_t. \quad (6.6)$$

The first term on the RHS of (6.6) is the predictive density from all past submodels that assume a break occurs prior to time t. The second term is the contribution assuming a break occurs at time t. In this case, past data are not useful and only the prior density is used to form the predictive distribution. The terms $p(M_i|I_{t-1})$, $i = 1, ..., t-1$ are the submodel probabilities, representing the probability of a break at time i give information I_{t-1}, and are updated each period after observing r_t as

$$p(M_i|I_t) = \begin{cases} \frac{p(r_t|I_{i,t-1}, M_i)p(M_i|I_{t-1})(1-\lambda_t)}{p(r_t|I_{t-1})} & 1 \le i < t \\ \frac{p(r_t|I_{t,t-1}, M_t)\lambda_t}{p(r_t|I_{t-1})} & i = t. \end{cases} \quad (6.7)$$

In addition to being inputs into (6.6) and other calculations below, the submodel probabilities also provide a distribution at each point in time of the *most recent structural break* inferred from the current data. Recall that submodels are indexed by their starting point. Therefore, if model $M_{t'}$ receives a high posterior weight given I_t with $t > t'$, this is evidence of the most recent structural break at t'.

Posterior estimates and model probabilities must be built up sequentially from $t = 1$ and updated as a new observation becomes available. At any given time, a posterior moment $g(\theta)$ which accounts for past structural breaks can be computed as,

$$E[g(\theta)|I_t] = \sum_{i=1}^{t} E[g(\theta)|I_{i,t}, M_i]p(M_i|I_i). \tag{6.8}$$

This is an average at time t of the model-specific posterior expectations of $g(\theta)$, weighted by the appropriate submodel probabilities. Submodels that receive large posterior probabilities will dominate this calculation.

Similarly, to compute an out-of-sample forecast of $g(r_{t+1})$ we include all the previous t submodels plus an additional submodel which conditions on a break occurring out-of-sample at time $t+1$ assuming $\lambda_{t+1} \neq 0$. The predictive mean of $g(r_{t+1})$ is

$$E[g(r_{t+1})|I_t] = \left[\sum_{i=1}^{t} E[g(r_{t+1})|I_{i,t}, M_i]p(M_i|I_t)\right](1-\lambda_{t+1}) + E[g(r_{t+1})|I_{t+1,t}, M_{t+1}]\lambda_{t+1} \tag{6.9}$$

Note that the predictive mean from the last term is based only on the prior as past data before $t+1$ are not useful in updating beliefs about θ give a break at time $t+1$.

In this paper, our main concern is with the equity premium. Using the mixture-of-Normals specification as our submodel with k fixed, this is $\gamma = \sum_{i=1}^{k} \mu_i \pi_i$. Given I_{t-1} we can compute the posterior distribution of the premium as well as the predictive distribution. It is important to note that even though our mixture of Normals submodel is not dynamic, allowing for a structural break at t differentiates the posterior and predictive distribution of the premium. Since we are concerned with forecasting the premium, we report features of the *predictive distribution of the premium* for period t given I_{t-1} defined as,

$$p(\gamma|I_{t-1}) = \left[\sum_{i=1}^{t-1} p(\gamma|I_{i,t-1}, M_i)p(M_i|I_{t-1})\right](1-\lambda_t) + p(\gamma|I_{t,t-1}, M_t)\lambda_t. \tag{6.10}$$

This equation is analogous to the predictive density of returns (6.6). From the Gibbs sampling output for each of the models we can compute the mean of the predictive distribution of the equity premium as,

$$E[\gamma|I_{t-1}] = \left[\sum_{i=1}^{t-1} E[\gamma|I_{i,t-1}, M_i]p(M_i|I_{t-1})\right](1-\lambda_t) + E[\gamma|I_{t,t-1}, M_t]\lambda_t. \tag{6.11}$$

In a similar fashion, the standard deviation of the predictive distribution of the premium can be computed from $\sqrt{E[\gamma^2|I_{t-1}] - (E[\gamma|I_{t-1}])^2}$. This provides a measure of uncertainty about the premium.

We can now clarify two of the estimators discussed in Section 3. Recall that $\hat{\gamma}_{ALL}$ uses all available data (submodel M_1) while $\hat{\gamma}_B$ optimally uses data after accounting for structural breaks. These are,

$$\hat{\gamma}_{ALL,t-1} = E[\gamma|I_{t-1}, M_1] \tag{6.12}$$
$$\hat{\gamma}_{B,t-1} = E[\gamma|I_{t-1}] \tag{6.13}$$

where the latter estimator integrates out all model uncertainty surrounding structural breaks through (6.11).

Finally, after estimation we can provide an estimate of the number of historical observations that are used at any given time to estimate the excess return distribution and hence the equity premium. Since submodels M_i define the time of a break, if a break occurs at $i < t$ we would only want to use the $(t-i+1)$ data points $r_i, r_{i+1}, ..., r_t$ after the break to estimate the premium. In practice, we do not know with certainty when a break occurs. However, we can use the submodel probabilities to infer the *mean useful observations* (MUO_t) defined as

$$\mathrm{MUO}_t = \sum_{i=1}^{t} (t-i+1)p(M_i|I_t). \tag{6.14}$$

A time series plot of MUO_t against time will indicate the number of useful historical observations at each point in time. If there are no structural breaks, we would expect MUO_t to follow the 45 degree line. In situations when breaks have been inferred, the MUO_t may dip substantially below the 45 degree line.

6.2 Calculations

Estimation of each submodel at each point in time follows the Gibbs sampler detailed in Section 5. After dropping the first 500 draws of the Gibbs sampler, we collect the next 5000 which are used to estimate various posterior quantities. We also require the submodel probabilities to form an out-of-sample forecast of the equity premium using (6.11). To calculate the marginal likelihood of a submodel, following Geweke (1995) we use a predictive likelihood decomposition,

$$p(r_i, ..., r_t|M_i) = \prod_{j=i}^{t} p(r_j|I_{i,j-1}, M_i). \tag{6.15}$$

Given a set of draws from the posterior distribution $\{\theta^{(i)}\}_{i=1}^{N}$, where $\theta^{(i)} = \{\mu_1, ..., \mu_k, \sigma_1^2, ..., \sigma_k^2, p_1, ..., p_k\}$, for submodel M_i, conditional on $I_{i,t-1}$, each of the individual terms in (6.15) can be estimated consistently as[15]

$$p(r_t|I_{i,t-1}, M_i) \approx \frac{1}{N} \sum_{i=1}^{N} p(r_t|\theta^{(i)}, I_{i,t-1}, M_i). \tag{6.16}$$

[15]This method of estimating the predictive likelihood provides accuracy similar to other methods such as Gelfand and Dey (1994).

This is calculated at the end of each Gibbs run, along with features of the predictive density, such as premium forecasts for each submodel. For the mixture-of-Normals specification, the data density is,

$$p(r_t|\theta^{(i)}, I_{i,t-1}, M_i) = \sum_{j=1}^{k} p_j \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{1}{2\sigma_j^2}(r_t - \mu_j)^2\right). \tag{6.17}$$

The predictive likelihood of submodel M_i is used in (6.7) to update the submodel probabilities at each point in time, and to compute the individual components $p(r_j|I_{j-1})$ of the structural break model through (6.6) and hence the marginal likelihood of the structural break model as,

$$p(r_1, ..., r_t) = \prod_{j=1}^{t} p(r_j|I_{j-1}). \tag{6.18}$$

6.3 Selecting Priors on the Premium

An advantage of Bayesian methods is that it is possible to introduce prior information into the analysis. This is particularly useful in our context as finance practitioners and academics have strong beliefs regarding the equity premium. Theory indicates the premium must be positive and from the wide range of estimates Derrig and Orr (2004) survey the vast majority of the reported estimates are well below 10%. The average survey response from U.S. Chief Financial Officers for recent years is below 5% (Graham and Harvey (2005)).

There are several issues involved in selecting priors when forecasting in the presence of structural breaks. Our model of structural breaks requires a proper predictive density for each submodel. This is satisfied if our prior $p(\theta|M_i)$ is proper.[16] There are also problems with using highly diffuse priors, as it may take many observations for the predictive density of a new submodel to receive any posterior support. In other words, the rate of learning about structural breaks is affected by the priors. Based on this, we use proper informative priors.

A second issue is the elicitation of priors in the mixture model. While it is straightforward for the one-component case, it is not obvious how priors on the component parameters affect features of the excess return distribution when $k > 1$. For two or more components, the likelihood of the mixture model is unbounded which make noninformative priors inappropriate (Koop (2003)).

In order to select informative priors based on features of excess returns, we conduct a *prior predictive* check on the submodel (Geweke (2003)). That is, we analyze moments of excess returns simulated from the submodel. We repeat the following steps

[16]Some of the submodels condition on very little data. For instance, at time $t-1$ submodel M_t uses no data and has a posterior equal to the prior.

1. draw $\theta \sim p(\theta)$ from the prior distribution
2. simulate $\{\tilde{r}_t\}_{t=1}^T$ from $p(r_t|I_{t-1}, \theta)$
3. using $\{\tilde{r}_t\}_{t=1}^T$ calculate the mean, variance, skewness and kurtosis

Table 2 reports summary statistics for the first four moments of excess returns from repeating the steps 1–3 many times. The prior associated with these results is listed in the second panel of Table 3. The prior can account for a range of empirically realistic sample statistics of excess returns. The 95% density region of the sample mean is approximately $[0, 0.1]$. The two-component model with this prior is also consistent with a wide range of skewness and excess kurtosis. In selecting a prior for the single-component model we tried to match, as far as possible, the features of the two-component model. This prior is listed in the top panel of Table 2. All prior specifications enforce a positive equity premium.

Although it is possible to have different priors for each submodel we use the same calibrated prior for all submodels in our analysis. Lastly, we set the probability of a break $\lambda_t = 0.01$. This favors infrequent breaks and allows the model to learn when breaks occur. We could introduce a new submodel for every observation but this would be computationally expensive. Instead, we restrict the number of submodels to one every year of data.[17] That is, our benchmark prior introduces a new submodel only every 12 months with $\lambda_t = 0.01$ and otherwise set $\lambda_t = 0$. This implies an expected duration of 100 years between structural breaks in the equity premium. We discuss other results for different specifications in the next section.

7 Results

This section discusses the out-of-sample model forecasts for the equity premium starting from the first observation to the last. First, we present results for a one component mixture submodel, and then in subsection 7.1 results for a two component mixture submodel. A summary of the model specifications, including priors, is reported in Table 3. The main results for the one-component specification are found in Figures 4 to 6, panel A of Figures 7 to 9, and Figure 10.

The out-of-sample forecasts of the equity premium from the one-component specification are found in Figure 4. For comparison purposes, the mean of the predictive distribution of the premium is displayed for both the structural break model and a no-break alternative. These are the forecasts $\hat{\gamma}_{B,t-1}$, computed from equation (6.13) which optimally uses past data, and $\hat{\gamma}_{ALL,t-1}$, computed from equation (6.12) using all available data at time $t-1$. The premium forecasts are similar until the start of the 1930s where

[17]Our first submodel starts in February 1885. Thereafter, new submodels are introduced in February of each year until 1914, after which new submodels are introduced in June of each year due to the missing 4 months of data in 1914 (see Schwert (1990) for details).

they begin to diverge. Thereafter, the premium from the structural break model rises over the 1950s and 1960s with a maximum value of 8.23 in 1962:1. Toward the end of the sample the premium decreases to values lower than the no-break model. The final premium forecast at the end of the sample is 3.53 for the structural break model and 4.65 for the no-break model.

The second panel of this figure displays the standard deviation of the predictive distribution of the premium. This is a measure of the uncertainty of our premium estimate in panel A. For the no-break model, uncertainty about the equity premium forecast originates from parameter uncertainty only, while for the structural break model it comes from both parameter and submodel uncertainty. Here again there are differences in the two specifications. The model that uses all data and ignores structural breaks shows a steady decline in the standard deviation of the premium's predictive distribution as more data become available. That is, for a structurally stable model, as we use more data we become more confident about our premium forecast. However, the standard deviation of the premium's predictive distribution from the break model shows that this increased confidence is misleading if structural breaks occur. As the second panel of Figure 4 illustrates, when a break occurs our uncertainty about the premium increases.

Figure 5 plots the mean and standard deviation of the posterior distribution of submodels for each date. Note that the standard deviation is a measure of submodel uncertainty, one of the two sources of uncertainty about the premium. Recall that submodels are indexed with the time period they start at, and their submodel probabilities identify the most recent structural break. Therefore, for any time period, there is a discrete probability distribution of possible submodels defined through (6.7). The mean and standard deviation of this distribution of submodels are

$$\text{mean}_t = \sum_{i=1885}^{t} iP(M_i|I_t); \quad \text{stdev}_t = \sqrt{\sum_{i=1885}^{t} i^2 P(M_i|I_t) - \text{mean}_t^2}. \tag{7.1}$$

These moments are calculated for each time t given the information set I_t. This calculation is repeated from the start of the sample to the end, and represents the inference that is available in real time.

There is a gradual increase in submodel uncertainty, measured by the standard deviation of the posterior distribution of submodels, starting in 1891 and a subsequent lowering after the 1930s and 1940s. It is interesting to note that in the early 1930s it takes less than one year for the uncertainty to drop by 97% from the highest levels in 1929. This indicates decisive evidence of the most recent structural break identified at 1929:6 and very fast learning about this change.[18] This is supported by the fact that the posterior mean of the submodel distribution jumps to the 1929 submodel at this time. There is a small increase in uncertainty during the 1930s but the posterior mean centers

[18]Therefore, the increase in the total uncertainty about the premium after 1929, shown in Figure 4:B, is mainly due to parameter uncertainty.

the distribution around the 1940 submodel until 1969 after which there is an increase in submodel uncertainty.

Figures 7 to 9 display the submodel probabilities through time for three different subperiods for the one-component specification ($k = 1$) in panels A. Figure 6 shows the probability of some selected submodels over time. These correspond to a slice through the submodel axis in panels A of Figures 7 to 9. The latter are 3-dimensional plots of (6.7) which is the probability of the most recent break point given data up to time t. The axis labelled Submodel M_i refers to the submodels identified by their starting observation i. Recall that the number of submodels is increasing with time, with a new submodel introduced every 12 months. The submodel probabilities at a point in time can be seen as a perpendicular line from the Time axis.

As shown in panel A of Figure 7, in the early part of the sample the first submodel, 1885, has probability close to 1. There was some preliminary evidence of a break early in the sample. For example, by 1902, that is, using data from 1885 to 1902, the first submodel M_{1885} received a probability of only 0.24 while submodel M_{1893} had a probability of 0.51. However, by 1907 the evidence for a break in 1893 diminished to 0.078, while the original submodel M_{1885} strengthened to 0.64. Thus learning as new data arrive can play an important role in revising previous beliefs regarding possible structural breaks. Recall that these probability assessments are based on data available in real time. As such, they represent the inference available to financial analysts at the time.

The first submodel of the sample, M_{1885} continues to receive most of the support in the 1910s and 1920s until 1929. As previously mentioned, there is very strong evidence of a structural break at 1929:6. This submodel has a probability of 0.94 based on data to 1929:11 which indicates fast learning about a change in the distribution of excess returns. The change in regime during this time and the subsequent crash in October of 1929 is likely identified as a sharp increase in volatility. As shown in Figure 4, during the 1930s the premium forecast is very similar to the no-break model, suggesting that the identified break in the excess return distribution in 1929 is due to higher-order moments such as volatility.

As mentioned previously, there is an increase in submodel uncertainty during the 1930s. Using data up to 1937, there is some evidence of a break in 1934[19] and in 1937. However, the next major break occurs in 1940. Until 1974, this submodel receives most of the weight with a probability for most of the time in excess of 0.90.[20] As shown in Figure 4, the 1940 structural break results in clear differences in the equity premium forecasts for the break and no-break models. Accounting for structural breaks indicates a larger equity premium after 1940 and more uncertainty about the premium. Note that by the mid-1950s the premium is almost double that obtained from the no-break model.

In the early 1970s there is weak evidence of a break in 1969, however, this subsequently declines during the mid-1970s, while the evidence for M_{1940} strengthens. By the mid-1970s there is uncertainty about submodels associated with 1969, 1973, and 1974,

[19] $M_{1934:6}$ has probability of 0.77 using data to 1937:6

[20] By 1969:5 the submodel still has a probability of 0.94.

which all receive significant support. By the mid-1980s we have learned that the most likely point of a break was 1969.[21] The strength of evidence for the 1969 submodel as the most recent break point is about 0.5 for the whole decade of the 1980s.

During the latter part of the 1990s there is some evidence of a break at 1988, with weaker evidence for the most recent break at 1991 and 1992. By the end of the sample the results support a recent break occurring sometime from 1996-1998 with the submodels M_{1996}, M_{1997}, and M_{1998}, possessing a combined probability of 0.77. In summary, we identify major breaks in 1929 and 1940, with weaker evidence for structural breaks in 1969 and 1988, and possibly a recent break in 1996-98.

Our results highlight several important points. First, the identification of structural breaks in the premium depends on the data used, and false assessments may occur which are later revised when more data become available.[22] This is an important aspect of learning about structural breaks. Second, our evidence of submodel uncertainty indicates the problem with using only one submodel. In a setting of submodel risk, the optimal approach is to model average as done in (6.11). There is overwhelming evidence for the structural break specification as measured by the marginal likelihood values found in Table 3 for the one-component models. A Bayes factor for the break model against the no-break model is around exp(155).

Finally, our discussion suggests that to forecast the premium we should not use all the data equally. The mean useful observations are displayed in Figure 10. The 45-degree line is the model that uses all data. Consistent with our discussion, the structural break model uses most of the data until around 1930 where the number of useful observations drops dramatically. Around 1940 the useful observations begin to steadily increase till further declining in the 1970s, 80s and 90s. In this figure, a rolling window model would be represented as a horizontal line. For example, a rolling window premium estimate using the most recent 10 years of data would be a horizontal line at 120. According to our model, this estimate would not be optimal during any historical time period.

7.1 Robustness

We now turn to the two-component submodel. Recall that this specification allows for higher-order moments in the distribution of excess returns. The results for this specification are found in panels B of Figures 7 to 9 and in Figures 11 to 13. The predictive mean for the equity premium, the standard deviation of the predictive distribution, and the mean useful observations are all broadly consistent with the one-component results. The two-component specification also identifies breaks in 1929 and 1940, and agrees with the previous analysis concerning a recent break in the late 1990s.

Table 3 records the marginal likelihood values of each of the models with and without

[21] For instance, M_{1968}, M_{1969}, M_{1973}, and M_{1974} receive probabilities of 0.13, 0.48, 0.06, and 0.01, respectively, based on data up to 1985:1.

[22] However, this false assessment of a structural break is still the optimal result given the data at hand.

breaks. Both the $k = 1$ and the $k = 2$ specifications provide strong evidence of structural breaks. However, the two-component break model has a log marginal likelihood value about 20 points larger than the one-component break model. According to the criteria in Section 5.3, this is very strong support for the two-component specification.

Figure 13 displays the posterior mean of the variance, skewness, and kurtosis of the excess returns distribution at each point in the sample using only information available to that time period. Since the skewness estimates are all less than zero and the kurtosis estimate is always greater than 3, there is clear evidence of higher-order moments that are inconsistent with the one-component specification for excess returns.

Panel B of Figures 7 to 9 display the submodel probabilities through time for the two-component specification. Note that this richer specification is much more decisive in favor of the 1885 submodel than the one-component version in panel A of Figure 7. Figure 9 also suggests that the simpler one-component specification tends to put more weight on more recent submodels. As mentioned earlier, these differences could be due to the fact that the two-component specification is more robust to fat tails (outliers) that, particularly with short samples, can be temporarily identified as probable structural breaks in the more restrictive one-component specification.

The modeling of asymmetries and fat tails results in some differences in submodel probabilities, and hence premium forecasts, mainly near the end of the sample. A comparison of the posterior mean and standard deviation of the distribution of submodels through time for $k = 1$, and $k = 2$ is shown in Figure 14. Both specifications are similar until the 1980s. Here the two-component specification always gives more probability to the 1940 submodel in the range of 0.04-0.15, while the one-component version essentially dismisses this from consideration and weights the submodel associated with 1969 much higher. In the 1990s, the probability of submodel 1940 increases steadily, so that by 1999 M_{1940} has a probability of 0.503.[23] The two-component specification, which can better accommodate outliers by capturing the fat tails and asymmetries in returns, places much more weight on submodel M_{1940}. This example underscores the importance of accurately modeling financial returns prior to an analysis of structural breaks.[24] There is still submodel uncertainty at the end of the sample consistent with a recent structural break. The final significant submodel probabilities, based on the full sample of data, are $M_{1940:6}$ 0.11, $M_{1998:6}$ 0.17, $M_{1999:6}$ 0.16, and $M_{2000:6}$ 0.14. The probability of a break in 1998-2000 is 0.47. The final forecast for the long-run equity premium, which averages over these submodels, is 4.02 percent.

As a further check on our results, Table 3 reports the marginal likelihood values for models which only allow for a structural break every 5 years as opposed to every year. The results favor allowing for structural breaks more frequently.

For the reasons discussed, we favor the structural break model with two-component mixture submodels as our preferred model in forecasting the premium. Our final compar-

[23]Submodel M_{1940} is not displayed in Figure 9.

[24]In other words, misspecified models may provide evidence of structural breaks when the underlying DGP is stable.

ison of the premium estimates from the alternative specifications is shown in Figure 15. Except for the end of the sample, the premium estimates are similar. However, other features of the predictive distribution of the premium do differ. For example, compare the standard deviations in panel B of Figures 4 and 11.

Also included in this figure is a 10-year rolling window based on the sample average. As we discussed above, and as shown in Figure 12, this ad hoc approach to dealing with structural breaks is nowhere optimal for the time period we consider. In addition, the simple rolling-window sample average is too volatile to produce realistic results. In some periods the sample average is negative while in other periods it is frequently in excess of 10%.

Although our figures show large differences in the premium forecasts with and without breaks, a natural question is how important these differences are for economic questions. As a simple example, consider a pension fund manager who must make a payment of \$1 twenty years from now. How much does the manager need to invest today in order to expect to meet this future liability? Based on current information, and assuming a zero riskfree rate, the investment required today is $E_t[1/(1+\gamma)^{20}]$, where the expectation is taken with respect to the predictive density of the equity premium at each point in time.[25] This is calculated by taking 1000 draws from the predictive distribution of the premium γ and calculating $1/(1+\gamma)^{20}$ for each. The average of these is the expected required investment. Figure 16 displays the required investment by the pension fund manager for each month through the whole sample for both models. Changes in the nobreak estimate only reflect learning about the model parameter as new data arrives while changes in the break model estimate reflect both learning about model parameters and structural breaks. In general, the shape of the predictive density for the premium affects the calculation of the required investment. This figure shows considerable differences after the first major break in 1929. For example, in 1950:1 the pension fund manager would need to invest 28% less under the structural break model to meet future liabilities.

Finally, it may be that structural breaks only affect the variance of excess returns. To better allow past data to contribute to premium forecasts after a structural break in volatility, we set the prior parameters for the premium in the one component specification to the previous posterior mean and variance of γ when a new submodel is introduced. Therefore, during any period a new submodel is introduced, the prior on γ begins centered on the most recent posterior for γ based on available data. The main difference in the premium forecasts for this case was that the premium was less variable and close to 6% from 1960 on, with a reduced standard deviation of the predictive distribution. However, the marginal likelihood is -1216.18 which is slightly worse than our original prior in Table 3 for $k = 1$, and still inferior to the $k = 2$ specification.

[25] Recall that the forward looking predictive density of the premium allows for breaks out-of-sample.

8 Conclusion

This paper makes several contributions to forecasting the long-run equity premium. First, we show that historical data are useful in updating our prior beliefs regarding the equity premium. In the presence of structural breaks, we provide an optimal approach to estimating and forecasting the equity premium using historical data on excess returns. Our evidence for structural breaks is strong and points toward at least 2 major breaks and possibly a more recent structural break. The paper has also shown the importance of submodel risk and the value of modeling higher-order moments of excess returns when inferring structural breaks and predicting the equity premium. Ignoring structural breaks leads to different premium estimates as well as overconfidence in the estimates.

Due to the presence of asymmetry and fat tails in excess returns, our statistical evidence clearly favors a mixture-of-Normals submodel specification with two components for the unconditional premium. For instance, the structural break model produces kurtosis values well above 3 and negative skewness throughout our sample of data. Interestingly, the premium forecasts (predictive mean) from the two-component model are quantitatively similar to the single-component model. Where they differ is in the shape of the predictive distribution of the premium. In general the two-component specification indicates that the predictive distribution of the premium is more disperse. This higher uncertainty associated with the equity premium will be important for investment decisions.

There is another important difference between the alternative specifications of the maintained submodel for the long-run equity premium. As we learn about the distribution governing excess returns, sometimes we infer a break that is later revised to be an outlier and not a structural break. The richer two-component submodel is more robust to these *false breaks*. One reason for this is that the two-component model is characterized by a high and low variance state. This allows for heteroskedasticity in excess returns. Therefore temporary outliers can be consistent with the maintained model and not evidence of a break in the distribution of excess returns.

Our evidence shows at least 2 major breaks (1929 and 1940), and possibly a more recent structural break in the late 1990s. We explicitly characterize the uncertainty with regard to break points which is clearly evident in our 3-dimensional plots (Figures 7 to 9) of the distribution of submodels.

Our model produces realistic forecasts of the premium over the entire 1885-2003 sample. The premium forecasts for the no-break and break alternatives are similar until the start of the 1930s where they begin to diverge. This divergence reflects the fact that the break model uses historical data optimally when breaks occur. In fact, the usefulness of historical data varies considerably over the sample. The premium from the structural break model rises over the 1950s and 1960s with a maximum value of 8.99 in 1961:12. Toward the end of the sample the premium decreases to values lower than the no-break model. The final premium forecast at the end of the sample is 4.02 for the structural break model and 5.10 for the no-break model.

Table 1: Summary Statistics for Annualized Monthly Excess Returns

Sample	Obs	Mean	Variance	Stdev	Skewness	Kurtosis
1885:02-2003:12	1423	0.0523	0.4007	0.6330	-0.4513	9.9871

Table 2: Sample Statistics for Excess Returns Implied by the Prior Distribution

	Mean	Median	Stdev	95% HPDI
γ	0.0369	0.0354	0.0320	(-0.0238, 0.1007)
γ_2	0.5808	0.5056	0.3312	(0.1519, 1.1786)
$\gamma_3/\gamma_2^{3/2}$	-0.3878	-0.3077	0.4718	(-1.4077, 0.3534)
γ_4/γ_2^2	8.1369	6.4816	5.9317	(2.7169, 18.7218)

This table reports summary measures of the empirical moments from the mixture model $k = 2$, when parameters are simulated from the prior distribution. First a draw from the prior distribution gives a parameter vector from which T observations of excess returns are simulated $\{\tilde{r}_t\}_{t=1}^T$. From these data we calculate the sample mean, variance, skewness and kurtosis of excess returns. This process is repeated a large number of times to produce a distribution of each of the excess return moments. Finally, from this empirical distribution we report the mean, median, standard deviation and the 95% highest posterior density interval (HPDI).

Table 3: Structural Break Model Specifications and Results

model	breaks	prior	log(ML)
$k=1$	$\lambda_t = 0$ none	$b = 0.03, B = 0.03^2$ $v = 9.0, s = 4.0$	-1371.22
$k=1$	$\lambda_t = 0.01$ every 5 years, otherwise $\lambda_t = 0$	$b = 0.03, B = 0.03^2$ $v = 9.0, s = 4.0$	-1235.33
$k=1$	$\lambda_t = 0.01$ every year, otherwise $\lambda_t = 0$	$b = 0.03, B = 0.03^2$ $v = 9.0, s = 4.0$	-1216.08
$k=2$	$\lambda_t = 0$ none	$b_1 = 0.05, b_2 = -0.30, B_{11} = 0.03^2, B_{22} = 0.05^2$ $v_1 = 10.0, s_1 = 3, v_2 = 8.0, s_2 = 20.0$ $\alpha_1 = 7, \alpha_2 = 1$	-1241.09
$k=2$	$\lambda_t = 0.01$ every 5 years, otherwise $\lambda_t = 0$	$b_1 = 0.05, b_2 = -0.30, B_{11} = 0.03^2, B_{22} = 0.05^2$ $v_1 = 10.0, s_1 = 3, v_2 = 8.0, s_2 = 20.0$ $\alpha_1 = 7, \alpha_2 = 1$	-1202.01
$k=2$	$\lambda_t = 0.01$ every year, otherwise $\lambda_t = 0$	$b_1 = 0.05, b_2 = -0.30, B_{11} = 0.03^2, B_{22} = 0.05^2$ $v_1 = 10.0, s_1 = 3, v_2 = 8.0, s_2 = 20.0$ $\alpha_1 = 7, \alpha_2 = 1$	-1196.30

This tables displays the number of components k, in the mixture model, the prior specification of the submodel parameters as well as the prior on the occurrence of structural breaks λ_t. Finally, the logarithm of the marginal likelihood is reported for all specifications based on the full sample of observations used in estimation.

Figure 1: Annualized Monthly Excess Returns

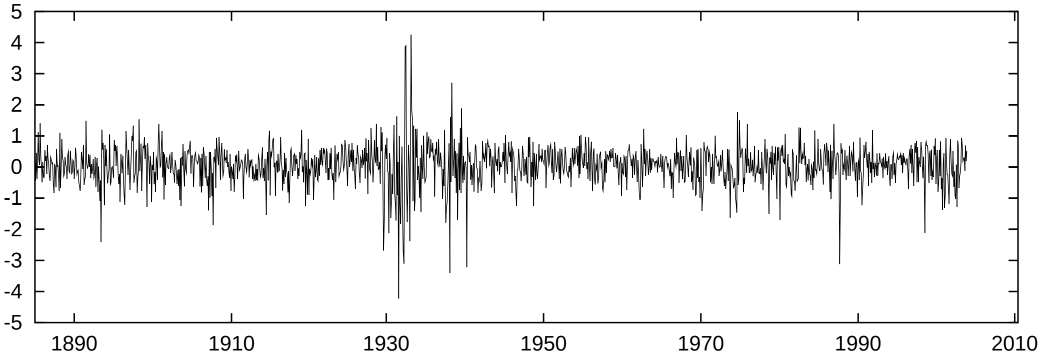
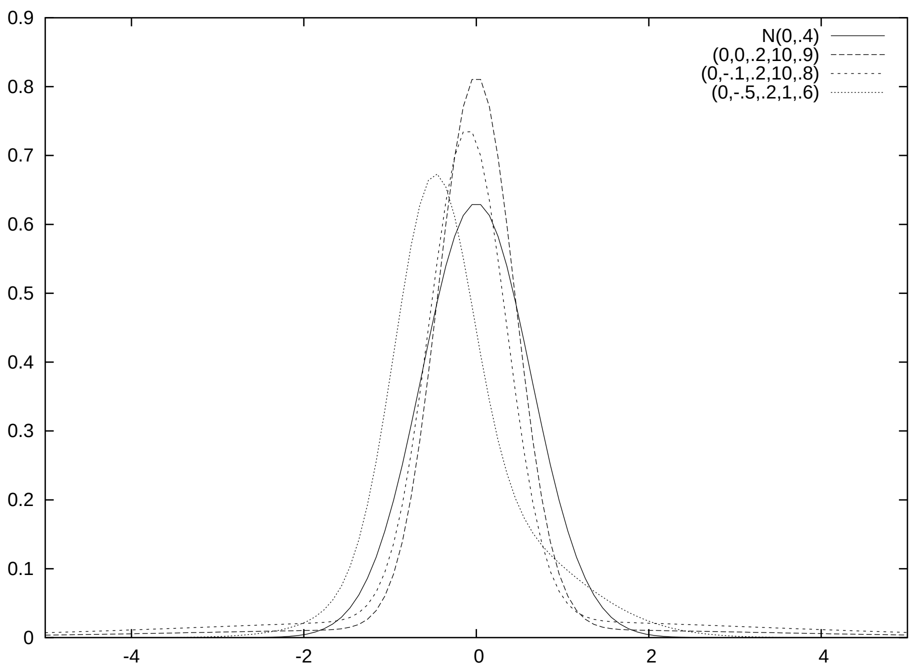


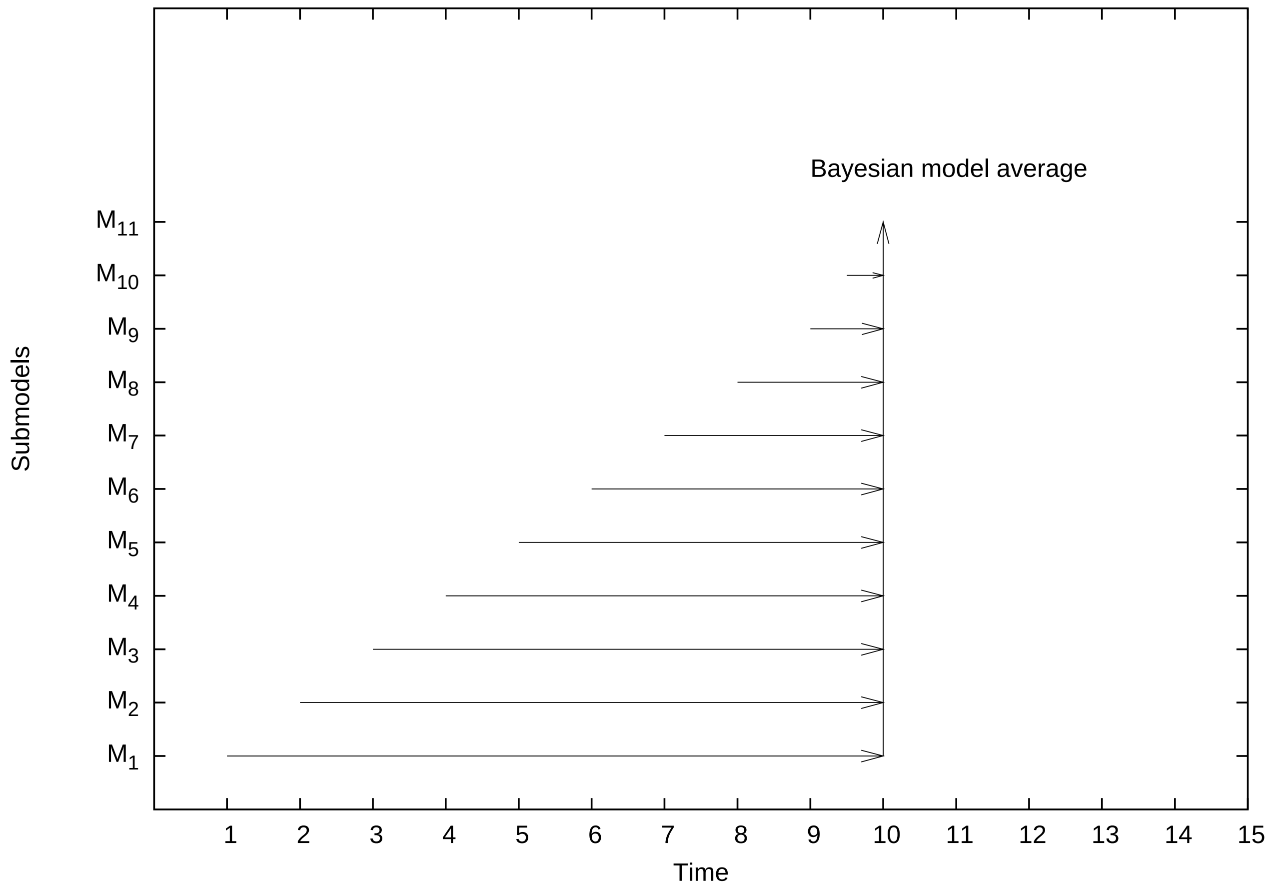
Figure 2: Some Examples of the Distribution From a Two-Component Mixture



This figure displays the density from various configurations of a mixture of two Normal densities. The parameters are $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, p_1)$ and correspond to the submodel in Section 4. Given the parameters the density is

$$\frac{p}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{1}{2\sigma_1^2}(r_t - \mu_1)^2\right) + \frac{1-p}{\sqrt{2\pi\sigma_2^2}} \exp\left(-\frac{1}{2\sigma_2^2}(r_t - \mu_2)^2\right).$$

Figure 3: Individual Submodels and the Bayesian Model Average



This figure is a graphical depiction of how the predictive density of excess returns is constructed for the structural break model. This corresponds to equation (6.6). The predictive density is computed for each of the submodels $M_1, ..., M_{10}$ given information up to $t = 10$. The final submodel M_{11}, postulates a break at $t = 11$ and uses no data but only a prior distribution. Each submodel is estimated using a smaller history of data (horizontal lines). Weighting these densities via Bayes rule (vertical line) gives the final predictive distribution (model average) of excess returns for $t = 11$.

Figure 4: Premium Forecasts through Time, $k = 1$.

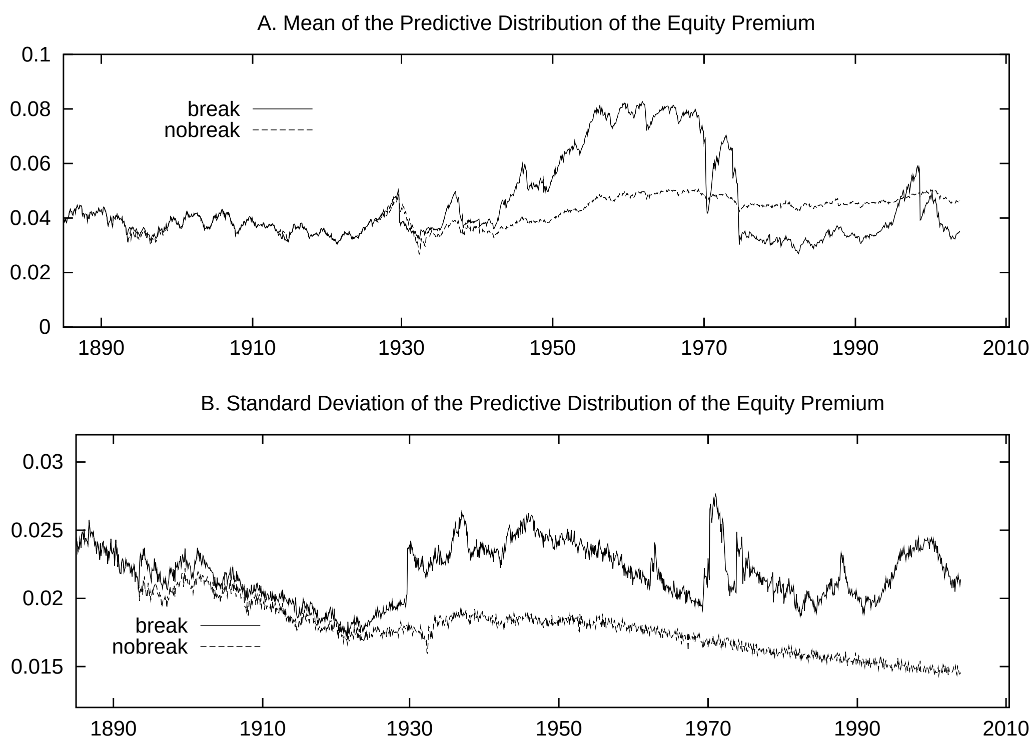
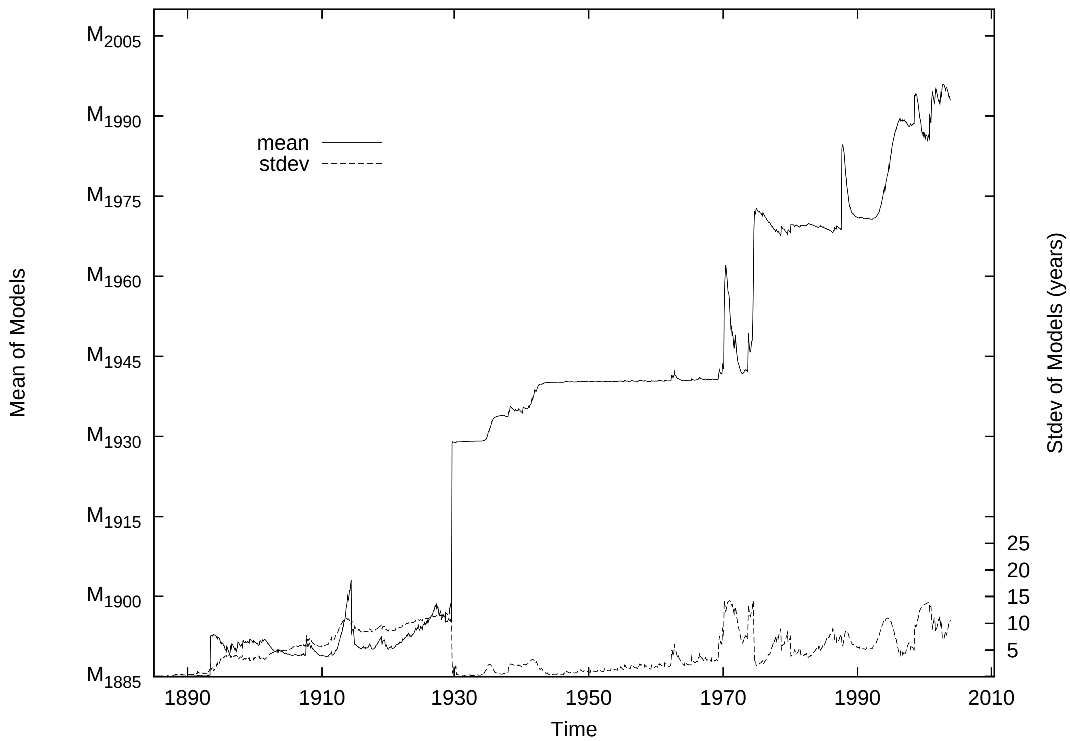


Figure A displays the out-of-sample forecasts (predictive mean) of the equity premium period by period for both the structural break model and the no break alternative. Figure B displays the corresponding standard deviation of the predictive distribution of the equity premium.

Figure 5: Posterior Mean and Standard Deviation of the Distribution of Submodels



This figure displays the posterior mean and the standard deviation of the distribution of submodels at each point in time. The moments are calculated from (6.7) for each observations $t = 1885:2 - 2003:12$, based on data up to and including t. The moments are

$$\text{mean}_t = \sum_{i=1885}^{t} iP(M_i|I_t); \quad \text{stdev}_t = \sqrt{\sum_{i=1885}^{t} i^2 P(M_i|I_t) - \text{mean}_t^2}$$

Submodels are indexed by the calendar time when they begin. The mean of the distribution of submodels is displayed on the vertical axis.

Figure 6: Submodel Probabilties over Time, $k = 1$

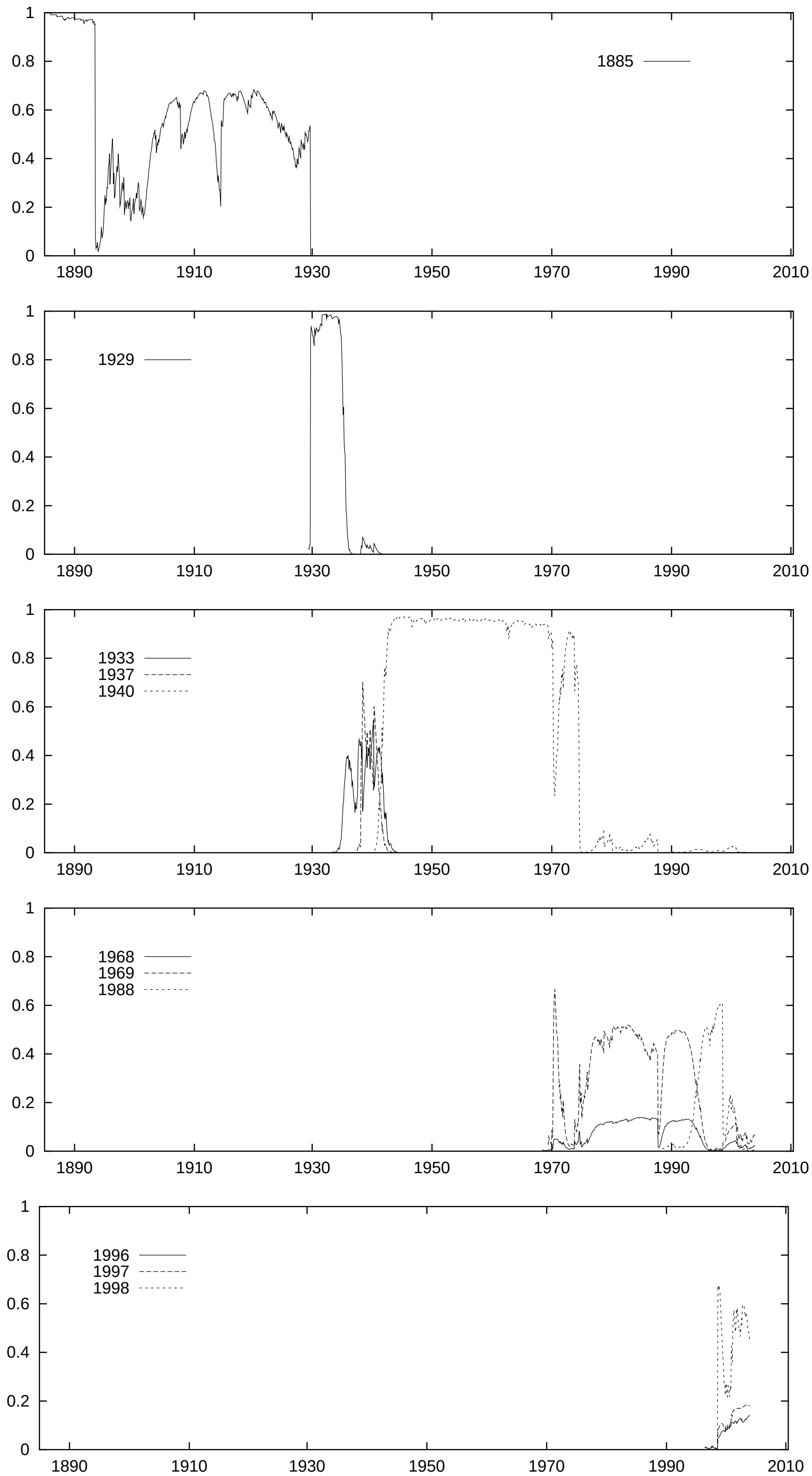
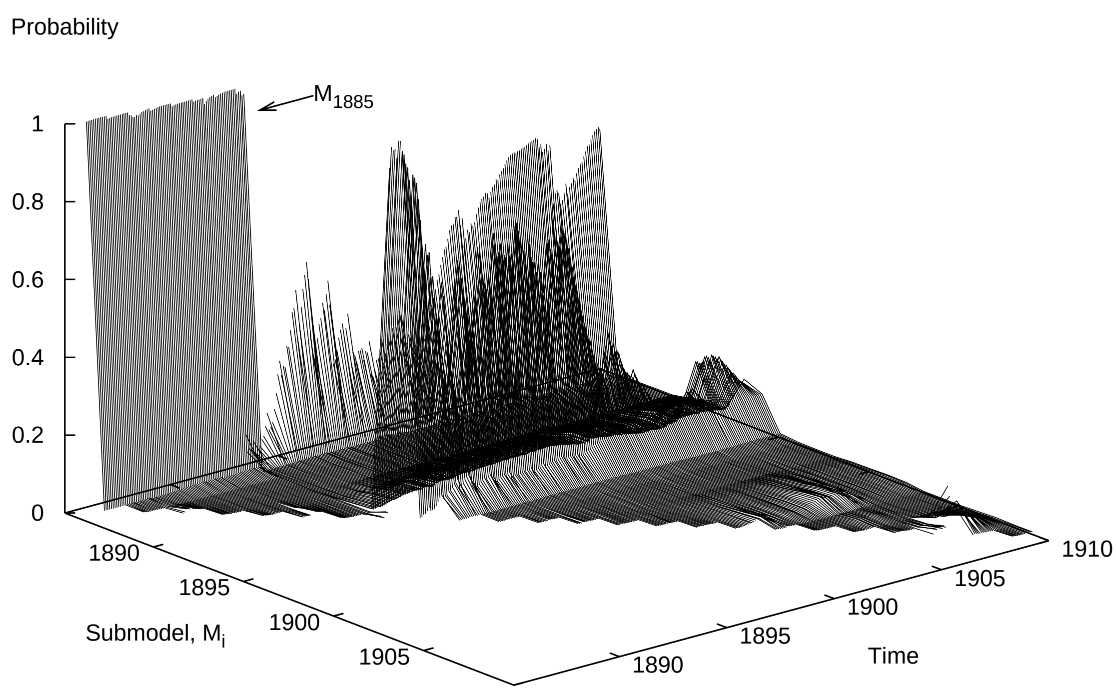


Figure 7: Submodel Probabilities through Time, 1885:2-1910:1

A. k=1



B. k=2

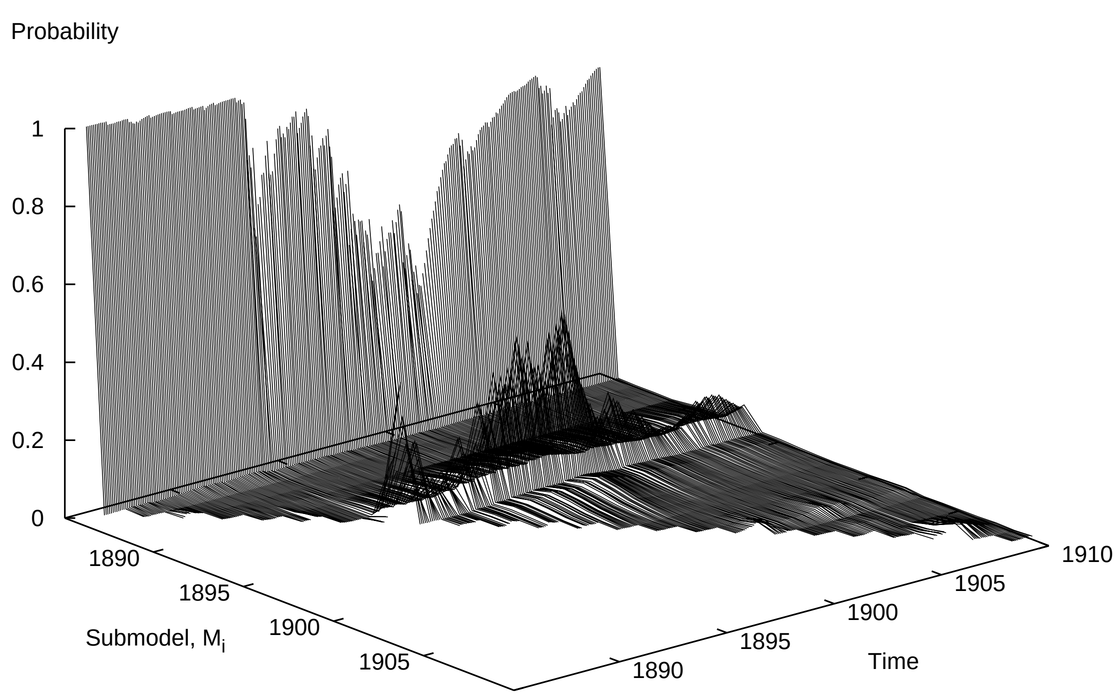
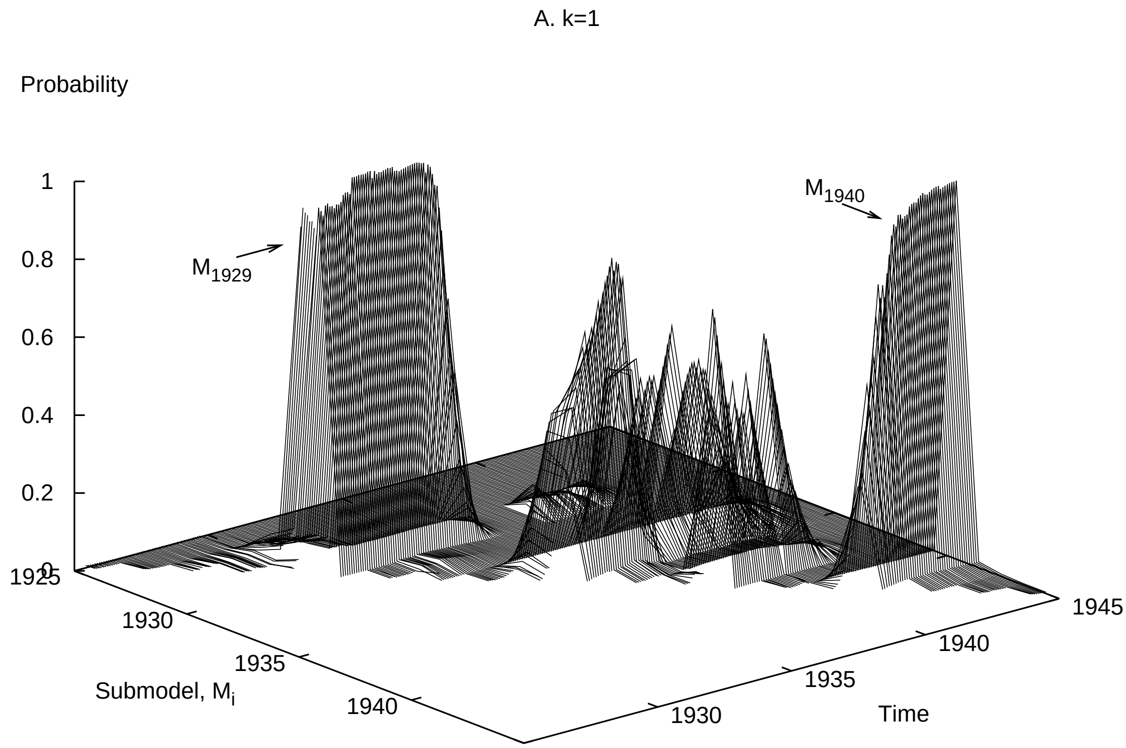


Figure 8: Submodel Probabilities through Time, 1925:1-1945:1



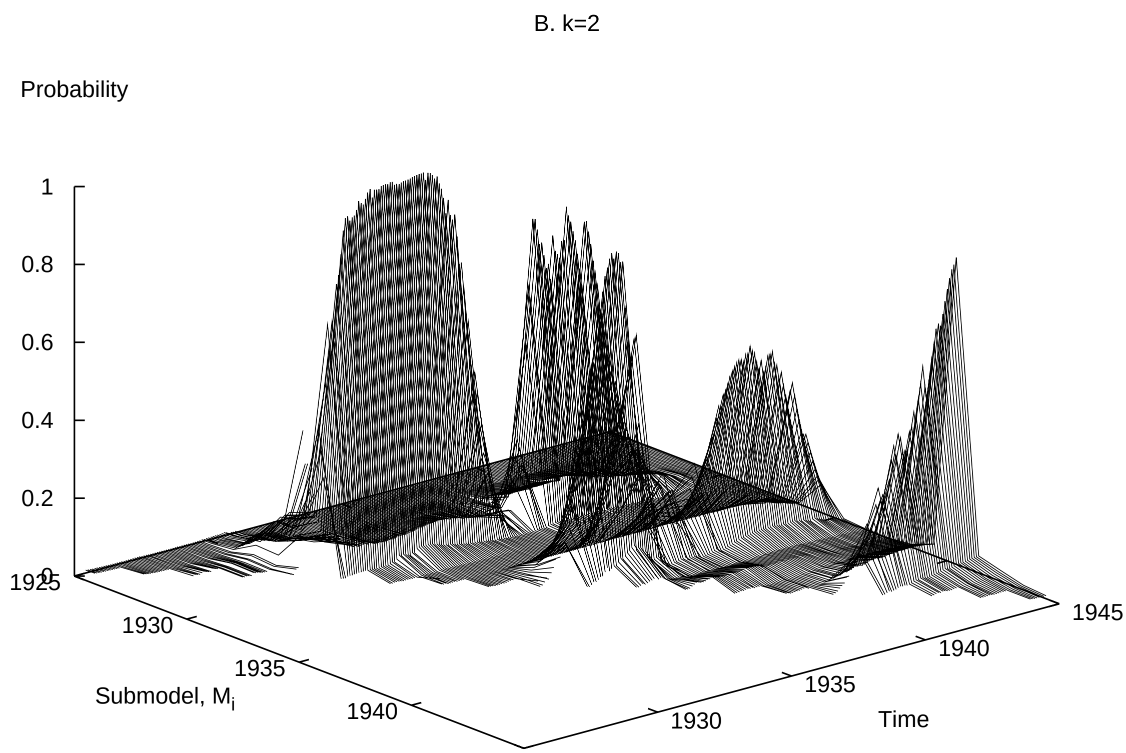
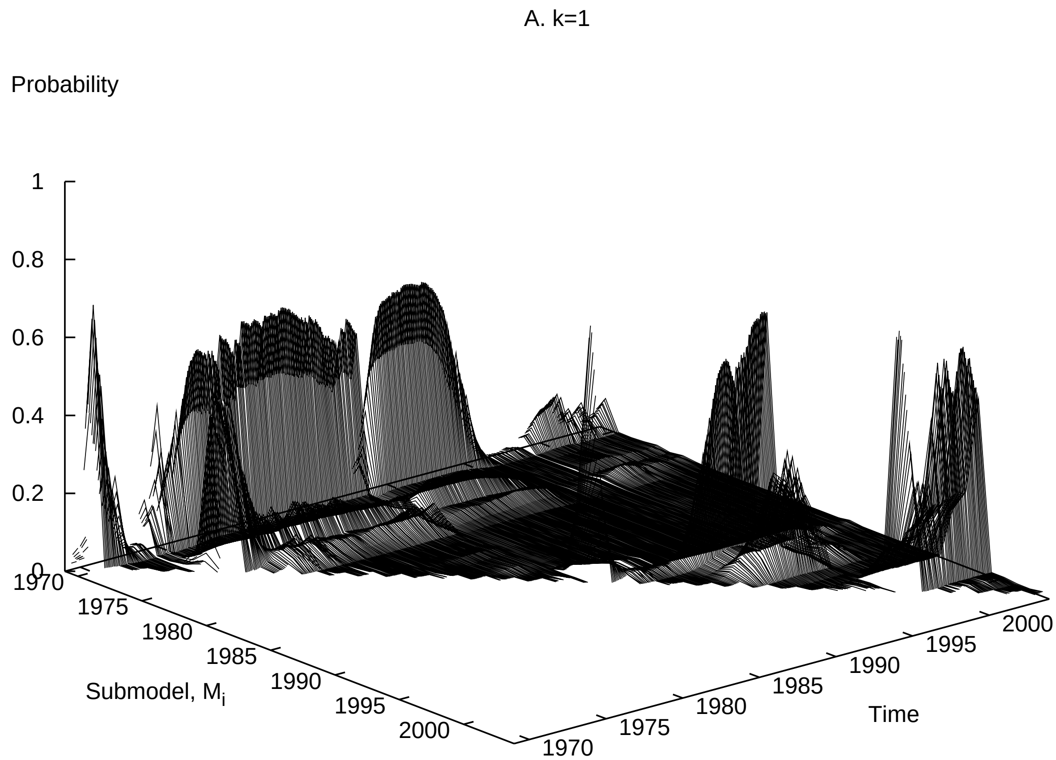


Figure 9: Submodel Probabilities through Time, 1970:1-2003:12



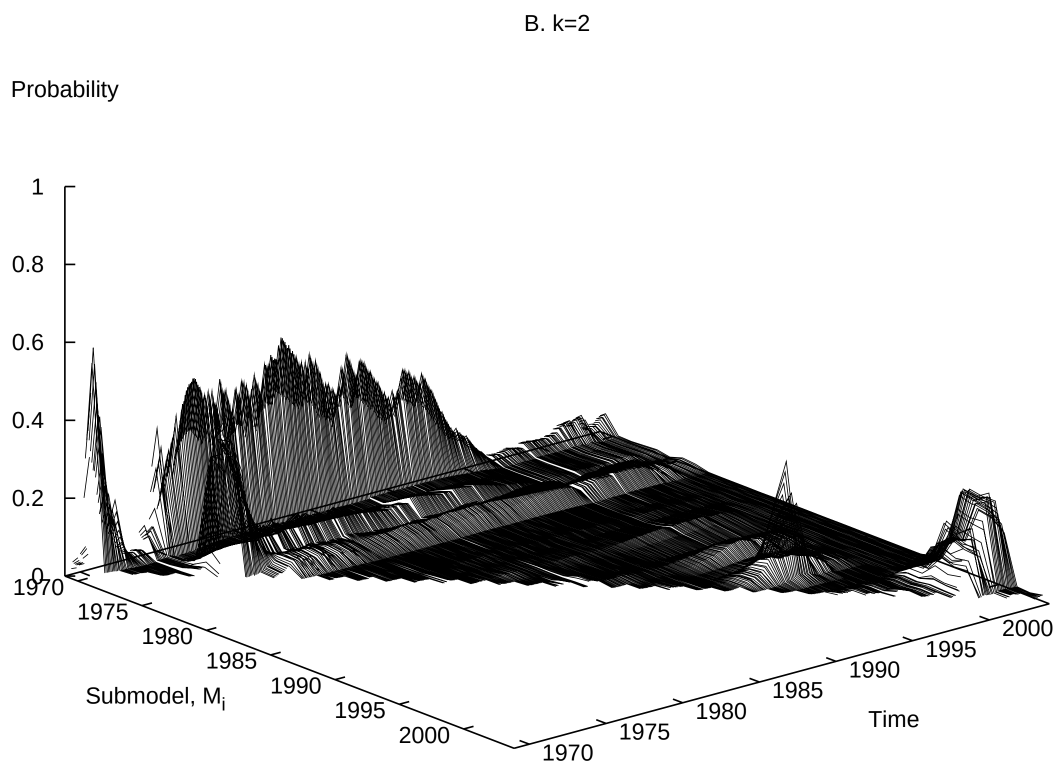
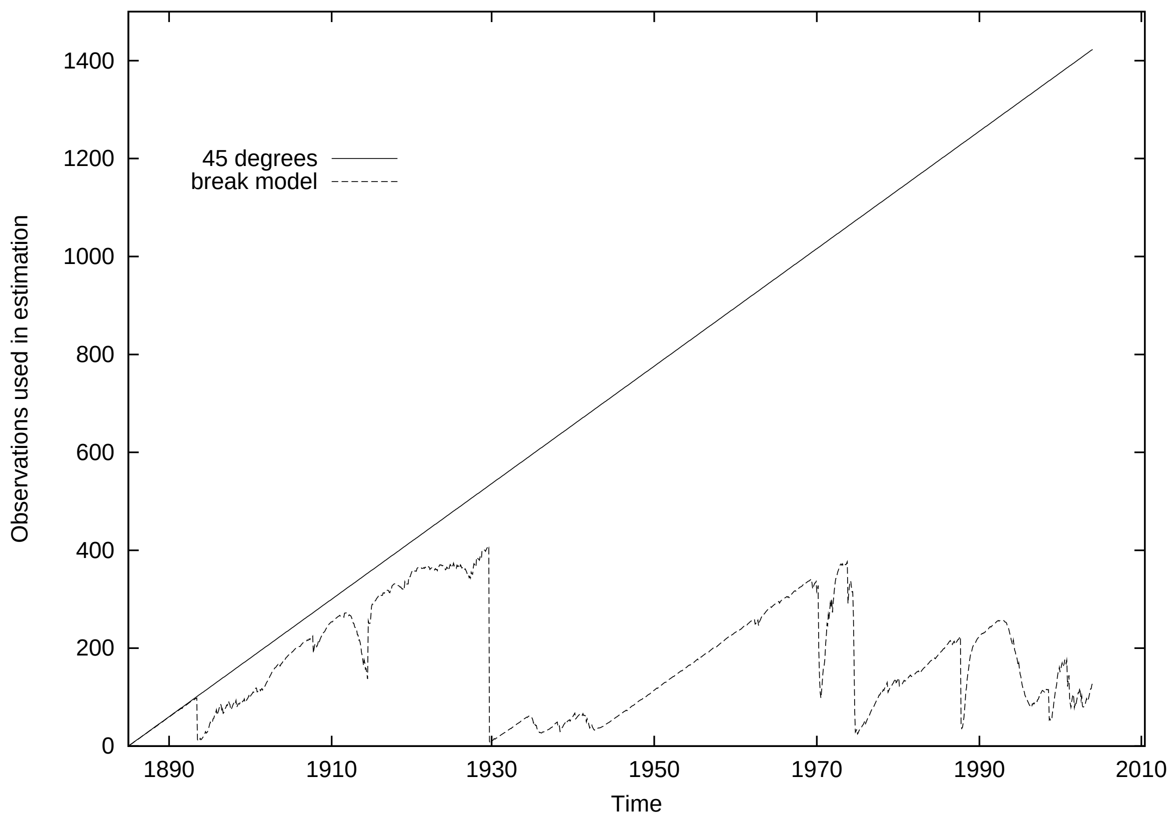


Figure 10: Mean useful Observations, $k = 1$



This figure shows the mean useful observations MUO_t defined as

$$\text{MUO}_t = \sum_{i=1}^{t} (t - i + 1) p(M_i | I_t).$$

which is the expected number of useful observation for model estimation at each point in time. $p(M_i|I_t)$ is the posterior submodel probability for M_i given the information set I_t. If there are no structural breaks then MUO_t would follow the 45 degree line.

Figure 11: Premium Forecasts through Time, $k = 2$.

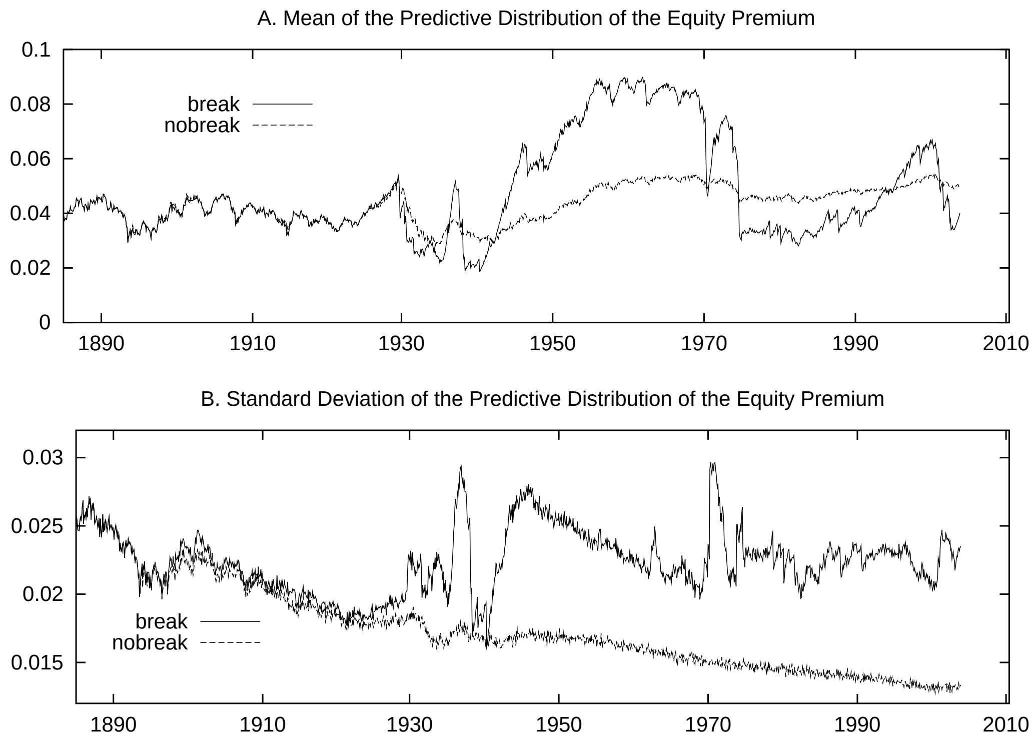
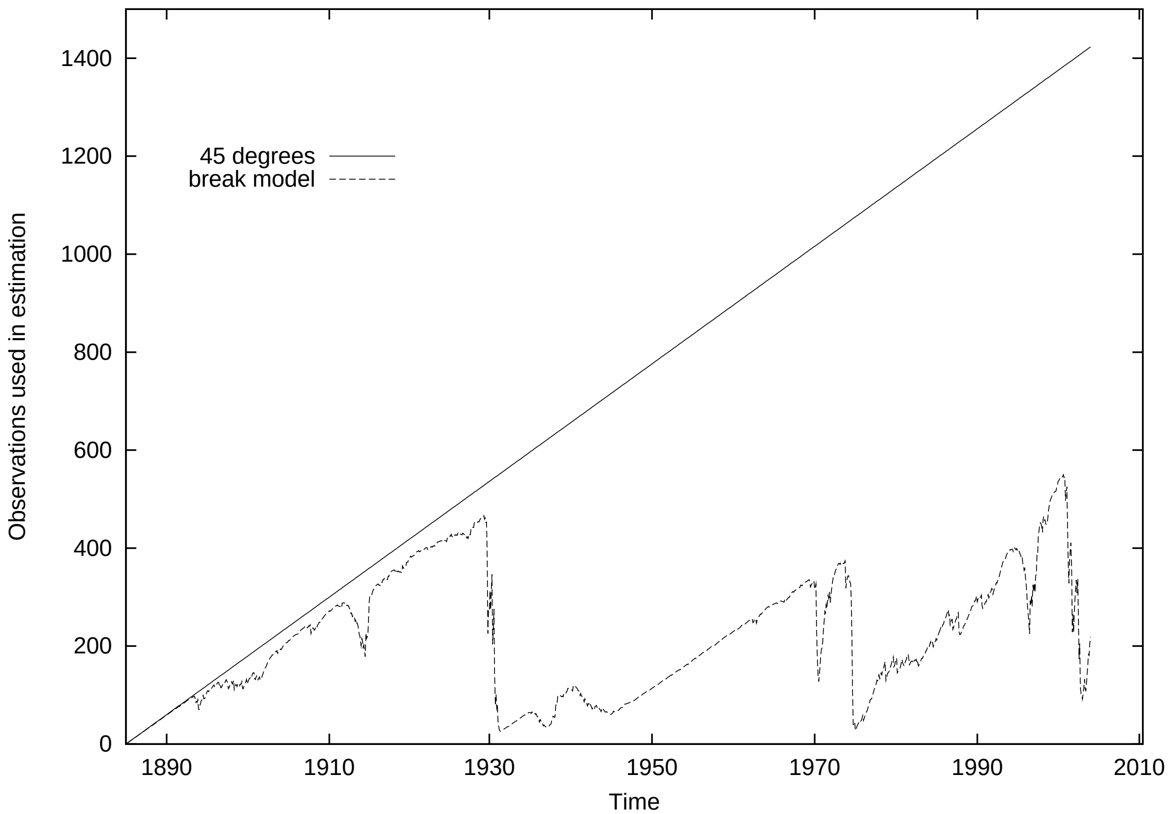


Figure A displays the out-of-sample forecasts (predictive mean) of the equity premium period by period for both the structural break model and the no break alternative. Figure B displays the corresponding standard deviation of the predictive distribution of the equity premium.

Figure 12: Mean useful Observations, $k = 2$.

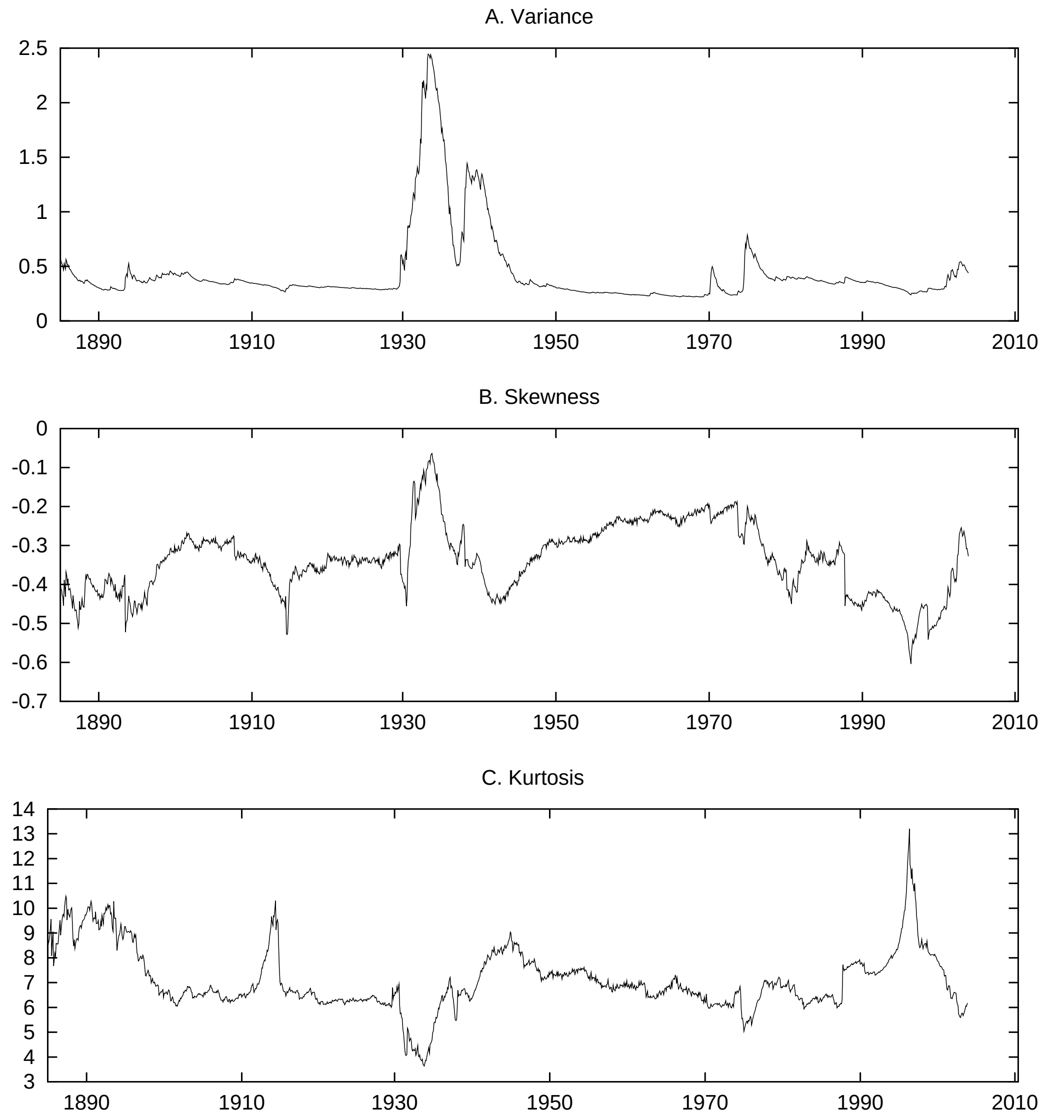


This figure shows the mean useful observations MUO_t defined as

$$\mathrm{MUO}_t = \sum_{i=1}^{t} (t - i + 1) p(M_i | I_t).$$

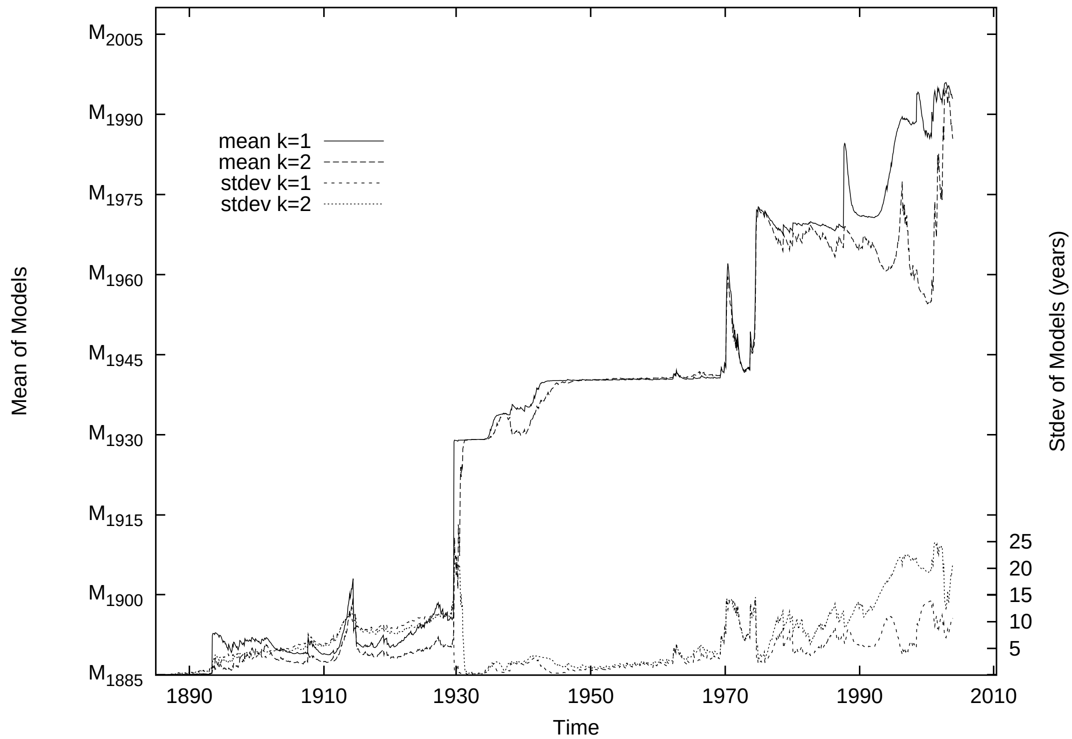
which is the expected number of useful observation for model estimation at each point in time. $p(M_i|I_t)$ is the posterior submodel probability for M_i given the information set I_t. If there are no structural breaks then MUO_t would follow the 45 degree line.

Figure 13: Higher-Order Moments of Excess Returns through Time



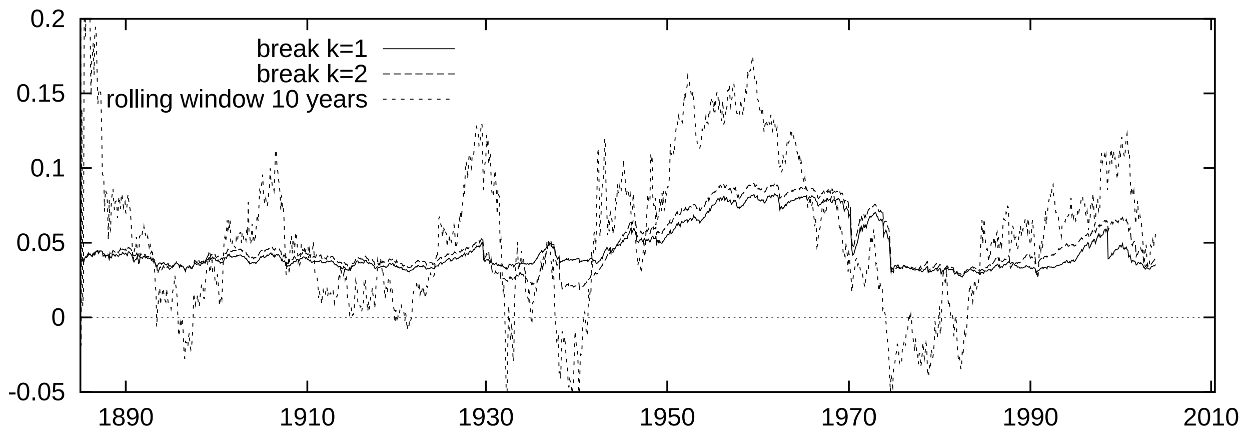
Displayed are the posterior means of the moments of the excess return distribution as inferred from the structural break model, $k = 2$. Each moment is estimated using only information in I_t at each point in time. The moments in (4.6)-(4.8) are computed for each Gibbs draw from the posterior distribution for each of the submodels M_i. The submodel specific moments are averaged using (6.8). This is repeated at each observation in the sample starting from $t = 1$. The evolution of the excess return moments reflect both learning (as more data arrive) and the effect of structural breaks.

Figure 14: Comparison of Posterior Mean and Standard Deviation of the Distribution of Submodels



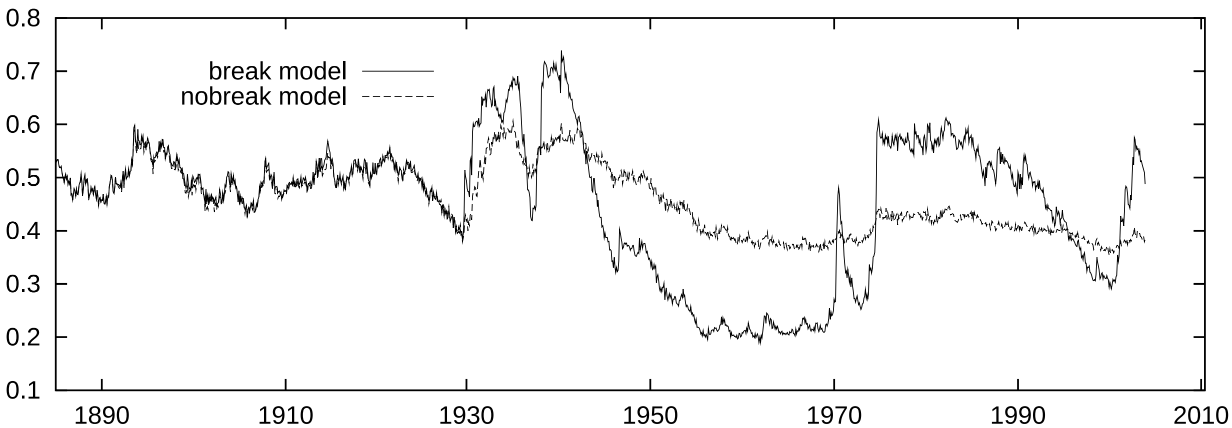
This figure compares the posterior mean and standard deviation of the distribution of submodels for $k = 1$, and 2 specifications. See the notes to Figure 5.

Figure 15: Comparison of Premium Forecasts



This figure compares the forecasts (predictive mean) of the equity premium from the structural break model with 1 and 2 components, along with the sample average that uses a rolling window of 10 years of data. The sample average at time t is defined as $\frac{1}{120}\sum_{i=1}^{120} r_{t-i+1}$.

Figure 16: Implications of Structural Breaks for a Pension Liability. $k = 2$



This figure compares the expected investment required today to receive $1 twenty years in the future. This is calculated as $E_t[1/(1+\gamma)^{20}]$ for both the break and no-break models at each point in time based on the most recent data available. The expectation is taken with respect to the predictive distribution of the equity premium γ, assuming a riskfree rate of 0.

References

Andreou, E., and E. Ghysels (2002): "Detecting Multiple Breaks in Financial Market Volatility Dynamics," *Journal of Applied Econometrics*, 17(5), 579–600.

Avramov, D. (2002): "Stock return predictability and model uncertainty," *Journal of Financial Economics*, 64(423-458).

Barberis, N. C. (2000): "Investing for the Long Run when Returns are Predictable," *Journal of Finance*, 55, 225–264.

Bernardo, J. M., and A. F. M. Smith (2000): *Bayesian Theory*. Wiley.

Campbell, J. Y., and S. B. Thompson (2004): "Predicting the Equity Premium Out of Sample: Can Anything Beat the Historical Average?," manuscript, Harvard University.

Chib, S. (1998): "Estimation and Comparison of Multiple Change Point Models," *Journal of Econometrics*, 86, 221–241.

Chib, S. (2001): "Markov Chain Monte Carlo Methods: Computation and Inference," in *Handbook of Econometrics*, ed. by Heckman, and Leamer. Elsevier Science.

Cremers, K. J. M. (2002): "Stock return predictability: A bayesian model selection perspective," *Review of Financial Studies*, 15(1223-1249).

Derrig, R. A., and E. D. Orr (2004): "Equity Risk Premium: Expectations Great and Small," *North American Actuarial Journal*, 8(1), 45–69.

Diebolt, J., and C. P. Robert (1994): "Estimation of Finite Mixture Distributions through Bayesian Sampling," *Journal of the Royal Stistical Society, Series B*, 56, 363–375.

Donaldson, R. G., M. Kamstra, and L. Kramer (2004): "Stare Down the Barrel and Center the Crosshairs: Targeting the Ex Ante Equity Premium," manuscript, York University.

Fama, E. F., and K. R. French (1992): "The Cross-Section of Expected Stock Returns," *Journal of Finance*, 47, 427–465.

——— (2002): "The Equity Premium," *Journal of Finance*, 57(2), 637–659.

Gelfand, A., and D. Dey (1994): "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of The Royal Statistical Society*, B, 56, 501–514.

Geweke, J. (1995): "Bayesian Comparison of Econometric Models," University of Iowa.

——— (1997): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18(1), 1–73.

——— (2003): "Contemporary Bayesian Econometrics and Statistics," Department of Economics and Statistics, University of Iowa.

Goyal, A., and I. Welch (2004): "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction," manuscript, Brown University.

Graham, J. R., and C. R. Harvey (2005): "The long-run equity risk premium," Fuqua School of Business, Duke University.

Jacquier, E., A. Kane, and A. Marcus (2005): "Optimal Estimation of the Risk Premium for the Long Run and Asset Allocation: A Case of Compounded Estimation Risk," *Journal of Financial Econometrics*, 3, 37–55.

Johannes, M., and N. Polson (2005): "MCMC Methods for Financial Econometrics," Handbook of Financial Econometrics, Elsevier.

Kandel, S., and R. Stambaugh (1996): "On the predictability of stock returns: an asset allocation perspective," *Journal of Finance*, 51, 385–424.

Kass, R. E., and A. E. Raftery (1995): "Bayes Factors," *Journal of the American Statistical Association*, 90(420), 773–795.

Kim, C.-J., J. C. Morley, and C. R. Nelson (2005): "The Structural Break in the Equity Premium," *Journal of Business & Economic Statistics*, 23(2), 181–191.

Koop, G. (2003): *Bayesian Econometrics*. Wiley, Chichester, England.

Maheu, J. M., and S. Gordon (2005): "Learning, Forecasting and Structural Breaks," manuscript, Department of Economics, University of Toronto.

Mayfield, E. S. (2004): "Estimating the Market Risk Premium," *Journal of Financial Economics*, 73(3), 465–496.

Mehra, R., and E. C. Prescott (2003): "The Equity Premium in Retrospect," in *Handbook of the Economics of Finance*, ed. by G. M. Constantinides, M. Harris, , and R. Stulz. North Holland, Amsterdam.

Pastor, L., and R. F. Stambaugh (2001): "The Equity Premium and Structural Breaks," *Journal of Finance*, 4, 1207–1231.

Pettenuzzo, D., and A. Timmermann (2004): "Optimal Asset Allocation under Structural Breaks," Department of Economics, University of California, San Diego.

Robert, C. P., and G. Casella (1999): *Monte Carlo Statistical Methods*. Springer, New York.

Roeder, K., and L. Wasserman (1997): "Practical Bayesian Density Estimation Using Mixtures of Normals," *Journal of the American Statistical Association*, 92(439), 894–902.

Ross, S. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341–360.

Schwert, G. W. (1990): "Indexes of U.S. Stock Prices from 1802 to 1987," *Journal of Business*, 63(3), 399–426.

Siegel, 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.

Turner, C., R. Startz, and C. Nelson (1989): "A Markov Model of Heteroskedasticity, Risk, and Learning in the Stock Market," *Journal of Financial Economics*, 25, 3–22.

MarketWatch

Yes, 100% of economists were dead wrong about yields

By Ben Eisen
Published: Oct 22, 2014 8:01 a.m. ET

Back in April every economist in a survey thought yields would rise. Guess what they did next

Getty Images
As it turns out, economists are not soothsayers.

NEW YORK (MarketWatch) — Just about six months ago, a headline flashed across the top of MarketWatch's home page. It read: "100% of economists think yields will rise within six months."

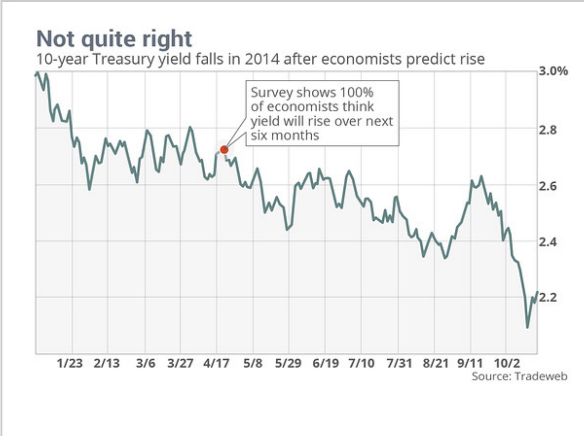
The April 22 report was based on a Bloomberg survey of 67 economists, all of whom expected the 10-year Treasury note 10_YEAR, +0.34% yield — which closed at 2.73% that day — to rise over the following half year.

"How quickly we would get to 4[%] was the discussion at the beginning of the year," said Mohamed El-Erian, chief economic adviser at Allianz SE, on CNBC Tuesday morning.

The market, however, has a funny way of leaning one way, just as the herd is heading in the other direction.

On Tuesday, the 10-year note traded at a yield of 2.21%, almost four-tenths of a percentage point lower than in April. Let's not forget that the yield unexpectedly dipped below 2%, just last week.

That underscores the difficulty of calling the direction of interest rates. It also makes all 67 economists wrong, as this chart of the benchmark yield shows:



Treasury yields tend to rise, and prices drop, as the U.S. economy grows and investors begin to expect the Federal Reserve to normalize monetary policy more quickly.

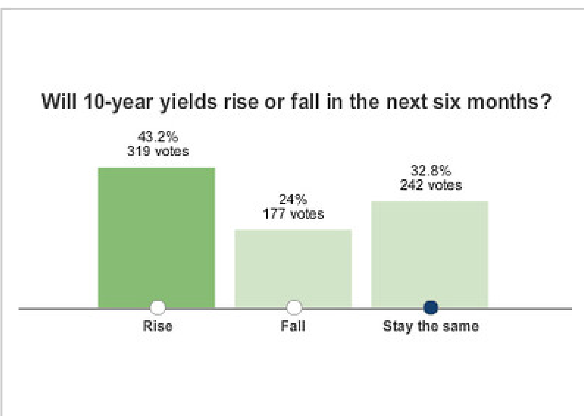
"There's an inherent bias out there that you can only get validation that the economy is improving if rates go up," said George Goncalves, head of interest-rate strategy at Nomura Securities. He was among the strategists saying in the spring that yields would keeping falling.

But the relationship between yields and the economy isn't always linear. Despite steady improvement in the economic numbers, yields have continued to fall. That's in part because of sluggish growth abroad, which has helped push back market views of when the central bank will begin hiking rates.

Goncalves added that falling yields have actually been a boon to the economy this year, keeping financial conditions loose and supporting the housing market. That creates a somewhat paradoxical situation where economic growth and yields are moving in the opposite direction.

The survey of economists' yield projections is generally skewed toward rising rates — only a few times since early 2009 have a majority of respondents to the Bloomberg survey thought rates would fall. But the unanimity of the rising rate forecasts in the spring was a stark reminder of how one-sided market views can become. It also teaches us that economists can be universally wrong.

Then again, the majority of MarketWatch readers weren't exactly expecting rates to fall either, judging by an informal survey taken at the time:



Looking forward, can you guess in which direction the most recent Bloomberg survey of economists shows yields are headed? Yep, the answer is up.

Do you think the 10-year yield will rise or fall in the next six months?

Rise OR Fall

MarketWatch

MarketWatch

Yes, 100% of economists were dead wrong about yields

By Ben Eisen
Published: Oct 22, 2014 8:01 a.m. ET

Back in April every economist in a survey thought yields would rise. Guess what they did next

Getty Images
As it turns out, economists are not soothsayers.

NEW YORK (MarketWatch) — Just about six months ago, a headline flashed across the top of MarketWatch's home page. It read: "100% of economists think yields will rise within six months."

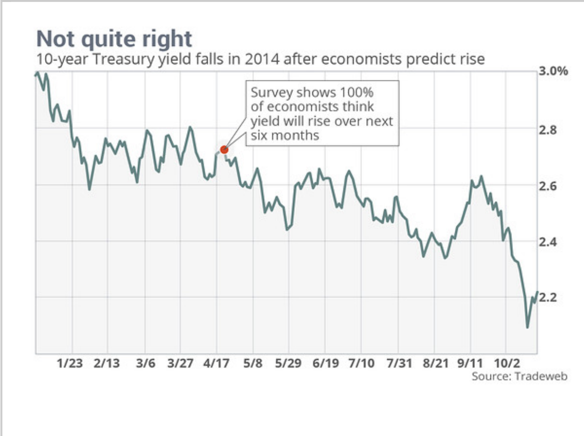
The April 22 report was based on a Bloomberg survey of 67 economists, all of whom expected the 10-year Treasury note 10_YEAR, +0.34% yield — which closed at 2.73% that day — to rise over the following half year.

"How quickly we would get to 4[%] was the discussion at the beginning of the year," said Mohamed El-Erian, chief economic adviser at Allianz SE, on CNBC Tuesday morning.

The market, however, has a funny way of leaning one way, just as the herd is heading in the other direction.

On Tuesday, the 10-year note traded at a yield of 2.21%, almost four-tenths of a percentage point lower than in April. Let's not forget that the yield unexpectedly dipped below 2%, just last week.

That underscores the difficulty of calling the direction of interest rates. It also makes all 67 economists wrong, as this chart of the benchmark yield shows:



Treasury yields tend to rise, and prices drop, as the U.S. economy grows and investors begin to expect the Federal Reserve to normalize monetary policy more quickly.

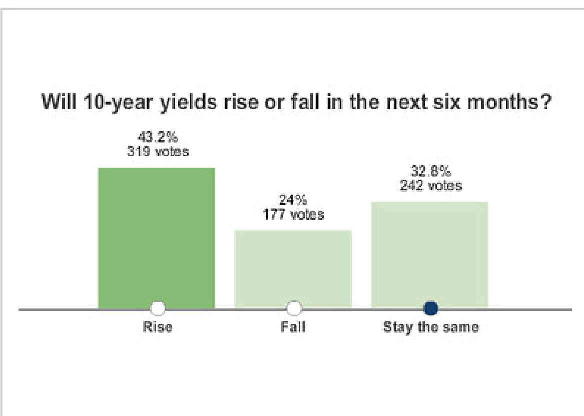
“There’s an inherent bias out there that you can only get validation that the economy is improving if rates go up,” said George Goncalves, head of interest-rate strategy at Nomura Securities. He was among the strategists saying in the spring that yields would keeping falling.

But the relationship between yields and the economy isn’t always linear. Despite steady improvement in the economic numbers, yields have continued to fall. That’s in part because of sluggish growth abroad, which has helped push back market views of when the central bank will begin hiking rates.

Goncalves added that falling yields have actually been a boon to the economy this year, keeping financial conditions loose and supporting the housing market. That creates a somewhat paradoxical situation where economic growth and yields are moving in the opposite direction.

The survey of economists’ yield projections is generally skewed toward rising rates — only a few times since early 2009 have a majority of respondents to the Bloomberg survey thought rates would fall. But the unanimity of the rising rate forecasts in the spring was a stark reminder of how one-sided market views can become. It also teaches us that economists can be universally wrong.

Then again, the majority of MarketWatch readers weren’t exactly expecting rates to fall either, judging by an informal survey taken at the time:



Looking forward, can you guess in which direction the most recent Bloomberg survey of economists shows yields are headed? Yep, the answer is up.

Do you think the 10-year yield will rise or fall in the next six months?

Rise OR Fall

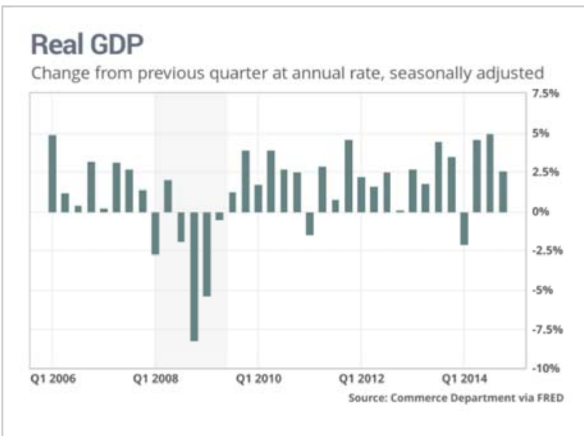
MarketWatch

MarketWatch

Economy downshifts to 2.6% rate in the fourth quarter

By Greg Robb
Published: Jan 30, 2015 12:24 p.m. ET

GDP below expectations of a 3.2% gain



WASHINGTON (MarketWatch) — The U.S. economy slowed a bit more than expected in the fourth quarter after expanding at the fastest pace in eleven years during the fall, according to data released Friday.

Gross domestic product — the value of all goods and services produced by the U.S. — grew at a 2.6% annual clip in the fourth quarter, the government said Friday. That's below the 5.0% pace recorded in the July-September period.

Economists polled by MarketWatch forecast GDP would grow by a seasonally adjusted 3.2% in the October-to-December period.

Stock traded lower all day Friday after the data was released. The S&P 500 index SPX, -0.78% was recently down 7 points to 2,014.

For all of 2014, the U.S. economy grew at a 2.4% rate, slightly faster than the 2.2% gain in the prior year.

Consumer spending was a major positive in the fourth quarter, expanding 4.3%, the fastest pace since before the financial crisis.

But growth was pulled down by weaker business spending, a drop in federal government spending and net exports.

Economists say the pattern of strong consumer spending and weak business spending should persist in the first quarter as a result of the sharp drop in oil prices.

"The economy is also showing more signs of lopsided growth, being too reliant on the consumer.," said Chris Williamson, chief economist at Markit.

And the stronger dollar DXY, +0.18% may also weaken the U.S. trade sector in coming quarters.

Economists were divided over what today's report signaled for coming quarters.

"This slowdown is nothing to worry about," said Paul Ashworth, chief U.S. economist at Capital Economics.

But Williamson said it might delay a Fed rate hike until late 2015 or 2016.

Prior to the release, economists polled by MarketWatch forecasted the U.S. will expand by roughly 3% in the first and second quarters. They based their optimism on a surge in hiring that's added 2.95 million new jobs in 2014, the largest gain since 1999.

Inflation as measured by the Federal Reserve's preferred price index, meanwhile, weakened in the fourth quarter to the lowest rate in almost six years, potentially making the central's bank effort at managing the U.S. recovery more difficult.

The PCE index fell at a 0.5% annual rate in the October-to-December period, compared to a 1.2% gain in the third quarter. That's the biggest drop since the first quarter of 2009. The core PCE that excludes food and energy rose at a 1.1% clip, down from 1.4%.

The Fed believes the slowdown in inflation will be temporary, but if the central bank is wrong, it could be forced to hold rates at zero longer than it would like.

The Ultimate Poison Pill: Closing the Value Gap

James M. McTaggart, Chairman & Chief Executive Officer

Seldom in the history of U.S. business has a structural change hit with the same force. Ten years ago, large-scale LBOs, raiders, and forced restructuring were virtually unknown. Today, they are commonplace and are rapidly changing the economic landscape. At the source of this structural change is a growing belief that many large diversified companies are not being managed to create the maximum value possible for their shareholders. It is also important to note that the gap between actual and potential market values, the "value gap," is so large for some companies that substantial profits can be made even after premiums of 30- 50% are paid to acquire control. This perception, combined with a flood of institutional money into junk bonds and LBO funds, has produced the takeover entrepreneur, who can now entice or threaten all but the very largest corporations.

Can it be true? Is the value gap of sufficient size to make a large number of diversified companies attractive takeover candidates? In general, the answer is yes, although the number of candidates has been declining recently due to the spread of value-based strategic management. More important, however, are the sources of the gap. There are three management shortcomings that we believe account for most of the gap between actual and potential market values:

1) A tendency to invest far too much capital in unprofitable businesses

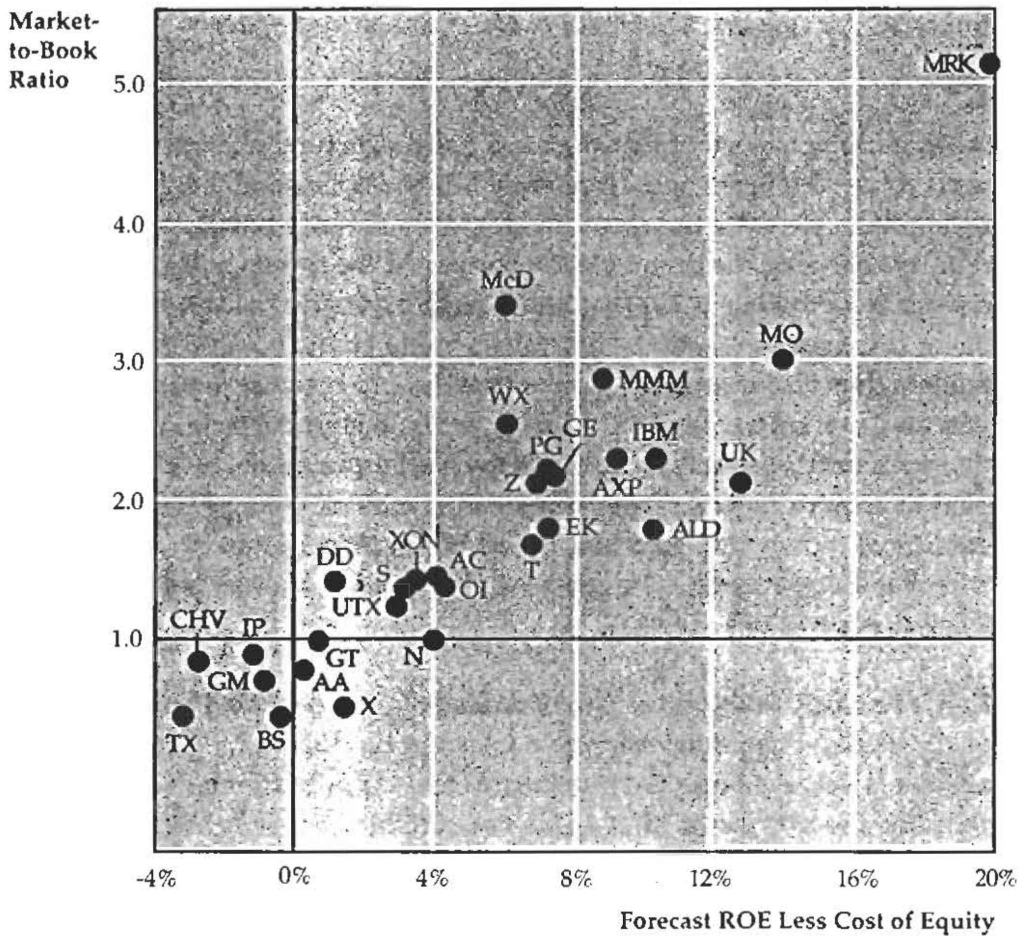
2) Poor balance sheet management, and

3) Tolerance of noneconomic overhead.

The Determinants of Value

In order to describe clearly the three sources of the value gap, it is necessary to first examine the factors that determine the market value of any business or company.

Exhibit 1: Profitability of Dow Jones Industrials - June 1986



ALD	Allied Corp.	IP	Int'l. Paper
AA	Aluminum Co. of Am.	McD	McDonald's Corp.
AC	American Can	MRK	Merck & Co.
AXP	American Express	MMM	Minnesota Mining
T	American Telephone	MO	Philip Morris
BX	Bethlehem Steel	OI	Owens-Illinois
CHV	Chevron	PG	Proctor & Gamble
DD	DuPont	S	Sears, Roebuck
EK	Eastman Kodak	TX	Texaco, Inc.
XON	Exxon Corp.	UK	Union Carbide
GE	General Electric	X	U.S. Steel
GM	General Motors	UTX	United Technologies
GT	Goodyear Tire	WX	Westinghouse
IBM	Int'l. Business Machines	ZZ	Woolworth (F.W.)
N	Inco Limited		

Fundamentally, the value of a company is determined by the cash flow it generates over time for its owners and the minimum acceptable rate of return required by investors to supply equity capital. This "cost of equity capital" is used to discount the expected equity cash flow, converting it to a present value. The cash flow is, in turn, produced by the interaction of a company's return on equity and the annual rate of equity growth. High-ROE companies in low-growth markets, such as Kellogg, are prodigious generators of cash flow, while low-ROE companies in high-growth markets, such as Texas Instruments, barely generate enough cash flow to finance growth.

A company's ROE over time relative to its cost of equity also determines whether it is worth more or less than its book value. If ROE is consistently greater than the cost of equity capital (the investor's minimum acceptable return), the business is economically profitable and its market value will exceed book value. If, however, the business earns an ROE consistently less than its cost of equity, it is economically unprofitable and its market value will be less than book value. These basic principles can be seen at work in **Exhibit I**, which plots the profitability of the Dow Jones Industrials, based on Value Line forecasts of ROE and Marakon estimates of the cost of equity capital.

Growth acts as a magnifier. If ROE remains constant and the growth rate of a profitable business increases, its market-to-book ratio rises. For an unprofitable business, increasing growth actually drives the market-to-book lower (unless growth causes ROE to rise). And in the case where ROE is just equal to the cost of equity, growth has no impact on the market-to-book ratio. The primary reason for the scattering of the observations in **Exhibit I** is differential growth rates.

The profitability of a company is determined primarily by the profitability of its businesses. The profitability of a business is, in turn, determined by economic forces affecting supply and demand in its product markets, its competitive position, and the effectiveness of its strategy. The interaction of constantly changing economic forces and competitive strategies produces a wide variation in both industry and company profitability, as can be seen in **Exhibits II and III**. Understanding how industry economics and competitive position determine profitability for a given business is the first step toward developing strategies to increase shareholder returns.

Exhibit II: Profitability of 14 U.S. Industries – Spring 1986

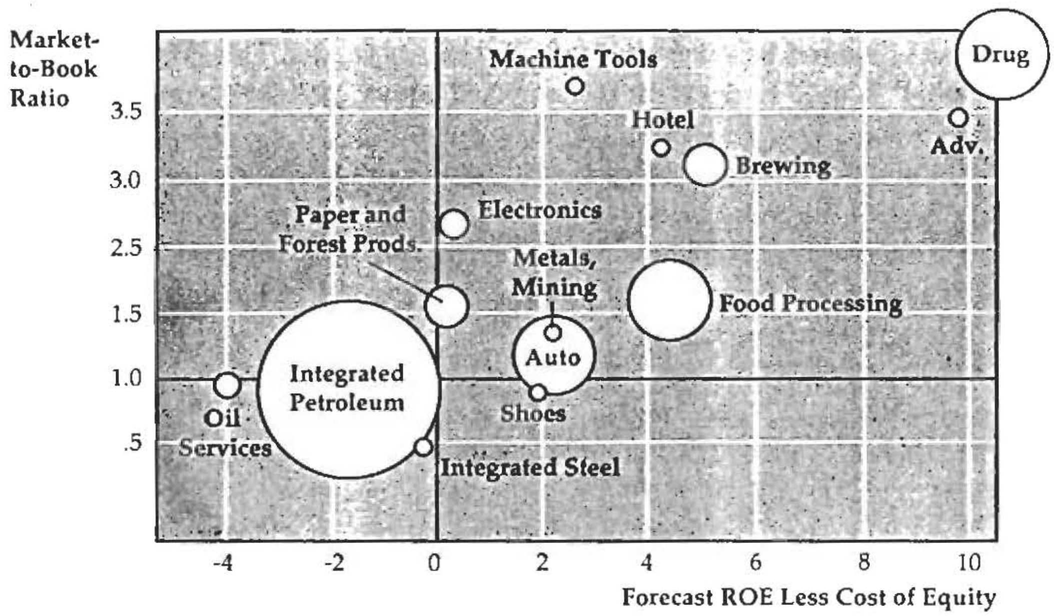
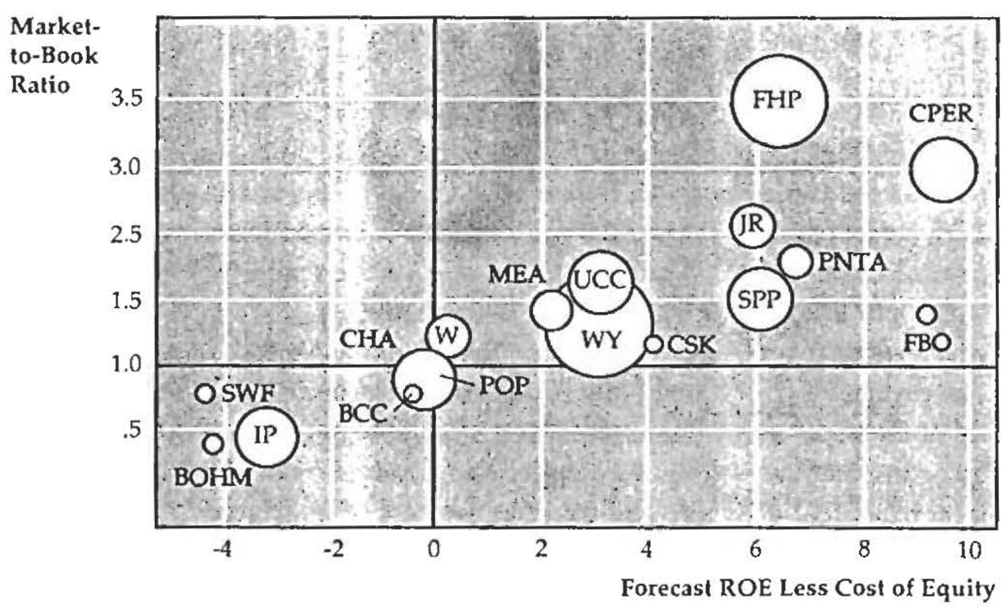


Exhibit III: Profitability of Paper and Forest Products Companies – Spring 1986



BOHM	Bohemia	FHP	Fort Howard Paper	SPP	Scott Paper
BCC	Boise Cascade	IP	International Paper	SWF	Southwest Forest
CHA	Champion International	JR	James River	UCC	Union Camp
CSK	Chesapeake	MEA	Mead	W	Westvaco
CPER	Consolidated Paper	PNTA	Pentair	WY	Weyerhaeuser
FBO	Federal Paper Board	POP	Pope & Talbot		

Sources of the Value Gap

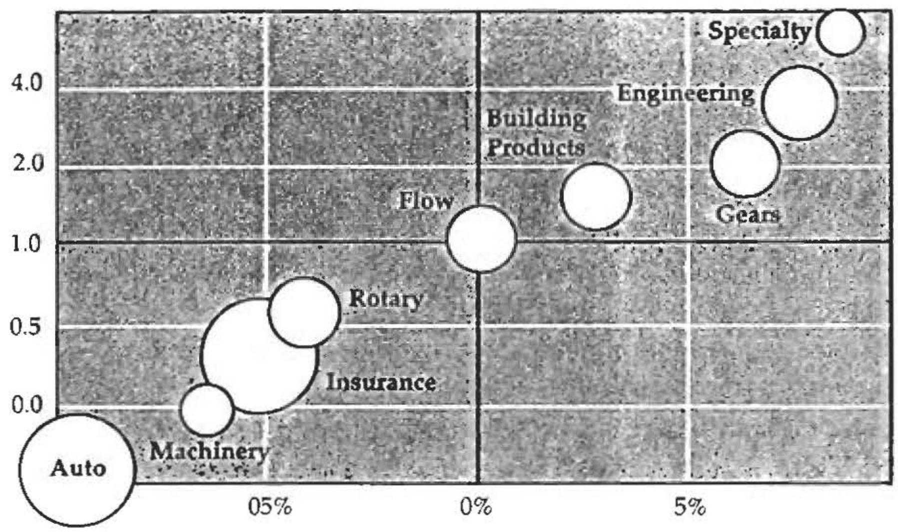
The wide variation in industry and company profitability also occurs within a typical diversified company's portfolio of businesses. Within a company, however, the capital allocation discipline provided by creditors and investors is replaced by management policies and strategies, which can significantly magnify the variation, particularly on the downside. The magnification can occur in either of two ways. The first is when management allows low-return businesses to invest too much capital, a process that can actually produce businesses with negative market values. The second is when management allows or causes high-return businesses to underinvest, which if prolonged usually results in a loss of competitive position and declining returns. In both instances, the business unit market values are significantly lower than they otherwise would be. This tendency to misallocate capital by allowing or causing businesses to pursue inappropriate strategies is the first of the three major sources of the gap between actual and potential market value.

The business portfolio shown in **Exhibit IV**, based on a recent engagement, illustrates the magnitude of the gap that can be produced by pursuing inappropriate business strategies. This company's sales were roughly $750 million, and its common stock was trading at about 80% of book value. Its portfolio contained five profitable and four unprofitable businesses. The operating value of each unprofitable business, based on the prevailing strategies, was less than 50% of its book value. All told, the four operating values summed to $115 million, versus a combined book value exceeding $300 million.

The most unprofitable business, machinery, was actually worth a negative $12 million; that is, the present value of its planned cash flow was negative $12 million. This was produced by an operating strategy whose primary objective was growth. The key element of the plan was a massive capital spending program designed to boost capacity and eliminate a competitive cost disadvantage. And while the program, if successful, would have significantly enhanced the unit's ROI (from 8% to 12%), the long-term positive impact on value was more than offset by the near-term negative cash flow.

Based on a thorough assessment of market economics and profitability relative to competitors, we concluded that by changing strategy at each

Exhibit IV: Profitability of Company Portfolio



of the four businesses to emphasize profitability rather than growth, their combined market values could be increased by at least $150 million within two years. In other words, the current value gap caused by over-investing in four unprofitable businesses was $120 million, or 40% of the company's market value.*

As a general rule, strategy changes at the business unit level emanating from improved capital allocation can enhance market values by anywhere from 20-100% within a few years. While this alone can provide impetus to takeover entrepreneurs, the value gap can, in fact, be further magnified by poor balance sheet management and tolerance of noneconomic overhead.

With respect to balance sheet management, substantial value can often be created by redeploying underperforming assets and reducing the cost of capital used to fund investments. On the asset side, two of the more prominent targets are excess cash and underutilized real estate. The source of value creation in the cash account is the low after-tax return it earns. To the extent that excess cash is held for long periods of time in

*The machinery business was subsequently sold in a leveraged buyout for book value and has since prospered.

taxable securities, it is worth less than its face value. Redeploying excess cash by repurchasing shares, for example, generates a capital gain equal to the present value of the tax savings. Excess pension fund reserves are also a source of funds that can be worth more if returned to shareholders. The source of value creation with corporate real estate is land or buildings that are not being put to their highest and best use. The capital tied up in undeveloped land, vacant office space, underutilized plants, or unprofitable retail outlets nearly always earns a return well below the cost of capital. To the extent that it can be redeployed into profitable businesses or, again, used to buy back stock, a substantial capital will occur.

On the liability side, value can be created for equity holders by increasing financial leverage up to a point. This, of course, is one of the sources of value that LBOs have utilized to recapture purchase price premiums. The source of the value creation is the tax saving due to the deductibility of interest. As a rule of thumb, each dollar of new debt should increase the firm's equity value by 20-25 cents until the firm's financial risk becomes excessive. At this point, the benefits from futher borrowing are offset by the restrictions placed on the firm, which limit its capital availability and increase the probability that the interest expense will not be tax deductible. This point, however, is significantly beyond the current leverage position of most U.S. companies.

The magnitude of the opportunity to increase returns through improved balance sheet management will, of course, depend on the amount of nonproductive assets on the company's books and its capacity to borrow. In the case of Gulf Oil, we estimated that redeployment of over $1 billion of excess cash and full utilization of the company's debt capacity would have produced a 20-25% increase in the market value of Gulf's stock. Focused efforts to reduce underperforming assets and improve liability management can result in increases to shareholder value of up to 50%.

With respect to overhead, our experience suggests that most large companies are overburdened and do not appreciate the magnitude of the overhead drag on equity values. The accumulation of overhead throughout most companies occurs for a variety of reasons. As companies grow, they face the continuing problem of how to decentralize operating responsibility while maintaining some centralized control. In many instances, the result is duplication of support functions at corporate, group, and business unit levels, such as accounting, personnel, and planning. In addition, the overriding objective of most people managing the support functions is to maximize the quality of their services, and their compensation is often closely correlated to the number of people under their stewardship. The result is excess staff and a service "quality-to-cost" ratio that is much lower than it should be.

The impact of noneconomic overhead on value can be staggering. For example, the overhead at Beatrice Corp. was estimated at roughly \$150 million annually, or 1.3% of its \$12 billion in sales. By contrast, Esmark, at roughly \$6 billion in sales, was spending only \$25 million on overhead functions, less than 0.5%. If Beatrice could have managed down its overhead to \$50 million, the resulting \$100 million in pretax earnings would have created roughly \$1 billion of shareholder value. This represents nearly 30% of Beatrice's preacquisition market value and 70% of the premium paid to acquire control of the company. This means that if the new owners can manage down Beatrice's overhead to Esmark's level, they will be two thirds of the way to recovering the acquisition premium, with potential divestments, strategy changes, and the impact of leverage and taxes yet to be considered.

Closing the Value Gap

In the current environment, with takeover financing readily available, no company can run for long with a large perceived gap between actual and potential market values. To close the gap, we recommend a five-step process:

First, develop accurate estimates of the operating and divestment values of each business in the portfolio. Few companies have this information, and yet it is the foundation of managing for shareholder value.

Second, incorporate profitability and operating values into both the strategic planning process and incentive compensation. The planning process should stress the relationships among market economics, competitive position, and profitability. Business unit managers cannot be expected to develop value-creating strategies if they don't know how much their units are worth or why they are either profitable or unprofitable. To ensure effective implementation, a significant portion of key executive compensation must be tied directly or indirectly to shareholder value.

Third, don't hoard cash or carry nonproductive assets on the books. At least once a year, a thorough analysis of asset productivity should be conducted.

Fourth, put in place an aggressive financial policy. The level of borrowing should be matched to the ability of business units to bear interest rate risk. Excess cash flow should be dedicated to profitable diversification, dividends, and repurchasing shares.

Fifth, don't tolerate noneconomic overhead. Support functions should be viewed as service businesses and where possible, subjected to both performance measurement and outside competition.

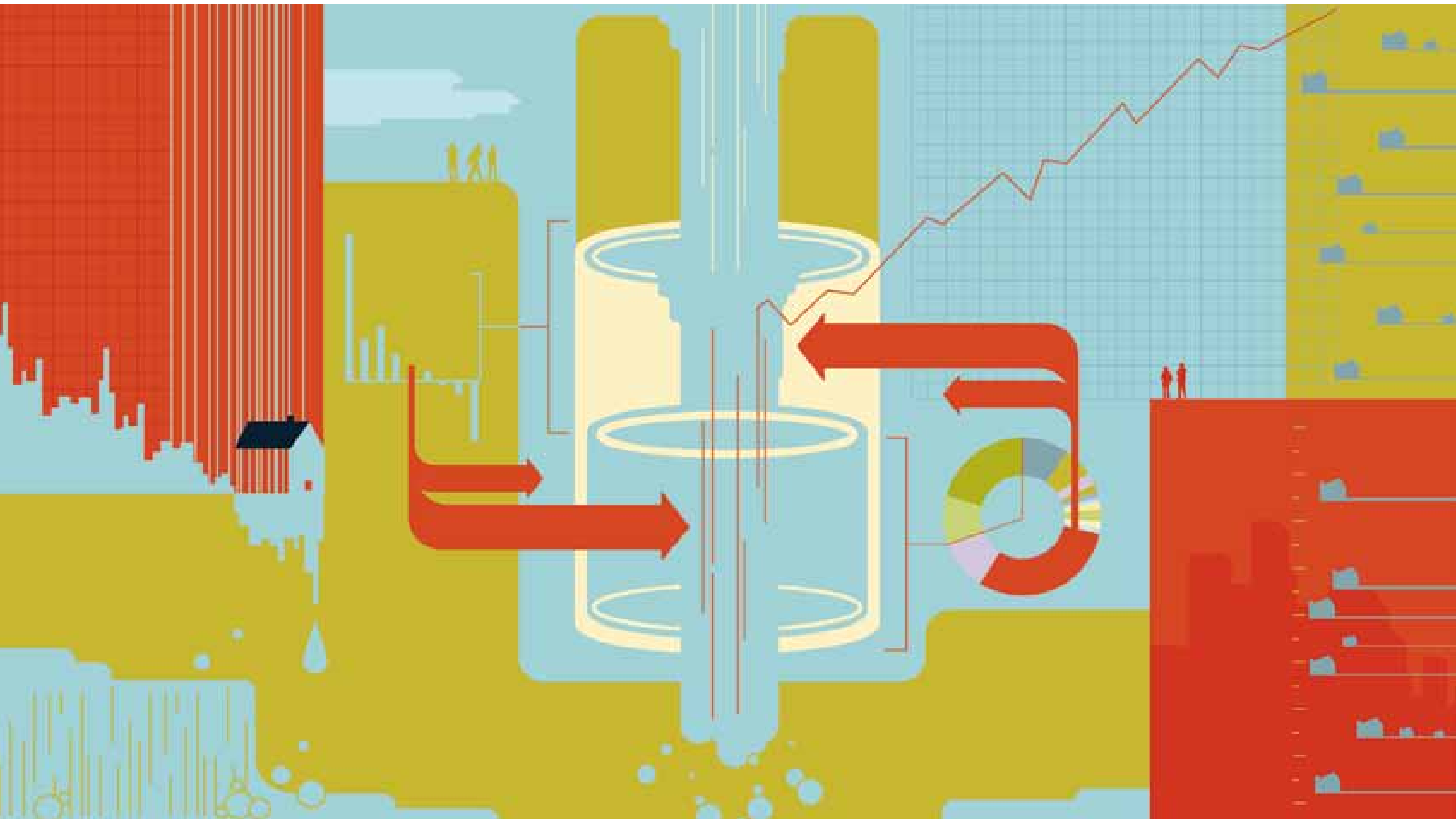
If managed well, a diversified company could be worth more than just the sum of its business unit values, owing to economies of scale and scope in support functions and to the increase in debt capacity produced by diversification. Those companies that can accomplish this feat will not only enrich shareholders but will also put in place the best possible poison pill.

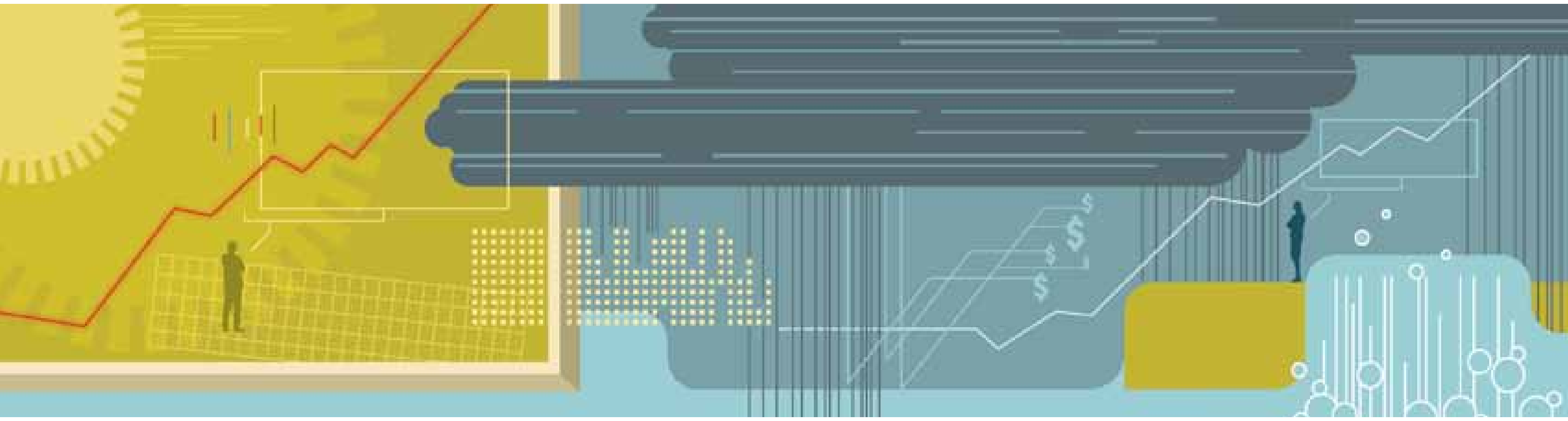
McKinsey&Company

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.

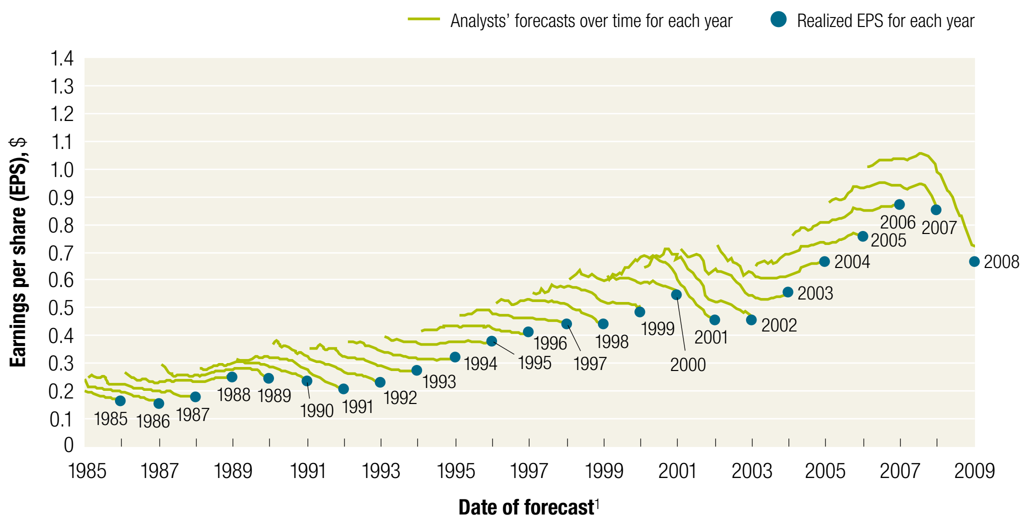
Exceptions to the long pattern of excessively optimistic forecasts are rare, as a progression of consensus earnings estimates for the S&P 500 shows (Exhibit 1). Only in years such as 2003 to 2006, when strong economic growth generated actual earnings that caught up with earlier predictions, do forecasts 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.

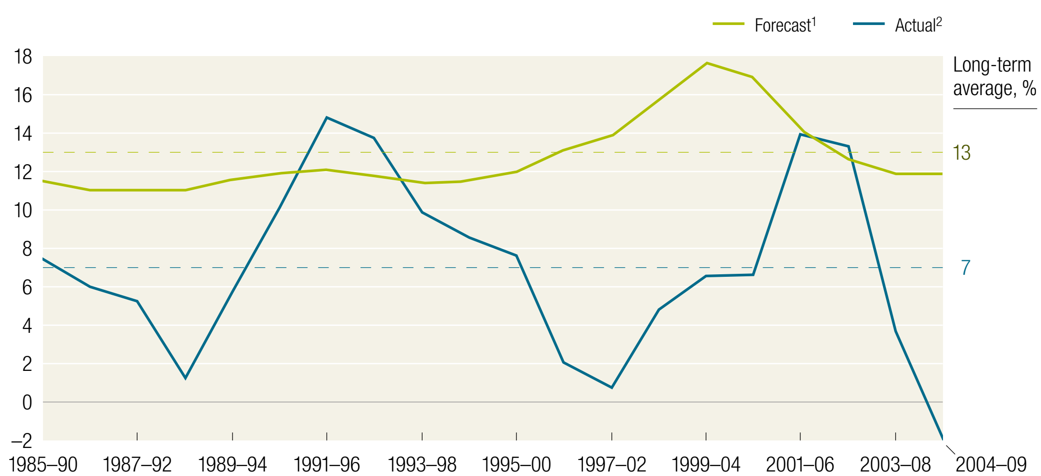
Source: Thomson Reuters I/B/E/S Global Aggregates; McKinsey analysis

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.

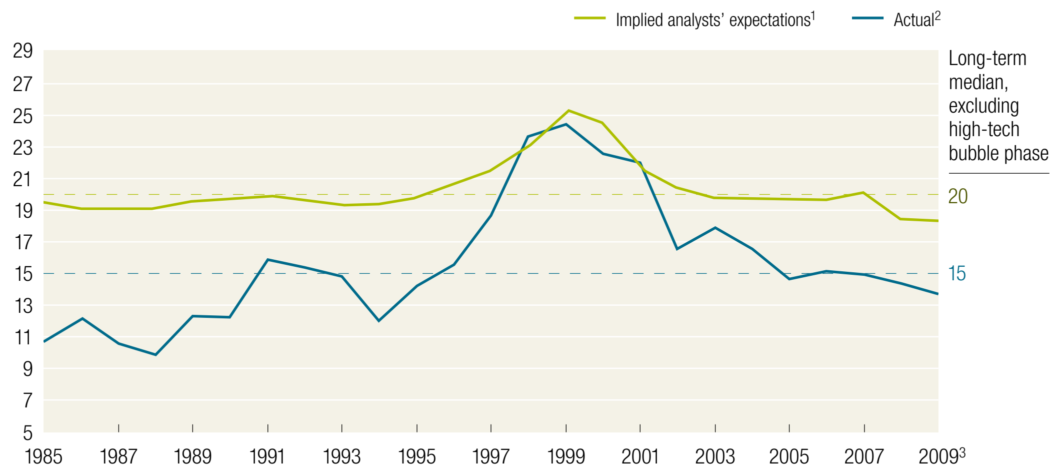
Source: Thomson Reuters I/B/E/S Global Aggregates; McKinsey analysis

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.

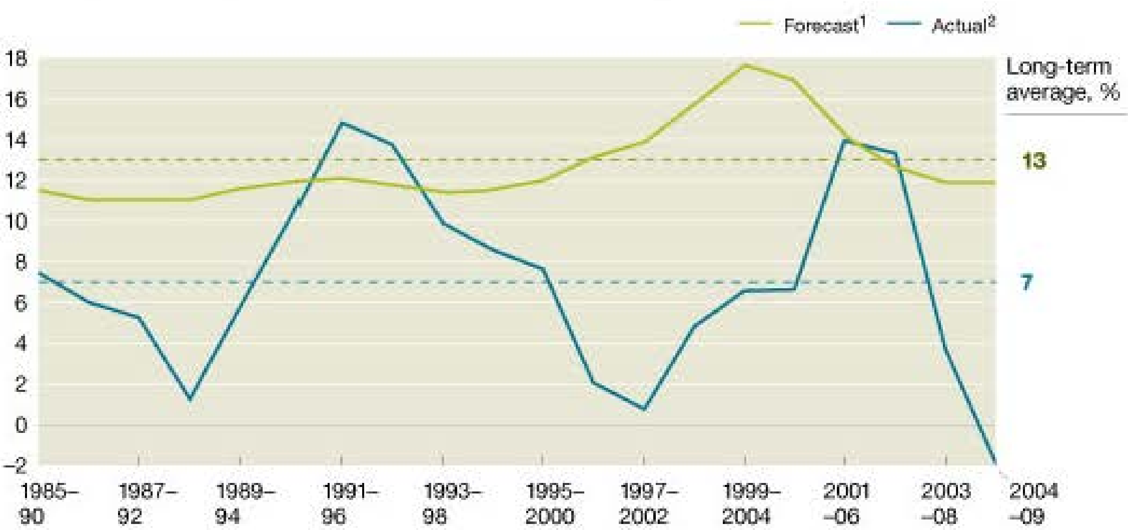
[10] Timothy Koller and Zane D. Williams, "What happened to the bull market?" mckinseyquarterly.com, November 2001.

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.

A generation of overoptimistic equity analysts

McKinsey research shows that equity analysts have been overoptimistic for the past quarter century: on average, their earnings-growth estimates—ranging from 10 to 12 percent annually, compared with actual growth of 6 percent—were almost 100 percent too high. Only in years of strong growth, such as 2003 to 2006, when actual earnings caught up with earlier predictions, do these forecasts hit the mark.

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

The McKinsey Quarterly

CORPORATE FINANCE

DECEMBER 2008

Why the crisis hasn't shaken the cost of capital

The cost of capital hasn't increased so far in the downturn—and didn't in past recessions.

Richard Dobbs, Bin Jiang, and Timothy M. Koller

McKinsey&Company

The cost of capital for companies reflects the attitudes of investors toward risk—specifically, the reward they expect for taking risks. If they become more averse to risk, companies have difficulty raising capital and may need to cancel or defer some investments or to forgo some mergers and acquisitions. So it's understandable that the current financial crisis has many executives concerned about what the price of risk—the cost of capital—will mean for their strategic decisions in the near term.

Yet our analysis finds no evidence that the long-term price of risk has increased over its historical levels—even though short-term capital is difficult to obtain. Anyone with a longer-term view won't find this surprising. At the peak of the tech bubble of 2000, when the media were awash with suggestions that the cost of capital had permanently declined, a deeper analysis suggested that it was remarkably stable—and has been for the past 40 years.[1]

Obviously, for companies that are concerned about survival and having difficulty raising capital, its cost is clearly irrelevant. We realize some companies just don't have access to new capital, period. Yet for companies that have more of it than they need to survive—either from internally generated funds or the long-term-debt markets—assumptions about its cost can make the difference between snapping up promising opportunities or being overtaken by competitors.

To understand changes in the weighted average cost of capital (WACC), we need to examine, in nominal terms, its component parts: the cost of equity and the cost of debt.

Cost of equity

We infer changes in the cost of equity by examining changes in equity values and in expected future profits and cash flows. Neither of these can be measured straightforwardly.

The S&P 500's climax—1,500, in 2007—reflected extraordinarily high profits in the financial, petroleum, and mining sectors and above-trend profits in many others.[2] To normalize the level of equity prices, we compared the long-term relationship between GDP growth and corporate profits. We estimated that, in mid-2008, the long-term sustainable level of corporate earnings would suggest a price level for the S&P 500 of about 1,100 to 1,200.[3] At the time of writing, the index was fluctuating in the 900-to-950 range, a decline of 15 to 25 percent from this sustainable level.

We can also calibrate this decline with the decline in share prices of those companies that did not experience the same earnings bubble, such as consumer goods companies and retailers. We find that these companies, which have had more stable earnings, are a stronger benchmark for assessing the economy-wide cost of capital. Their share prices at the time of this writing were down by about 15 to 20 percent from peak levels. Admittedly, this calculation isn't exact, and prices change daily.

The second factor in assessing the cost of equity capital is the ongoing level of corporate profits, which typically falls in recessions as GDP trend growth declines. History suggests that a recession involving a 5 to 10 percent decline in the cumulative long-term GDP trend would permanently reduce the corporate-profits trend line also by 5 to 10 percent.

Now let's pull these variables together into a discounted-cash-flow model. A midpoint estimate of the share-price decline—20 percent—and a 7.5 percent decline in the profit trend line translate into a hike in the cost of equity capital of about half of a percentage point. That is within the usual allowances for measurement error and within the range of annual market fluctuations.

Note that this analysis does not make allowance for the expected sharper short-term drop in corporate profits or for the market's tendency to overreact to recessions. Taking all these factors into account, we think there has been no significant change in the long-term cost of equity capital.

EXHIBIT 1

Minimal impact

Change in cost of equity, percentage point

Changes in earnings (each year in perpetuity), %

Changes in share prices,%	**–10.0**	**–7.5**	**–5.0**	**–2.5**	**0**
–25	0.8	0.9	1.1	1.2	1.3
–20	0.5	0.6	0.7	0.9	1.0
–15	0.2	0.4	0.5	0.6	0.7
–10	0.0	0.1	0.2	0.3	0.4
–5	–0.2	–0.1	0.0	0.1	0.2
0	–0.4	–0.3	–0.2	–0.1	0

Cost of equity increases by ≥1 percentage point.

20% reduction in share price combined with 7.5% profit decline = 0.6 percentage point increase in cost of equity capital.

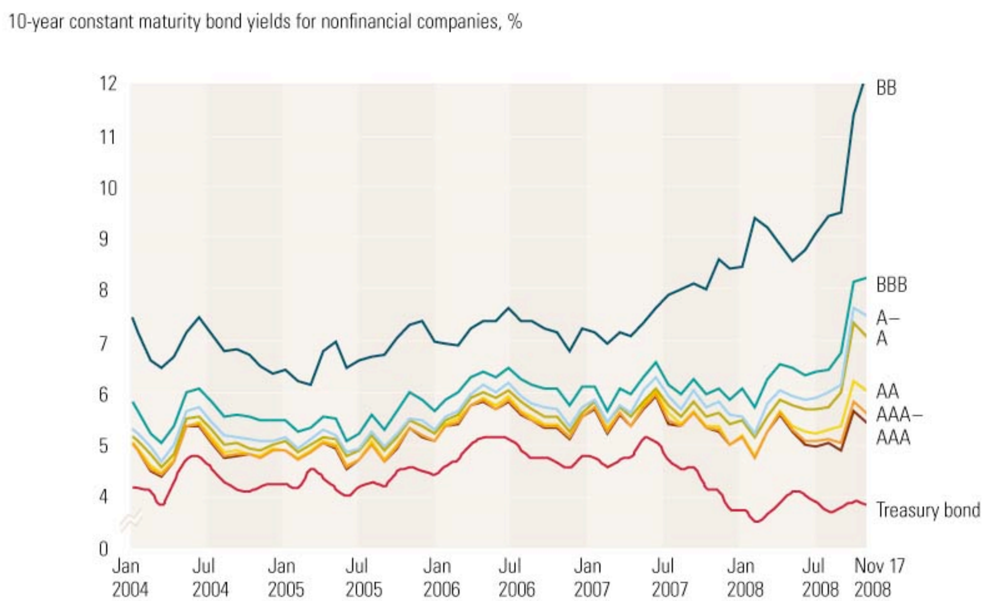
Cost of equity decreases

But this is based on our assumptions: Exhibit 1 allows you to construct your own estimate of the change in the cost of equity capital. For it to increase by a full percentage point, share prices would have to decline by 25 percent from their normal levels while profits remained relatively stable. Mathematically, a bigger drop in profits, which some expect, would mean an even smaller increase in the cost of capital.

Some might object that very few public offerings of equity have been floated recently. Our answer is that prices of liquid shares on stock exchanges are the best indicator of what investors will pay for shares. Others might counter that the economy faces extraordinarily high uncertainty right now. That is true, but uncertainty affects industries differently and therefore ought to be built into cash flow projections rather than the cost of equity. A single uncertainty risk premium should not apply to the entire economy.

EXHIBIT 2

A growing spread



Source: Bloomberg

Cost of long-term debt

The cost of debt is the second component of the cost of capital. It's easy to assume the cost of debt has increased, considering the increase in absolute rates on corporate bonds and the spread between Treasury and corporate bonds in recent months (Exhibit 2). As a benchmark, the yield to maturity on

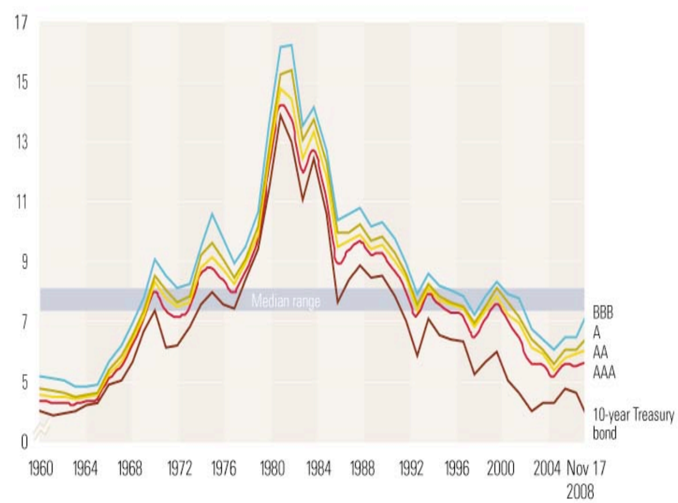
A-rated bonds rose a little more than one percentage point, to about 7 percent, from September to November 2008.

When you take a longer-term perspective, though, 7 percent isn't unusually high. Only during 6 of the past 20 years has the cost of debt for A-rated companies been lower than that (Exhibit 3).

EXHIBIT 3

Cheaper debt?

Moody's average annual bond index yields for nonfinancial companies, %



Source: Moody's; Bloomberg

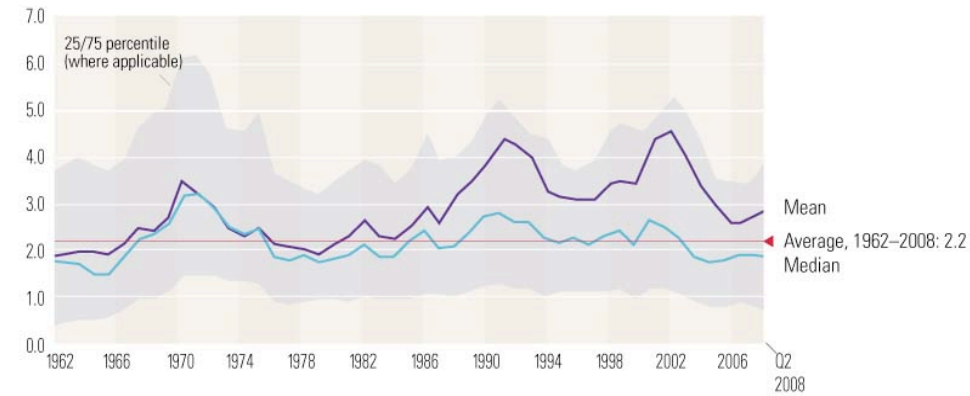
In all likelihood, the spread is increasing as a result of high demand for Treasury bonds—a demand that depresses their yields—not because investment-grade corporate bonds are becoming more risky. The rates and spreads of the past several years were probably unsustainably low and current levels are simply a reversion to normality.

The impact of the increasing cost of debt on a company's WACC is mitigated by the tax deductibility of debt and by the conservatism of the capital structures of most investment-grade companies, which means that the cost of debt is a smaller proportion of the WACC. Indeed, nonfinancial S&P 500 companies have less debt today than they have had for most of the past 40 years (Exhibit 4).

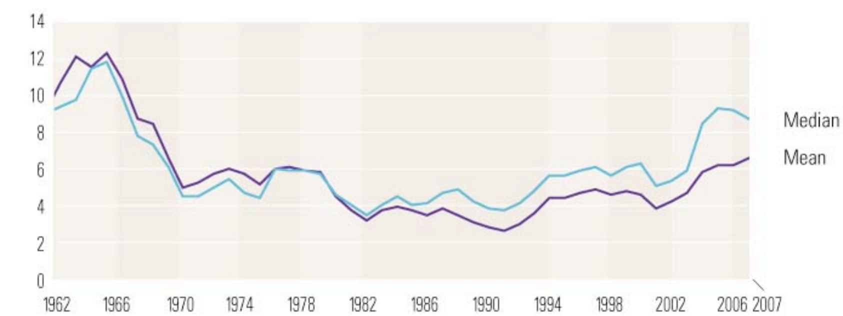
EXHIBIT 4

From a point of strength

Ratio of debt to EBITA[1] for nonfinancial S&P 500 companies



Interest coverage ratio for nonfinancial S&P 500 companies



[1]Earnings before interest, taxes, and amortization.

Implications

In sum, despite the decline in equity values and the increasing spreads on corporate debt, there is no evidence of a substantial increase in the cost of long-term capital. Of course, we cannot be certain that its cost will not increase over the next several years as the recession develops.

One unknown that demands caution is the outlook for inflation or deflation. The analysis above is on a nominal basis. For real cost of capital not to change, we need to assume that long-term inflation remains stable, at 2 to 3 percent. Some analysts are concerned about deflation, at least in the short term; others about inflation as governments around the world flood their economies with money. Deflation or high levels of inflation for an extended period could change investors' appetite for risk and the real cost of capital, along with other economic relationships.

Nonetheless, as with all valuations, the uncertainty of cash flows has a much bigger effect on value than changes in the cost of capital. That uncertainty has increased significantly. It is particularly unclear what a normal level of growth and returns on capital will be in the future. The credit bubble has distorted both during the past few years.

About the Authors

Richard Dobbs is a director in McKinsey's Seoul office and **Bin Jiang** is a consultant in the New York office, where **Tim Koller** is a principal.

Notes

[1] See Marc H. Goedhart, Timothy M. Koller, and Zane D. Williams, "The real cost of equity," mckinseyquarterly.com, October 2002.

[2] See Marc H. Goedhart, Bin Jiang, and Timothy M. Koller, "Market fundamentals: 2000 versus 2007," mckinseyquarterly.com, September 2007.

[3] See Richard Dobbs, Bin Jiang, and Timothy M. Koller, "Preparing for a slump in earnings," mckinseyquarterly.com, March 2008.

Related Articles on mckinseyquarterly.com

"The real cost of equity"

"Investing when interest rates are low"

Journal of Monetary Economics 15 (1985) 145–161. North-Holland

THE EQUITY PREMIUM
A Puzzle*

Rajnish MEHRA

Columbia University, New York, NY 10027, USA

Edward C. PRESCOTT

Federal Reserve Bank of Minneapolis

University of Minnesota, Minneapolis, MN 55455, USA.

Restrictions that a class of general equilibrium models place upon the average returns of equity and Treasury bills are found to be strongly violated by the U.S. data in the 1889–1978 period. This result is robust to model specification and measurement problems. We conclude that, most likely, an equilibrium model which is not an Arrow–Debreu economy will be the one that simultaneously rationalizes both historically observed large average equity return and the small average risk-free return.

1. Introduction

Historically the average return on equity has far exceeded the average return on short-term virtually default-free debt. Over the ninety-year period 1889–1978 the average real annual yield on the Standard and Poor 500 Index was seven percent, while the average yield on short-term debt was less than one percent. The question addressed in this paper is whether this large differential in average yields can be accounted for by models that abstract from transactions costs, liquidity constraints and other frictions absent in the Arrow–Debreu set-up. Our finding is that it cannot be, at least not for the class of economies considered. Our conclusion is that most likely some equilibrium model with a

*This research was initiated at the University of Chicago where Mehra was a visiting scholar at the Graduate School of Business and Prescott a Ford foundation visiting professor at the Department of Economics. Earlier versions of this paper, entitled 'A Test of the Intertemporal Asset Pricing Model', were presented at the University of Minnesota, University of Lausanne, Harvard University, NBER Conference on Intertemporal Puzzles in Macroeconomics, and the American Finance Meetings. We wish to thank the workshop participants, George Constantinides, Eugene Fama, Merton Miller, and particularly an anonymous referee, Fischer Black, Stephen LeRoy and Charles Plosser for helpful discussions and constructive criticisms. We gratefully acknowledge financial support from the Faculty Research Fund of the Graduate School of Business, Columbia University, the National Science Foundation and the Federal Reserve Bank of Minneapolis.

friction will be the one that successfully accounts for the large average equity premium.

We study a class of competitive pure exchange economies for which the equilibrium *growth* rate process on consumption and equilibrium asset returns are stationary. Attention is restricted to economies for which the elasticity of substitution for the composite consumption good between the year t and year $t+1$ is consistent with findings in micro, macro and international economics. In addition, the economies are constructed to display equilibrium consumption growth rates with the same mean, variance and serial correlation as those observed for the U.S. economy in the 1889–1978 period. We find that for such economies, the average real annual yield on equity is a maximum of four-tenths of a percent higher than that on short-term debt, in sharp contrast to the six percent premium observed. Our results are robust to non-stationarities in the means and variances of the economies' growth processes.

The simple class of economies studied, we think, is well suited for the question posed. It clearly is poorly suited for other issues, in particular issues such as the volatility of asset prices.[1] We emphasize that our analysis is not an estimation exercise, which is designed to obtain better estimates of key economic parameters. Rather it is a quantitative theoretical exercise designed to address a very particular question.[2]

Intuitively, the reason why the low average real return and high average return on equity cannot simultaneously be rationalized in a perfect market framework is as follows: With real per capita consumption growing at nearly two percent per year on average, the elasticities of substitution between the year t and year $t+1$ consumption good that are sufficiently small to yield the six percent average equity premium also yield real rates of return far in excess of those observed. In the case of a growing economy, agents with high risk aversion effectively discount the future to a greater extent than agents with low risk aversion (relative to a non-growing economy). Due to growth, future consumption will probably exceed present consumption and since the marginal utility of future consumption is less than that of present consumption, real interest rates will be higher on average.

This paper is organized as follows: Section 2 summarizes the U.S. historical experience for the ninety-year period 1889–1978. Section 3 specifies the set of economies studied. Their behavior with respect to average equity and short-term debt yields, as well as a summary of the sensitivity of our results to the specifications of the economy, are reported in section 4. Section 5 concludes the paper.

[1] There are other interesting features of time series and procedures for testing them. The variance bound tests of LeRoy and Porter (1981) and Shiller (1980) are particularly innovative and constructive. They did indicate that consumption risk was important [see Grossman and Shiller (1981) and LeRoy and LaCavita (1981)].

[2] See Lucas (1980) for an articulation of this methodology.

Table 1

Time periods	% growth rate of per capita real consumption: Mean	% growth rate of per capita real consumption: Standard deviation	% real return on a relatively riskless security: Mean	% real return on a relatively riskless security: Standard deviation	% risk premium: Mean	% risk premium: Standard deviation	% real return on S&P 500: Mean	% real return on S&P 500: Standard deviation
1889–1978	1.83 (Std error = 0.38)	3.57	0.80 (Std error = 0.60)	5.67	6.18 (Std error = 1.76)	16.67	6.98 (Std error = 1.74)	16.54
1889–1898	2.30	4.90	5.80	3.23	1.78	11.57	7.58	10.02
1899–1908	2.55	5.31	2.62	2.59	5.08	16.86	7.71	17.21
1909–1918	0.44	3.07	−1.63	9.02	1.49	9.18	−0.14	12.81
1919–1928	3.00	3.97	4.30	6.61	14.64	15.94	18.94	16.18
1929–1938	−0.25	5.28	2.39	6.50	0.18	31.63	2.56	27.90
1939–1948	2.19	2.52	−5.82	4.05	8.89	14.23	3.07	14.67
1949–1958	1.48	1.00	−0.81	1.89	18.30	13.20	17.49	13.08
1959–1968	2.37	1.00	1.07	0.64	4.50	10.17	5.58	10.59
1969–1978	2.41	1.40	−0.72	2.06	0.75	11.64	0.03	13.11

2. Data

The data used in this study consists of five basic series for the period 1889–1978.[3] The first four are identical to those used by Grossman and Shiller (1981) in their study. The series are individually described below:

(i) *Series P*: Annual average Standard and Poor's Composite Stock Price Index divided by the Consumption Deflator, a plot of which appears in Grossman and Shiller (1981, p. 225, fig. 1).
(ii) *Series D*: Real annual dividends for the Standard and Poor's series.
(iii) *Series C*: Kuznets–Kendrik–USNIA per capita real consumption on non-durables and services.
(iv) *Series PC*: Consumption deflator series, obtained by dividing real consumption in 1972 dollars on non-durables and services by the nominal consumption on non-durables and services.
(v) *Series RF*: Nominal yield on relatively riskless short-term securities over the 1889–1978 period; the securities used were ninety-day government Treasury Bills in the 1931–1978 period, Treasury Certificates for the

[3] We thank Sanford Grossman and Robert Shiller for providing us with the data they used in their study (1981).

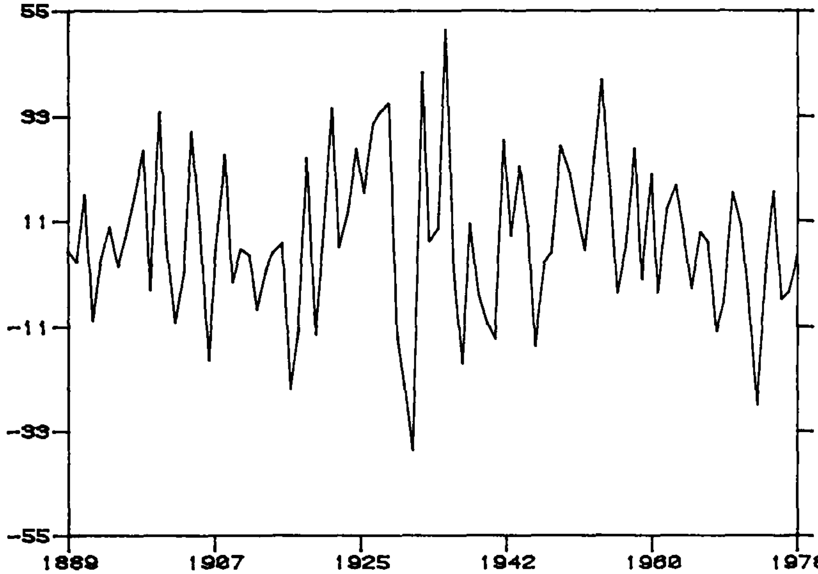


Fig. 1. Real annual return on S&P 500, 1889–1978 (percent).

1920–1930 period and sixty-day to ninety-day Prime Commercial Paper prior to 1920.[4]

These series were used to generate the series actually utilized in this paper. Summary statistics are provided in table 1.

Series P and D above were used to determine the average annual real return on the Standard and Poor's 500 Composite Index over the ninety-year period of study. The annual return for year t was computed as $(P_{t+1}+D_t-P_t)/P_t$. The returns are plotted in fig. 1. Series C was used to determine the process on the growth rate of consumption over the same period. Model parameters were restricted to be consistent with this process. A plot of the percentage growth of real consumption appears in fig. 2. To determine the real return on a relatively riskless security we used the series RF and PC. For year t this is calculated to be $RF_t-(PC_{t+1}-PC_t)/PC_t$.

This series is plotted in fig. 3. Finally, the Risk Premium (RP) is calculated as the difference between the Real Return on Standard and Poor's 500 and the Real Return on a Riskless security as defined above.

[4]The data was obtained from Homer (1963) and Ibbotson and Singuefield (1979).

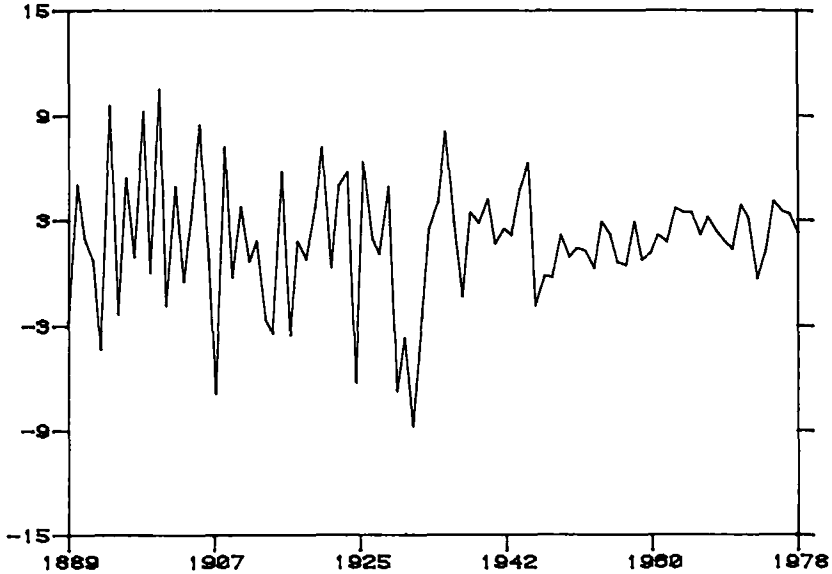


Fig. 2. Growth rate of real per capita consumption, 1889–1978 (percent).

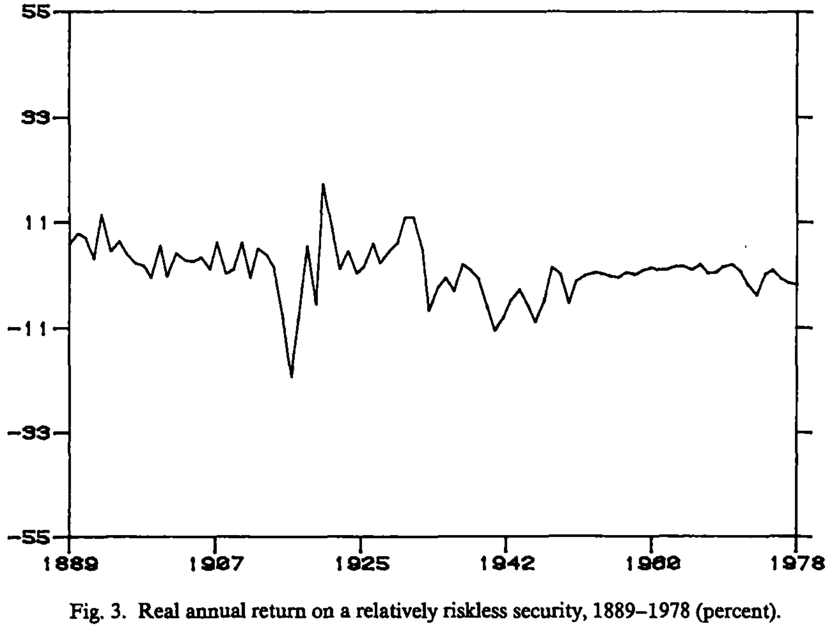


Fig. 3. Real annual return on a relatively riskless security, 1889–1978 (percent).

3. The economy, asset prices and returns

In this paper, we employ a variation of Lucas' (1978) pure exchange model. Since per capita consumption has grown over time, we assume that the *growth rate* of the endowment follows a Markov process. This is in contrast to the assumption in Lucas' model that the endowment *level* follows a Markov process. Our assumption, which requires an extension of competitive equilibrium theory, enables us to capture the non-stationarity in the consumption series associated with the large increase in per capita consumption that occurred in the 1889–1978 period.

The economy we consider was judiciously selected so that the joint process governing the growth rates in aggregate per capita consumption and asset prices would be stationary and easily determined. The economy has a single representative 'stand-in' household. This unit orders its preferences over random consumption paths by

$$\mathrm{E}_0\left\{\sum_{t=0}^{\infty}\beta^t U(c_t)\right\}, \qquad 0<\beta<1, \tag{1}$$

where c_t is per capita consumption, β is the subjective time discount factor, $\mathrm{E}_0\{\cdot\}$ is the expectation operator conditional upon information available at time zero (which denotes the present time) and U: $R_+ \to R$ is the increasing concave utility function. To insure that the equilibrium return process is stationary, the utility function is further restricted to be of the constant relative risk aversion class,

$$U(c,\alpha)=\frac{c^{1-\alpha}-1}{1-\alpha}, \qquad 0<\alpha<\infty. \tag{2}$$

The parameter α measures the curvature of the utility function. When α is equal to one, the utility function is defined to be the logarithmic function, which is the limit of the above function as α approaches one.

We assume that there is one productive unit producing the perishable consumption good and there is one equity share that is competitively traded. Since only one productive unit is considered, the return on this share of equity is also the return on the market. The firm's output is constrained to be less than or equal to y_t. It is the firm's dividend payment in the period t as well.

The growth rate in y_t is subject to a Markov chain; that is,

$$y_{t+1}=x_{t+1}y_t, \tag{3}$$

where $x_{t+1} \in \{\lambda_1, \dots, \lambda_n\}$ is the growth rate, and

$$\Pr\{x_{t+1} = \lambda_j;\ x_t = \lambda_i\} = \phi_{ij}. \tag{4}$$

It is also assumed that the Markov chain is ergodic. The λ_i are all positive and $y_0 > 0$. The random variable y_t is observed at the beginning of the period, at which time dividend payments are made. All securities are traded ex-dividend. We also assume that the matrix A with elements $a_{ij} \equiv \beta\phi_{ij}\lambda_j^{1-\alpha}$ for $i, j = 1, \dots, n$ is stable; that is, $\lim A^m$ as $m \to \infty$ is zero. In Mehra and Prescott (1984) it is shown that this is necessary and sufficient for expected utility to exist if the stand-in household consumes y_t every period. They also define and establish the existence of a Debreu (1954) competitive equilibrium with a price system having a dot product representation under this condition.

Next we formulate expressions for the equilibrium time t price of the equity share and the risk-free bill. We follow the convention of pricing securities ex-dividend or ex-interest payments at time t, in terms of the time t consumption good. For any security with process $\{d_s\}$ on payments, its price in period t is

$$P_t = \mathrm{E}_t\left\{\sum_{s=t+1}^{\infty} \beta^{s-t} U'(y_s) d_s / U'(y_t)\right\}, \tag{5}$$

as equilibrium consumption is the process $\{y_s\}$ and the equilibrium price system has a dot product representation.

The dividend payment process for the equity share in this economy is $\{y_s\}$. Consequently, using the fact that $U'(c) = c^{-\alpha}$,

$$\begin{aligned} P_t^e &= P^e(x_t, y_t) \\ &= \mathrm{E}\left\{\sum_{s=t+1}^{\infty} \beta^{s-t} \frac{y_t^\alpha}{y_s^\alpha} y_s | x_t, y_t\right\}. \end{aligned} \tag{6}$$

Variables x_t and y_t are sufficient relative to the entire history of shocks up to, and including, time t for predicting the subsequent evolution of the economy. They thus constitute legitimate state variables for the model. Since $y_s = y_t \cdot x_{t+1} \cdot \dots \cdot x_s$, the price of the equity security is homogeneous of degree one in y_t, which is the current endowment of the consumption good. As the equilibrium values of the economies being studied are time invariant functions of the state (x_t, y_t), the subscript t can be dropped. This is accomplished by redefining the state to be the pair (c, i), if $y_t = c$ and $x_t = \lambda_i$. With this convention, the price of the equity share from (6) satisfies

$$p^{e}(c,i)=\beta\sum_{j=1}^{n}\phi_{ij}(\lambda_{j}c)^{-\alpha}\left[p^{e}(\lambda_{j}c,j)+c\lambda_{j}\right]c^{\alpha}. \tag{7}$$

Using the result that $p^{e}(c,i)$ is homogeneous of degree one in c, we represent this function as

$$p^{e}(c,i)=w_{i}c, \tag{8}$$

where w_i is a constant. Making this substitution in (7) and dividing by c yields

$$w_{i}=\beta\sum_{j=1}^{n}\phi_{ij}\lambda_{j}^{(1-\alpha)}(w_{j}+1)\quad\text{for}\quad i=1,\dots,n. \tag{9}$$

This is a system of n linear equations in n unknowns. The assumption that guaranteed existence of equilibrium guarantees the existence of a unique positive solution to this system.

The period return if the current state is (c,i) and next period state $(\lambda_{j}c,j)$ is

$$\begin{aligned} r_{ij}^{e} &= \frac{p^{e}(\lambda_{j}c,j)+\lambda_{j}c-p^{e}(c,i)}{p^{e}(c,i)} \\ &= \frac{\lambda_{j}(w_{j}+1)}{w_{i}}-1, \end{aligned} \tag{10}$$

using (8).

The equity's expected period return if the current state is i is

$$R_{i}^{e}=\sum_{j=1}^{n}\phi_{ij}r_{ij}^{e}. \tag{11}$$

Capital letters are used to denote expected return. With the subscript i, it is the expected return conditional upon the current state being (c,i). Without this subscript it is the expected return with respect to the stationary distribution. The superscript indicates the type of security.

The other security considered is the one-period real bill or riskless asset, which pays one unit of the consumption good next period with certainty.

From (6),

$$\begin{aligned} p_i^{\mathrm{f}} &= p^{\mathrm{f}}(c, i) \\ &= \beta \sum_{j=1}^{n} \phi_{ij} U'(\lambda_j c)/U'(c) \\ &= \beta \sum_{j=1}^{n} \phi_{ij} \lambda_j^{-\alpha}. \end{aligned} \tag{12}$$

The certain return on this riskless security is

$$R_i^{\mathrm{f}} = 1/p_i^{\mathrm{f}} - 1, \tag{13}$$

when the current state is (c, i).

As mentioned earlier, the statistics that are probably most robust to the modelling specification are the means over time. Let $\pi \in R^n$ be the vector of stationary probabilities on i. This exists because the chain on i has been assumed to be ergodic. The vector π is the solution to the system of equations

$$\pi = \phi^T \pi,$$

with

$$\sum_{i=1}^{n} \pi_i = 1 \quad \text{and} \quad \phi^T = \{\phi_{ji}\}.$$

The expected returns on the equity and the risk-free security are, respectively,

$$R^{\mathrm{e}} = \sum_{i=1}^{n} \pi_i R_i^{\mathrm{e}} \quad \text{and} \quad R^{\mathrm{f}} = \sum_{i=1}^{n} \pi_i R_i^{\mathrm{f}}. \tag{14}$$

Time sample averages will converge in probability to these values given the ergodicity of the Markov chain. The risk premium for equity is $R^{\mathrm{e}} - R^{\mathrm{f}}$, a parameter that is used in the test.

4. The results

The parameters defining preferences are α and β while the parameters defining technology are the elements of $[\phi_{ij}]$ and $[\lambda_i]$. Our approach is to assume two states for the Markov chain and to restrict the process as follows:

$$\lambda_1 = 1 + \mu + \delta, \qquad \lambda_2 = 1 + \mu - \delta,$$

$$\phi_{11} = \phi_{22} = \phi, \qquad \phi_{12} = \phi_{21} = (1 - \phi).$$

The parameters μ, ϕ, and δ now define the technology. We require $\delta > 0$ and $0 < \phi < 1$. This particular parameterization was selected because it permitted us to independently vary the average growth rate of output by changing μ, the variability of consumption by altering δ, and the serial correlation of growth rates by adjusting ϕ.

The parameters were selected so that the average growth rate of per capita consumption, the standard deviation of the growth rate of per capita consumption and the first-order serial correlation of this growth rate, all with respect to the model's stationary distribution, matched the sample values for the U.S. economy between 1889–1978. The sample values for the U.S. economy were 0.018, 0.036 and -0.14, respectively. The resulting parameter's values were $\mu = 0.018$, $\delta = 0.036$ and $\phi = 0.43$. Given these values, the nature of the test is to search for parameters α and β for which the model's averaged risk-free rate and equity risk premium match those observed for the U.S. economy over this ninety-year period.

The parameter α, which measures peoples' willingness to substitute consumption between successive yearly time periods is an important one in many fields of economics. Arrow (1971) summarizes a number of studies and concludes that relative risk aversion with respect to wealth is almost constant. He further argues on theoretical grounds that α should be approximately one. Friend and Blume (1975) present evidence based upon the portfolio holdings of individuals that α is larger, with their estimates being in the range of two. Kydland and Prescott (1982), in their study of aggregate fluctuations, found that they needed a value between one and two to mimic the observed relative variabilities of consumption and investment. Altug (1983), using a closely related model and formal econometric techniques, estimates the parameter to be near zero. Kehoe (1984), studying the response of small countries balance of trade to terms of trade shocks, obtained estimates near one, the value posited by Arrow. Hildreth and Knowles (1982) in their study of the behavior of farmers also obtain estimates between one and two. Tobin and Dolde (1971), studying life cycle savings behavior with borrowing constraints, use a value of 1.5 to fit the observed life cycle savings patterns.

Any of the above cited studies can be challenged on a number of grounds but together they constitute an *a priori* justification for restricting the value of α to be a maximum of ten, as we do in this study. This is an important restriction, for with large α virtually any pair of average equity and risk-free returns can be obtained by making small changes in the process on consump-

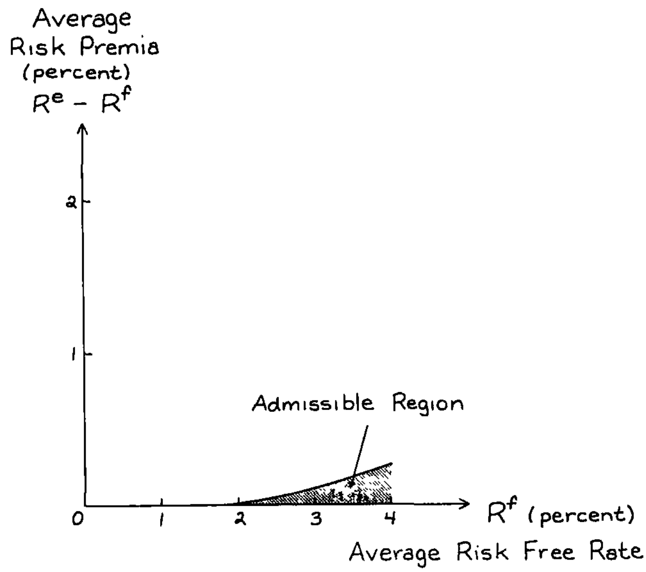


Fig. 4. Set of admissible average equity risk premia and real returns.

tion.[5] With α less than ten, we found the results were essentially the same for very different consumption processes, provided that the mean and variances of growth rates equaled the historically observed values. An advantage of our approach is that we can easily test the sensitivity of our results to such distributional assumptions.

The average real return on relatively riskless, short-term securities over the 1889–1978 period was 0.80 percent. These securities do not correspond perfectly with the real bill, but insofar as unanticipated inflation is negligible and/or uncorrelated with the growth rate x_{t+1} conditional upon information at time t, the expected real return for the nominal bill will equal $\dot{R}_i^f$. Litterman (1980), using vector autoregressive analysis, found that the innovation in the inflation rate in the post-war period (quarterly data) has standard deviation of only one-half of one percent and that his innovation is nearly orthogonal to the subsequent path of the real GNP growth rate. Consequently, the average realized real return on a nominally denoted short-term bill should be close to that which would have prevailed for a real bill if such a security were traded. The average real return on the Standard and Poor's 500 Composite Stock

[5]In a private communication, Fischer Black using the Merton (1973) continuous time model with investment opportunities constructed an example with a curvature parameter (α) of 55. We thank him for the example.

Index over the ninety years considered was 6.98 percent per annum. This leads to an average equity premium of 6.18 percent (standard error 1.76 percent).

Given the estimated process on consumption, fig. 4 depicts the set of values of the average risk-free rate and equity risk premium which are both consistent with the model and result in average real risk-free rates between zero and four percent. These are values that can be obtained by varying preference parameters α between zero and ten and β between zero and one. The observed real return of 0.80 percent and equity premium of 6 percent is clearly inconsistent with the predictions of the model. The largest premium obtainable with the model is 0.35 percent, which is not close to the observed value.

4.1. Robustness of results

One set of possible problems are associated with errors in measuring the inflation rate. Such errors do not affect the computed risk premium as they bias both the real risk-free rate and the equity rate by the same amount. A potentially more serious problem is that these errors bias our estimates of the growth rate of consumption and the risk-free real rate. Therefore, only if the tests are insensitive to biases in measuring the inflation rate should the tests be taken seriously. A second measurement problem arises because of tax considerations. The theory is implicitly considering effective after-tax returns which vary over income classes. In the earlier part of the period, tax rates were low. In the latter period, the low real rate and sizable equity risk premium hold for after-tax returns for all income classes [see Fisher and Lorie (1978)].

We also examined whether aggregation affects the results for the case that the growth rates were independent between periods, which they approximately were, given that the estimated ϕ was near one-half. Varying the underlying time period from one one-hundredths of a year to two years had a negligible effect upon the admissible region. (See the appendix for an exact specification of these experiments.) Consequently, the test appears robust to the use of annual data in estimating the process on consumption.

In an attempt to reconcile the large discrepancy between theory and observation, we tested the sensitivity of our results to model misspecification. We found that the conclusions are not at all sensitive to changes in the parameter μ, which is the average growth rate of consumption, with decreases to 1.4 percent or increases to 2.2 percent not reducing the discrepancy. The sensitivity to δ, the standard deviation of the consumption growth rate, is larger. The average equity premium was roughly proportional to δ squared. As the persistence parameter ϕ increased ($\phi = 0.5$ corresponds to independence over time), the premium decreased. Reducing ϕ (introducing stronger negative serial correlation in the consumption growth rate) had only small effects. We also modified the process on consumption by introducing additional states that permitted us to increase higher moments of the stationary distribution of the growth rate without varying the first or second moments. The maximal equity premium increased by 0.04 to 0.39 only. These exercises lead us to the conclusion that the result of the test is not sensitive to the specification of the process generating consumption.

That the results were not sensitive to increased persistence in the growth rate, that is to increases in ϕ, implies low frequency movements or non-stationarities in the growth rate do *not* increase the equity premium. Indeed, by assuming stationarity, we biased the test *towards* acceptance.

4.2. Effects of firm leverage

The security priced in our model does not correspond to the common stocks traded in the U.S. economy. In our model there is only one type of capital, while in an actual economy there is virtually a continuum of capital types with widely varying risk characteristics. The stock of a typical firm traded in the stock market entitles its owner to the residual claim on output after all other claims including wages have been paid. The share of output accruing to stockholders is much more variable than that accruing to holders of other claims against the firm. Labor contracts, for instance, may incorporate an insurance feature, as labor claims on output are in part fixed, having been negotiated prior to the realization of output. Hence, a disproportionate part of the uncertainty in output is probably borne by equity owners.

The firm in our model corresponds to one producing the entire output of the economy. Clearly, the riskiness of the stock of this firm is not the same as that of the Standard and Poor's 500 Composite Stock Price Index. In an attempt to match the two securities we price and calculate the risk premium of a security whose dividend next period is actual output less a fraction of expected output. Let θ be the fraction of expected date $t+1$ output committed at date t by the firm. Eq. (7) then becomes

$$p^{e}(c,i)=\beta\sum_{j=1}^{n}\phi_{ij}(\lambda_j c)^{-\alpha}\left[p^{e}(\lambda_j c,j)+c\lambda_j-\theta\sum_{k=1}^{n}\phi_{ik}c\lambda_k\right]c^{\alpha}. \quad (15)$$

As before, it is conjectured and verified that $p^{e}(c,i)$ has the functional form $w_i c$. Substituting $w_i c$ for $p^{e}(c,i)$ in (15) yields the set of linear equations

$$w_i=\beta\sum_{j=1}^{n}\phi_{ij}\lambda_j^{-\alpha}\left[\lambda_j w_j+\lambda_j-\theta\sum_{k=1}^{n}\phi_{ik}\lambda_k\right], \quad (16)$$

for $i=1,\dots,n$. This system was solved for the equilibrium w_i and eqs. (10), (11), and (14) used to determine the average equity premium.

As the corporate profit share of output is about ten percent, we set $\theta = 0.9$. Thus, ninety percent of expected output is committed and all the risk is borne by equity owners who receive ten percent of output on average. This increased the equity risk premium by less than one-tenth percent. This is the case because financial arrangements have no effect upon resource allocation and, therefore, the underlying Arrow–Debreu prices. Large fixed payment commitments on the part of the firm do not reverse the test's outcome.

4.3. *Introducing production*

With our structure, the process on the endowment is exogenous and there is neither capital accumulation nor production. Modifying the technology to admit these opportunities cannot overturn our conclusion, because expanding the set of technologies in this way does not increase the set of joint equilibrium processes on consumption and asset prices [see Mehra (1984)]. As opposed to standard testing techniques, the failure of the model hinges not on the acceptance/rejection of a statistical hypothesis but on its inability to generate average returns even close to those observed. If we had been successful in finding an economy which passed our not very demanding test, as we expected, we planned to add capital accumulation and production to the model using a variant of Brock's (1979, 1982), Donaldson and Mehra's (1984) or Prescott and Mehra's (1980) general equilibrium stationary structures and to perform additional tests.

5. Conclusion

The equity premium puzzle may not be why was the average equity return so high but rather why was the average risk-free rate so low. This conclusion follows if one accepts the Friend and Blume (1975) finding that the curvature parameter α significantly exceeds one. For $\alpha = 2$, the model's average risk-free rate is at least 3.7 percent per year, which is considerably larger than the sample average 0.80 given the standard deviation of the sample average is only 0.60. On the other hand, if α is near zero and individuals nearly risk-neutral, then one would wonder why the average return of equity was so high. This is not the only example of some asset receiving a lower return than that implied by Arrow–Debreu general equilibrium theory. Currency, for example, is dominated by Treasury bills with positive nominal yields yet sizable amounts of currency are held.

We doubt whether heterogeneity, per se, of the agents will alter the conclusion. Within the Debreu (1954) competitive framework, Constantinides (1982) has shown heterogeneous agent economies also impose the set of restrictions tested here (as well as others). We doubt whether non-time-additivity separable preferences will resolve the puzzle, for that would require consumptions near in time to be poorer substitutes than consumptions at widely separated dates. Perhaps introducing some features that make certain types of intertemporal trades among agents infeasible will resolve the puzzle. In the absence of such markets, there can be variability in individual consumptions, yet little variability in aggregate consumption. The fact that certain types of contracts may be non-enforceable is one reason for the non-existence of markets that would otherwise arise to share risk. Similarly, entering into contracts with as yet unborn generations is not feasible.[6] Such non-Arrow–Debreu competitive equilibrium models may rationalize the large equity risk premium that has characterized the behavior of the U.S. economy over the last ninety years. To test such theories it would probably be necessary to have consumption data by income or age groups.

Appendix

The procedure for determining the admissible region depicted in fig. 4 is as follows. For a given set of parameters μ, δ and ϕ, eqs. (10)–(14) define an algorithm for computing the values of R^e, R^f and $R^e - R^f$ for any (α, β) pair belonging to the set

$$x = \{(\alpha, \beta): 0 < \alpha \leq 10, 0 < \beta < 1, \text{ and the existence condition of section 3 is satisfied}\}.$$

Letting $R^f = h_1(\alpha, \beta)$ and $R^e - R^f = h_2(\alpha, \beta)$, h: $X \to R^2$, the range of h is the region depicted in fig. 4. The function h was evaluated for all points of a fine grid in X to determine the admissible region.

The experiments to determine the sensitivity of the results to the period length have model time periods $n = 2, 1, 1/2, 1/4, 1/8, 1/16, 1/64$ and $1/128$ years. The values of the other parameters are $\mu = 0.018/n$, $\delta = 0.036/\sqrt{n}$ and $\phi = 0.5$. With these numbers the mean and standard deviation of annual growth rates are 0.018 and 0.036 respectively as in the sample period. This follows because $\phi = 0.5$ implies independence of growth rates over periods. The change in the admissible region were hundredths of percent as n varied.

The experiments to test the sensitivity of the results to μ consider $\mu = 0.014$, 0.016, 0.018, 0.020 and 0.022, $\phi = 0.43$ and $\delta = 0.036$. As for the period length, the growth rate's effects upon the admissible region are hundredths of percent.

The experiments to determine the sensitivity of results to δ set $\phi = 0.43$, $\mu = 0.018$ and $\delta = 0.21$, 0.26, 0.31, 0.36, 0.41, 0.46 and 0.51. The equity premium varied approximately with the square of δ in this range.

[6] See Wallace (1980) for an exposition on the use of the overlapping generations model and the importance of legal constraints in explaining rate of return anomalies.

Similarly, to test the sensitivity of the results to variations in the parameter ϕ, we held δ fixed at 0.036 and μ at 0.018 and varied ϕ between 0.005 and 0.95 in steps of 0.05. As ϕ increased the average equity premium declined.

The test for the sensitivity of results to higher movements uses an economy with a four-state Markov chain with transition probability matrix

$$\begin{bmatrix} \phi/2 & \phi/2 & 1-\phi/2 & 1-\phi/2 \\ \phi/2 & \phi/2 & 1-\phi/2 & 1-\phi/2 \\ 1-\phi/2 & 1-\phi/2 & \phi/2 & \phi/2 \\ 1-\phi/2 & 1-\phi/2 & \phi/2 & \phi/2 \end{bmatrix}.$$

The values of the λ are $\lambda_1 = 1 + \mu$, $\lambda_2 = 1 + \mu + \delta$, $\lambda_3 = 1 + \mu$, and $\lambda_4 = 1 + \mu - \delta$. Values of μ, δ and ϕ are 0.018, 0.051 and 0.36, respectively. This results in the mean, standard deviation and first-order serial correlations of consumption growth rates for the artificial economy equaling their historical values. With this Markov chain, the probability of above average changes is smaller and magnitude of changes larger. This has the effect of increasing moments higher than the second without altering the first or second moments. This increases the maximum average equity premium from 0.35 percent to 0.39 percent.

References

Altug, S.J., 1983, Gestation lags and the business cycle: An empirical analysis, Carnegie-Mellon working paper, Presented at the Econometric Society meeting, Stanford University (Carnegie-Mellon University, Pittsburgh, PA).

Arrow, K.J., 1971, Essays in the theory of risk-bearing (North-Holland, Amsterdam).

Brock, W.A., 1979, An integration of stochastic growth theory and the theory of finance, Part 1: The growth model, in: J. Green and J. Scheinkman, eds., General equilibrium, growth & trade (Academic Press, New York).

Brock, W.A., 1982, Asset prices in a production economy, in: J.J. McCall, ed., The economics of information and uncertainty (University of Chicago Press, Chicago, IL).

Constantinides, G., 1982, Intertemporal asset pricing with heterogeneous consumers and no demand aggregation, Journal of Business 55, 253–267.

Debreu, G., 1954, Valuation equilibrium and Pareto optimum, Proceedings of the National Academy of Sciences 70, 588–592.

Donaldson, J.B. and R. Mehra, 1984, Comparative dynamics of an equilibrium, intertemporal asset pricing model, Review of Economic Studies 51, 491–508.

Fisher, L. and J.H. Lorie, 1977, A half century of returns on stocks and bonds (University of Chicago Press, Chicago, IL).

Friend, I. and M.E. Blume, 1975, The demand for risky assets, American Economic Review 65, 900–922.

Grossman, S.J. and R.J. Shiller, 1981, The determinants of the variability of stock market prices, American Economic Review 71, 222–227.

Hildreth, C. and G.J. Knowles, 1982, Some estimates of Farmers' utility functions, Technical bulletin 335 (Agricultural Experimental Station, University of Minnesota, Minneapolis, MN).

Homer, S., 1963, A history of interest rates (Rutgers University Press, New Brunswick, NJ).

Ibbotson, R.G. and R.A. Singuefield, 1979, Stocks, bonds, bills, and inflation: Historical returns (1926–1978) (Financial Analysts Research Foundation, Charlottesville, VA).

Kehoe, P.J., 1983, Dynamics of the current account: Theoretical and empirical analysis, Working paper (Harvard University, Cambridge, MA).

Kydland, F.E. and E.C. Prescott, 1982, Time to build and aggregate fluctuations, Econometrica 50, 1345–1370.

LeRoy, S.F. and C.J. LaCivita, 1981, Risk-aversion and the dispersion of asset prices, Journal of Business 54, 535–548.

LeRoy, S.F. and R.D. Porter, 1981, The present-value relation: Tests based upon implied variance bounds, Econometrica 49, 555–574.

Litterman, R.B. 1980, Bayesian procedure for forecasting with vector autoregressions, Working paper (MIT, Cambridge, MA).

Lucas, R.E., Jr., 1978, Asset prices in an exchange economy, Econometrica 46, 1429–1445.

Lucas, R.E., Jr., 1981, Methods and problems in business cycle theory, Journal of Money, Credit, and Banking 12, Part 2. Reprinted in: R.E. Lucas, Jr., Studies in business-cycle theory (MIT Press, Cambridge, MA).

Mehra, R., 1984, Recursive competitive equilibrium: A parametric example, Economics Letters 16, 273–278.

Mehra, R. and E.C. Prescott, 1984, Asset prices with nonstationary consumption, Working paper (Graduate School of Business, Columbia University, New York).

Merton, R.C., 1973, An intertemporal asset pricing model, Econometrica 41, 867–887.

Prescott, E.C. and R. Mehra, 1980, Recursive competitive equilibrium: The case of homogeneous households, Econometrica 48, 1365–1379.

Shiller, R.J., 1981, Do stock prices move too much to be justified by subsequent changes in dividends?, American Economic Review 71, 421–436.

Tobin, J. and W. Dolde, 1971, Wealth, liquidity and consumption, in: Consumer spending and monetary policy: The linkage (Federal Reserve Bank of Boston, Boston, MA) 99–146.

Wallace, N., 1980, The overlapping generations model of fiat money, in: J.H. Kareken and N. Wallace, eds., Models of monetary economies (Federal Reserve Bank of Minneapolis, Minneapolis, MN) 49–82.

The History of Finance

An eyewitness account.

Merton H. Miller

At five years, the German Finance Association is not very old as professional societies go, but then neither is the field of finance itself. Finance in its modern form really dates only from the 1950s. In the forty years since then, the field has come to surpass many, perhaps even most, of the more traditional fields of economics in terms of the numbers of students enrolled in finance courses, the numbers of faculty teaching finance courses, and above all in the quantity and quality of their combined scholarly output.

The huge body of scholarly research in finance over the last forty years falls naturally into two main streams. And no, I don't mean "asset pricing" and "corporate finance," but instead a deeper division that cuts across both. The division I have in mind is the more fundamental one between what I will call the *business school* approach to finance and the *economics department* approach. Let me say immediately, however, that my distinction is purely "notional," not physical — a distinction over what the field is really all about, not where the offices of the faculty happen to be located.

In the United States, the vast majority of academics in finance teach in business schools, not economics departments, and always have. At the same time, in the elite schools at least, a substantial fraction of the finance faculties have been trained in — that is, have received their Ph.D.s from — economics departments. Habits of thought acquired in graduate school have a tendency to stay with you.

MERTON H. MILLER is Robert R. McCormick distinguished service professor emeritus at the University of Chicago (IL 60637).

The characteristic business school approach tends to be what we would call in our jargon "micro normative." That is, a decision-maker, whether an individual investor or a corporate manager, is seen as maximizing some objective function, be it utility, expected return, or shareholder value, taking the prices of securities in the market as given. In a business school, after all, that's what you're supposed to be doing: teaching your charges how to make better decisions.

To someone trained in the classical traditions of economics, however, the dictum of the great Alfred Marshall stands out: "It is not the business of the economist to tell the brewer how to make beer." The characteristic economics department approach thus is not micro, but macro normative. The models assume a world of micro optimizers, and deduce from that how market prices, which the micro optimizers take as given, actually evolve.

Note that I am differentiating the stream of research in finance along macro versus micro lines, and not along the more familiar normative versus positive line. Both streams of research in finance are thoroughly positivist in outlook in that they try to be, or at least claim to be, concerned with testable hypotheses. The normal article in finance journals over the last forty years has two main sections: the first presenting the model, and the second an empirical section showing that real-world data are consistent with the model (which is hardly surprising, because had that not been so, the author would never have submitted the paper in the first place, and the editors would never have accepted the article for publication).

The interaction of these two streams, the business school stream and the economics department stream — the micro normative and the macro normative — has largely governed the history of the field of finance to date. I propose to review some of the highpoints of this history, taking full advantage of a handy organizing principle nature has given us: to wit, the Nobel Prizes in Finance.

Let me emphasize that I will not be offering a comprehensive survey of the field — the record is far too extensive for that — but rather a selective view of what I see as the highlights, an eyewitness account, as it were, and always with special emphasis on the tensions between the business school and the economics department streams.

After my overview, I offer some very personal views on where I think the field is heading, or at least where I would be heading were I just entering the field today.

MARKOWITZ AND THE THEORY OF PORTFOLIO SELECTION

The tension between the micro and macro approaches was visible from the very beginning of modern finance — from our big bang, as it were — which I think we can all agree today dates to the year 1952 with the publication in the *Journal of Finance* of Harry Markowitz's article, "Portfolio Selection." Markowitz in this remarkable paper gave, for the first time, a precise definition of what had hitherto been just vague buzzwords: risk and return.

Specifically, Markowitz then identified the yield or return on an investment with the expected value or probability-weighted mean value of its possible outcomes; and its risk with the variance or squared deviations of those outcomes around the mean. This identification of return and risk with mean and variance, so instinctive to finance professionals these days, was far from obvious then. The common perception of risk even today focuses on the likelihood of losses — on what the public thinks of as the "downside" risk — not just on the *variability* of returns.

Markowitz's choice of the variance as his measure of risk, counterintuitive as it may have appeared to many at the time, turns out to have been inspired. It not only subsumes the more intuitive view of risk — because in the normal or at least the symmetric distributions we use in practice the downside risk is essentially the mirror image of the upside — but it also has a property even more important for the development of the field. By identifying return and risk with mean and variance, Markowitz makes the powerful algebra of mathematical statistics available for the study of portfolio selection.

The immediate contribution of that algebra is the famous formula for the variance of a *sum* of random variables; that is, the weighted sum of the variance *plus* twice the weighted sum of the covariances. We in finance have been living on that formula, literally, for more than forty years now. That formula shows, among other things, that for the individual investor, the relevant unit of analysis must always be the whole portfolio, not the individual share. The risk of an individual share cannot be defined apart from its relation to the whole portfolio and, in particular, its covariances with the other components. Covariances, and not mere numbers of securities held, govern the risk-reducing benefits of diversification.

The Markowitz mean-variance model is the perfect example of what I call the business school or micro normative stream in finance. And this is somewhat ironic, in that the Markowitz paper was originally a thesis in the University of Chicago's economics department. Markowitz even notes that Milton Friedman, in fact, voted against the thesis initially on the grounds that it wasn't really economics.

And indeed, the mean-variance model, as visualized by Markowitz, really *wasn't* economics. Markowitz saw investors as actually applying the model to pick their portfolios using a combination of past data and personal judgment to select the needed means, variances, and covariances.

For the variances and covariances, at least, past data probably *could* provide at least a reasonable starting point. The precision of such estimates can always be enhanced by cutting the time interval into smaller and smaller intervals. But what of the means? Simply averaging the returns of the last few years, along the lines of the examples in the Markowitz paper (and later book) won't yield reliable estimates of the return *expected* in the future. And running those unreliable estimates of the means through the computational algorithm can lead to weird, corner portfolios that hardly seem to offer the presumed benefits of diversification, as any finance instructor who has assigned the portfolio selection model as a classroom exercise can testify.

If the Markowitz mean-variance algorithm is useless for selecting optimal portfolios, why do I take its publication as the starting point of modern finance? Because the essentially business school model of Markowitz was transformed by William Sharpe, John Lintner, and Jan Mossin into an economics department model of enormous reach and power.

WILLIAM SHARPE AND THE CAPITAL ASSET PRICING MODEL

That William Sharpe was so instrumental in transforming the Markowitz business school model into an economics department model continues the irony. Markowitz, it will be recalled, submitted his thesis to an economics department, but Sharpe was always a business school faculty member, and much of his earlier work had been in the management science/operations research area. Sharpe also maintains an active consulting practice advising pension funds on their portfolio selection problems. Yet his capital asset pricing model is almost as perfect an example as you can find of an economists' macro normative model of the kind I have described.

Sharpe starts by imagining a world in which every investor is a Markowitz mean-variance portfolio selector. And he supposes further that these investors all share the same expectation as to returns, variances, and covariances. But if the inputs to the portfolio selection are the same, then every investor will hold exactly the same portfolio of risky assets. And because all risky assets must be held by somebody, an immediate implication is that every investor holds the "market portfolio," that is, an aliquot share of every risky security in the proportions in which they are outstanding.

At first sight, of course, the proposition that everyone holds the same portfolio seems too unrealistic to be worth pursuing. Keep in mind first, however, that the proposition applies only to the holdings of risky assets. It does not assume that every investor has the same degree of risk aversion. Investors can always reduce the degree of risk they bear by holding riskless bonds along with the risky stocks in the market portfolio; and they can increase their risk by holding negative amounts of the riskless asset; that is, by borrowing and leveraging their holdings of the market portfolio.

Second, the idea of investing in the market portfolio is no longer strange. Nature has imitated art, as it were. Shortly after Sharpe's work appeared, the market created mutual funds that sought to hold all the shares in the market in their outstanding proportions. Such index funds, or "passive" investment strategies, as they are often called, are now followed by a large and increasing number of investors, particularly by U.S. pension funds.

The realism or lack of realism of the assumptions underlying the Sharpe CAPM has never been a subject of serious debate within the profession, unlike the case of the Modigliani and Miller propositions to be considered later. The profession, from the outset, wholeheartedly adopted the Friedman positivist view: that what counts is not the literal accuracy of the assumptions, but the *predictions* of the model.

In the case of Sharpe's model, these predictions are striking indeed. The CAPM implies that the distribution of expected rates of return across all risky assets is a *linear* function of a single variable, namely, each asset's sensitivity to or covariance with the market portfolio, the famous beta, which becomes the natural measure of a security's risk. The aim of science is to explain a lot with a little, and few models in finance or economics do so more dramatically than the CAPM.

The CAPM not only offers new and powerful theoretical insights into the nature of risk, but also lends itself admirably to the kind of in-depth empirical investigation so necessary for the development of a new field like finance. And its benefits have not been confined narrowly to the field of finance. The great volume of empirical research testing the CAPM has led to major innovations in both theoretical and applied econometrics.

Although the single-beta CAPM managed to withstand more than thirty years of intense econometric investigation, the current consensus within the profession is that a single risk factor, although it takes us an enormous length of the way, is not quite enough for describing the cross-section of expected returns. Besides the market factor, two other pervasive risk factors have by now been identified for common stocks.

One is a size effect; small firms seem to earn higher returns than large firms, on average, even after controlling for beta or market sensitivity. The other is a factor, still not fully understood, but that seems reasonably well captured by the ratio of a firm's accounting book value to its market value. Firms with high book-to-market ratios appear to earn higher returns on average over long horizons than those with low book-to-market ratios after controlling for size and for the market factor.

That a three-factor model has now been shown to describe the data somewhat better than the single-factor CAPM should detract in no way, of course, from appreciation of the enormous influence of the original CAPM on the theory of asset pricing.

THE EFFICIENT MARKETS HYPOTHESIS

The mean-variance model of Markowitz and the CAPM of Sharpe et al. are contributions whose great scientific value was recognized by the Nobel Committee in 1990. A third major contribution to finance was recognized at the same time. But before describing it, let me mention a fourth major contribution that has done much to shape the development of the field of finance in the last twenty-five years, but that has so far not received the attention from the Nobel Committee I believe it deserves.

I refer, of course, to the efficient markets hypothesis, which says, in effect, that no simple rule based on already published and available information can generate above-normal rates of return. On this score of whether mechanical profit opportunities exist, the conflict between the business school tradition in finance and the economics department tradition has been and still remains intense.

The hope that studying finance might open the way to successful stock market speculation served to support interest in the field even before the modern scientific foundations were laid in the 1950s. The first systematic collection of stock market prices, in fact, was compiled under the auspices of the Alfred Cowles Foundation in the 1930s.

Cowles had a lifelong enthusiasm for the stock market, dimmed only slightly by the catastrophic crash of 1929. The Cowles Foundation, currently an adjunct of the Yale University economics department, was the source of much fundamental research on econometrics in the 1940s and '50s.

The Cowles indexes of stock prices have long since been superseded by much more detailed and computerized data bases, such as those of the Center for Research in Security Prices at the University of Chicago. And to those computer data bases, in turn, goes much of the credit for stimulating the empirical research in finance that has given the field its distinctive flavor.

Even before these new computerized data bases came into widespread use in the early 1960s, however, the mechanical approach to above-normal investment returns was already being seriously challenged. The challenge was delivered, curiously enough, not by economists, but by statisticians like M.G. Kendall and my colleague, Harry Roberts — who argued that stock prices are essentially random walks. This implies, among other things, that the record of past stock prices, however rich in "patterns" it might appear, has no predictive power for future stock returns.

By the late 1960s, however, the evidence was accumulating that stock prices are not random walks by the strictest definition of that term. Some elements of predictability *could* be detected, particularly in long-run returns. The issue of whether publicly available information could be used for successful stock market speculation had to be rephrased — a task in which my colleague, Eugene Fama, played the leading role — as whether the observed departures from randomness in the time series of returns on common stocks represent true profit opportunities after transaction costs and after appropriate compensation for changes in risk over time. With this shift in focus from returns to cost- and risk-adjusted returns, the efficient markets debate becomes no longer a matter of statistics, but one of economics.

This connection with economics helps explain why the efficient markets hypothesis of finance remains as strong as ever, despite the steady drumbeat of empirical studies directed against it. If you find some mechanical rule that seems to earn above-normal returns — and with thousands of researchers spinning through the mountains of tapes of past data, anomalies, like the currently fashionable "momentum effects," are bound to keep turning up — then imitators will enter and compete away those above-normal returns exactly as in any other setting in economics. Above-normal profits, wherever they are found, inevitably carry with them the seeds of their own decay.

THE MODIGLIANI-MILLER PROPOSITIONS

Still other pillars on which the field of finance rests are the *Modigliani-Miller propositions* on capital structure. Here, the tensions between the micro normative and the macro normative approaches were evident from the outset, as is clear from the very title of the first M&M paper, "The Cost of Capital, Corporation Finance and the Theory of Investment." The theme of that paper, and indeed of the whole field of corporate finance at the time, is capital budgeting.

The micro normative wing was concerned with finding the "cost of capital," in the sense of the optimal cutoff rate for investment when the firm can finance the project either with debt or equity or some combination of both. The macro normative or economics wing sought to express the aggregate demand for investment by corporations as a function of the cost of capital that firms are actually using as their optimal cutoffs, rather than just the rate of interest on long-term government bonds.

The M&M analysis provided answers, but ones that left both wings of the profession dissatisfied. At the macro normative level, the M&M measure of the cost of capital for aggregate investment functions never really caught on, and, indeed, the very notion of estimating aggregate demand functions for investment has long since been abandoned by macro economists. At the micro level, the M&M propositions imply that the choice of financing instrument is irrelevant for the optimal cutoff. Such a cutoff is seen to depend solely on the risk (or "risk class") of the investment, regardless of how it is financed, hardly a happy position for professors of finance to explain to their students being trained, presumably, in the art of selecting optimal capital structures.

Faced with the unpleasant action consequences of the M&M model at the micro level, the tendency of many at first was to dismiss the assumptions underlying M&M's then-novel arbitrage proof as unrealistic. The assumptions underlying the CAPM, of course, are equally or even more implausible, as noted earlier, but the profession seemed far more willing to accept Friedman's "the assumptions don't matter" position for the CAPM than for the M&M propositions.

The likely reason is that the second blade of the Friedman positivism slogan — what *does* count is the descriptive power of the model itself — was not followed up. Tests by the hundreds of the CAPM fill the literature. But direct calibration tests of the M&M propositions and their implications do not.

One fundamental difficulty of testing the M&M propositions shows up in the initial M&M paper itself. The capital structure proposition says that if you could find two firms whose underlying earnings are identical, then so would be their market values, regardless of how much of the capital structure takes the form of equity as opposed to debt.

But how do you find two companies whose earnings are identical? M&M tried using industry as a way of holding earnings constant, but this sort of filter is far too crude. Attempts to exploit the power of the CAPM for testing M&M were no more successful. How do you compute a beta for the underlying real assets?

One way to avoid the difficulty of not having two identical firms, of course, is to see what happens when the *same* firm changes its capital structure. If a firm borrows and uses the proceeds to pay its shareholders a huge dividend or to buy back shares, does the value of the firm increase? Many studies have suggested that it does. But the interpretation of such results faces a hopeless identification problem.

The firm, after all, never issues a press release saying "we are just conducting a purely scientific investigation of the M&M propositions." The market, which is forward-looking, has every reason to believe that the capital structure decisions are conveying management's views about changes in the firm's prospects for the future. These confounding "information effects," present in every dividend and capital structure decision, render indecisive all tests based on specific corporate actions.

Nor can we hope to refute the M&M propositions indirectly by calling attention to the multitude of new securities and of variations on old securities that are introduced year after year. The M&M propositions say only that no gains could be earned from such innovations if the market were in fact "complete." But the new securities in question may well be serving to complete the market, earning a first-mover's profit to the particular innovation. Only those in Wall Street know how hard it is these days to come by those innovator's profits.

If all this seems reminiscent of the efficient markets hypothesis, that is no accident. The M&M propositions are also ways of saying that there is no free lunch. Firms cannot hope to gain by issuing what looks like low-cost debt rather than high-cost equity. They just make the cost of higher-cost equity even higher. And if any substantial number of firms, at the same time, seek to replace what they think is their high-cost equity with low-cost debt (even tax-advantaged debt), then the interest costs of debt will rise, and the required yields on equity will fall until the perceived incentives to change capital structures (or dividend policies for that matter) are eliminated.

The M&M propositions, in short, like the efficient markets hypothesis, are about *equilibrium* in the capital markets — what equilibrium looks like, and what forces are set in motion once it is disturbed. And this is why neither the efficient markets hypothesis nor the Modigliani-Miller propositions have ever set well with those in the profession who see finance as essentially a branch of management science.

OPTIONS

Fortunately, however, recent developments in finance, also recognized by the Nobel Committee, suggest that the conflict between the two traditions in finance, the business school stream and the economics department stream, may be on the way to reconciliation.

This development, of course, is the field of options, whose pioneers, recently honored by the Nobel Committee, were Robert Merton and Myron Scholes (with the late Fischer Black everywhere acknowledged as the third pivotal figure). Because the intellectual achievement of their work has been commemorated over and over — and rightly so — I will not seek to review it here. Instead, in keeping with my theme, I want to focus on what options mean for the history of finance.

Options mean, among other things, that for the first time in its close to fifty-year history, the field of finance can be built, or as I will argue be rebuilt, on the basis of "observable" magnitudes. 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 parry by reminding them of their neutrino — a particle with no mass whose presence is 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.

To say that option prices are based on observables is not strictly true, of course. The option price in the Black-Scholes-Merton formula depends on the current market value of the underlying share, the striking price, the time to maturity of the contract, and the risk-free rate of interest, all of which are observable either exactly or very closely. But the option price depends also, and very critically, on the variance of the distribution of returns on the underlying share, which is not directly observable; it must be estimated.

Still, as Fischer Black always reminded us, estimating variances is orders of magnitude easier than estimating the means or expected returns that are central to the models of Markowitz, Sharpe, or Modigliani-Miller. The precision of an estimate of the variance can be improved, as noted earlier, by cutting time into smaller and smaller units — from weeks to days to hours to minutes. For means, however, the precision of estimate can be enhanced only by lengthening the sample period, giving rise to the well-known dilemma that by the time a high degree of precision in estimating the mean from past data has been achieved, the mean itself has almost surely shifted.

Having a base in observable quantities — or virtually observable quantities — on which to value securities might seem at first sight to have benefited primarily the management science stream in finance. And indeed, recent years have seen the birth of a new and rapidly growing specialty area within the profession, that of financial engineering (and the recent establishment of a journal with that name is a clear sign that the field is here to stay). The financial engineers have already reduced the original Black-Scholes-Merton formula to Model-T status.

Nor has the micro normative field of *corporate* finance been left out. When it comes to capital budgeting, long a major focus of corporate finance, the decision impact of what have come to be called "real" options — even simple ones like the right to close down a mine when the output price falls and reopen it when it rises — is substantially greater than that of variations in the cost of capital.

The options revolution, if I may call it that, is also transforming the macro normative or economics stream in finance. The hint of things to come in that regard is prefigured in the title of the original Black-Scholes paper, "The Pricing of Options and Corporate Liabilities." The latter phrase was added to the title precisely to convince the editors of the *Journal of Political Economy* — about as economics a journal as you can get — that the original (rejected) version of the paper was not just a technical *tour de force* in mathematical statistics, but an advance with wide application for the study of market prices.

And indeed, the Black-Scholes analysis shows, among other things, how options serve to "complete the market" for securities by eliminating or at least substantially weakening the constraints on high leverage obtainable with ordinary securities. The Black-Scholes demonstration that the shares in highly leveraged corporations are really call options also serves in effect to complete the M&M model of the pricing of corporate equities subject to the prior claims of the debtholders. We can go even further: *Every* security can be thought of as a package of component Arrow-Debreu state-price contingent claims (options, for short), just as every physical object is a package of component atoms and molecules.

RECONSTRUCTION OF FINANCE?

I will speculate no further about these and other exciting prospects for the future. Let me close rather with a question: What would I advise a young member of the German Finance Association to specialize in? What would I specialize in if I were starting over and entering the field today?

Well, I certainly wouldn't go into asset pricing or corporate finance. Research in those subfields has already reached the phase of rapidly diminishing returns. Agency theory, I would argue, is best left to the legal profession, and behavioral finance is best left to the psychologists. So, at the risk of sounding a bit like the character in the movie "The Graduate," I reduce my advice to a single word: options.

When it comes to research potential, options have much to offer both the management science/business school wing within the profession *and* the economics wing. In fact, so vast are the research opportunities for both wings that the field is surely due for a total reconstruction as profound as that following the original breakthrough by Harry Markowitz in 1952.

The shift toward options as the center of gravity of finance that I foresee should be particularly welcomed by the members of the German Finance Association. I can remember when research in finance in Germany was just beginning and tended to consist of replication of American studies using German data. But when it comes to a relatively new area like options, we all stand roughly equal at the starting line. And this is an area in which the rigorous and mathematical German academic training may even offer a comparative advantage.

It is no accident, I believe, that the Deutsche Termin Borse (or Eurex, as it has now become after merging with the Swiss exchange) has taken the high-tech road to a leading position among the world's futures exchanges only eight years after a great conference in Frankfurt where Hartmut Schmidt, Fischer Black, and I sought to persuade the German financial establishment that allowing futures and options trading would not threaten the German economy. Hardware and electronic trading were the key to DTB's success, but I see no reason why the German scholarly community cannot duplicate that success on the more abstract side of research in finance as well.

Whether it can should be clear by the time of the twenty-fifth annual meeting. I'm only sorry I won't be able to see that happy occasion.

ENDNOTE

This is a slightly modified version of an address delivered at the Fifth Annual Meeting of the German Finance Association in Hamburg on September 25, 1998.

Morgan Stanley

Journal of Applied Corporate Finance

SUMMER 2000 VOLUME 13.2

The History of Finance: An Eyewitness Account

by Merton H. Miller,
University of Chicago

THE HISTORY OF FINANCE: AN EYEWITNESS ACCOUNT

by Merton H. Miller,
*University of Chicago**

I am honored indeed to be Keynote Speaker at the Fifth Anniversary of the German Finance Association. Five years, of course, is not very old as professional societies go, but then neither is the field of finance itself. That field in its modern form really dates from the 1950s. In the 40 years since then, the field has come to surpass many, perhaps even most, of the more traditional fields of economics in terms of the number of students enrolled in finance courses, the number of faculty teaching finance courses and, above all, in the quantity and quality of their combined scholarly output.

The huge body of scholarly research in finance over the last 40 years falls naturally into two main streams. And no, I don't mean "asset pricing" and "corporate finance," but a deeper division that cuts across both those conventional subdivisions of the field. The division I have in mind is the more fundamental one between what I will call the Business School approach to finance and the Economics Department approach. Let me say immediately, however, that my distinction is purely "notional" not physical—a distinction over what the field is really all about, not where the offices happen to be located. In the U.S., as I am sure you are aware, the vast majority of academics in finance are, and always have been, teaching in Business Schools, not Economics Departments. I should add immediately, however, that in the elite schools at least, a substantial fraction of the finance faculties have been trained in—that is, have received their Ph.D.s from—Economics Departments. Habits of thought acquired in graduate school have a tendency to stay with you.

The characteristic Business School approach tends to be what we would call in our jargon "micro normative." That is, a decision-maker, be it an individual investor or a corporate manager, is seen as maximizing some objective function, be it utility, expected return or shareholder value, taking the prices of securities in the market as given. In a Business School, after all, that's what you're supposed to be doing: teaching your charges how to make better decisions. To someone trained in the classical traditions of economics, however, the famous dictum of the great Alfred Marshall stands out: "It is not the business of the economist to tell the brewer how to make beer." The characteristic Economics Department approach thus is not micro, but *macro* normative. Their models assume a world of micro optimizers, and deduce from that how the market prices, which the micro optimizers take as given, actually evolve.

Note that I am differentiating the stream of research in finance along macro versus micro lines and not along the more familiar normative versus positive line. Both streams of research in finance are thoroughly positivist in outlook in that they try to be, or at least claim to be, concerned with testable hypotheses. The normal article in finance journals over the last 40 years has two main sections: one where the model is presented, and the second an empirical section showing that real-world data are consistent with the model (which is hardly surprising because had that not been so, the author would never have submitted the paper in the first place and the editors would never have accepted it for publication).

The interaction of these two streams, the Business School stream and the Economics Department

*A Keynote Address presented at the Fifth Annual Meeting of the German Finance Association in Hamburg, Germany, September 25, 1998. It was first published in the Summer 1999 issue of the *Journal of Portfolio Management*, a publication of Institutional Investor.

stream—the micro normative and the macro normative—has largely governed the history of the field of finance to date. I propose to review some of the highpoints of that history, taking full advantage of a handy organizing principle nature has given us—to wit, the Nobel prizes in finance. Let me emphasize again that I will not be offering a comprehensive survey of the field—the record is far too large for that—but rather a selective view of what I see as the highlights, an eyewitness account, as it were, and always with special emphasis on the tensions between the Business School and the Economics Department streams. After that overview I will offer some very personal views on where I think the field is heading, or at least where I would be heading were I just entering the field today.

MARKOWITZ AND THE THEORY OF PORTFOLIO SELECTION

The tension between the micro and macro approaches was visible from the very beginning of modern finance—from our big bang, as it were—which I think we can all agree today dates to the year 1952 with the publication in the *Journal of Finance* of Harry Markowitz's article "Portfolio Selection." Markowitz in that remarkable paper gave, for the first time, a precise definition of what had hitherto been just vague buzzwords, "risk" and "return." Specifically, Markowitz identified the yield or return on an investment with the expected value or probability-weighted mean value of its possible outcomes; and its risk with the variance or squared deviations of those outcomes around the mean. This identification of return and risk with Mean and Variance, so instinctive to finance professionals these days, was far from obvious then. The common perception of risk even today focuses on the likelihood of losses—on what the public thinks of as the "downside" risk—not just on the *variability* of returns. Yet Markowitz's choice of the Variance as his measure of risk, counterintuitive as it may have appeared to many at the time, turned out to be inspired. It not only subsumed the more intuitive view of risk—because in the normal (or at least the symmetric) distributions we use in practice the downside risk is essentially the mirror image of the upside—but it had a property even more important for the development of the field. By identifying return and risk with Mean and Variance, Markowitz made the powerful algebra of mathematical statistics available for the study of portfolio selection.

The immediate contribution of that algebra was the famous formula for the variance of a *sum* of random variables: the weighted sum of the variance *plus* twice the weighted sum of the covariances. We in finance have been living off that formula, literally, for more than 40 years now. That formula shows, among other things, that for the individual investor, the relevant unit of analysis must always be the whole portfolio, not the individual share. The risk of an individual share cannot be defined apart from its relation to the whole portfolio and, in particular, its covariances with the other components. Covariances, and not mere numbers of securities held, govern the risk-reducing benefits of diversification.

The Markowitz Mean-Variance model is the perfect example of what I have called the Business School or micro normative stream in finance. And that is somewhat ironic in that the Markowitz paper was originally a thesis in the University of Chicago's Economics Department. Markowitz even notes that Milton Friedman, in fact, voted against the thesis initially on the grounds that it wasn't really economics. And indeed, the Mean-Variance model, as visualized by Markowitz, really *wasn't* economics. Markowitz saw investors as actually applying the model to pick their portfolios using a combination of past data and personal judgment to select the needed Means, Variances, and Covariances.

For the Variances and Covariances, at least, past data probably *could* provide at least a reasonable starting point. The precision of such estimates can always be increased by cutting the time interval into smaller and smaller intervals. But what of the Means? Simply averaging the returns of the last few years, along the lines of the examples in the Markowitz paper (and later book) won't yield reliable estimates of the return *expected* in the future. And running those unreliable estimates of the Means through the computational algorithm can lead to weird, corner portfolios that hardly seem to offer the presumed benefits of diversification, as any finance instructor who has assigned the portfolio selection model as a classroom exercise can testify.

But if the Markowitz Mean-Variance algorithm is useless for selecting optimal portfolios, why have I taken its publication as the starting point of modern finance? Because that essentially Business School model of Markowitz was transformed by William Sharpe, John Lintner, and Jan Mossin into an Economics Department model of enormous reach and power.

WILLIAM SHARPE AND THE CAPITAL ASSET PRICING MODEL

That William Sharpe was so instrumental in transforming the Markowitz Business School model into an Economics Department model continues the irony noted earlier. Markowitz, it will be recalled, submitted his thesis to an Economics Department, but Sharpe was always a business school faculty member and much of his earlier work had been in the management science/operations research area. Sharpe also maintains an active consulting practice advising pension funds on their portfolio selection problems. Yet his Capital Asset Pricing Model is almost as perfect an example as you can find of an economists' macro-normative model of the kind I described.

Sharpe starts by imagining a world in which every investor is a Markowitz Mean-Variance portfolio selector. And he supposes further that these investors all share the same expectation as to returns, variances, and covariances. But if the inputs to the portfolio selection are the same, then every investor will hold exactly the same portfolio of risky assets. And because all risky assets must be held by somebody, an immediate implication is that every investor holds the "market portfolio," that is an aliquot share of every risky security in the proportions in which they are outstanding.

At first sight, of course, the proposition that everyone holds the same portfolio seems too unrealistic to be worth pursuing. Keep in mind first, however, that the proposition applies only to the holdings of risky assets. It does not assume that every investor has the same degree of risk aversion. Investors can always reduce the degree of risk they bear by holding riskless bonds along with the risky stocks in the market portfolio; and they can increase their risk by holding negative amounts of the riskless asset, that is by borrowing and leveraging their holdings of the market portfolio.

Second, the idea of investing in the market portfolio is no longer strange. Nature has imitated art, as it were. Shortly after Sharpe's work appeared, the market created mutual funds that sought to hold all the shares in the market in their outstanding proportions. Such index funds, or "passive" investment strategies, as they are often called, are now followed by a large and increasing number of investors, particularly, but by no means only, those of U.S. pension funds.

The realism or lack of realism of the assumptions underlying the Sharpe CAPM was never a subject of serious debate within the profession, unlike the case of the M&M propositions to be considered later. The profession, from the outset, wholeheartedly adopted the Friedman positivist view that what counts is not the literal accuracy of the assumptions, but the *predictions* of the model. And in the case of Sharpe's model, those predictions were striking indeed. The CAPM implies that the distribution of expected rates of return across all risky assets is a *linear* function of a single variable—namely each asset's sensitivity to or covariance with the market portfolio, the famous ß, which becomes the natural measure of a security's risk. The aim of science is to explain a lot with a little and few models in finance or economics do so more dramatically than the CAPM.

The CAPM not only offered new and powerful theoretical insights into the nature of risk, but also lent itself admirably to the kind of in-depth empirical investigation so necessary for the development of a new field like finance. Nor have the benefits been confined narrowly to the field of finance. The great volume of empirical research testing the CAPM has led to major innovations in both theoretical and applied econometrics.

Although the single-ß CAPM managed to withstand more than 30 years of intense econometric investigation, the current consensus within the profession is that a single risk factor, though it takes us an enormous length of the way, is not quite enough for describing the cross-section of expected returns. In addition to the market factor, two other pervasive risk factors have by now been identified for common stocks. One is a size effect: small firms seem to earn higher returns than large firms, on average, even after controlling for ß or market sensitivity. The other is a factor, still not fully understood, but which seems reasonably well captured by the ratio of a firm's accounting book value to its market value. Firms with high book-to-market ratios appear to earn higher returns on average over long horizons than those with low book-to-market ratios, after controlling for size and for the market factor. That a three-factor model has now been shown to describe the data somewhat better than the single factor CAPM should detract in no way, of course, from our appreciating the enormous influence on the theory of asset pricing exerted by the original CAPM.

In the past 50 years, the field of finance has come to surpass many, perhaps even most, of the more traditional fields of economics in terms of the number of students enrolled in finance courses, the number of faculty teaching finance courses, and, above all, in the quantity and quality of their combined scholarly output.

THE EFFICIENT MARKETS HYPOTHESIS

The Mean-Variance model of Markowitz and the CAPM of Sharpe et al. were contributions whose great scientific value were recognized by the Nobel Committee in 1990. A third major contribution to finance was recognized at the same time. But before describing it, let me mention a fourth major contribution that has done much to shape the development of the field of finance in the last 25 years, but which has so far not received the attention from the Nobel Committee I believe it deserves. I refer, of course, to the Efficient Markets Hypothesis, which says, in effect, that no simple rule based on already published and available information can generate above-normal rates of return. On this score of whether mechanical profit opportunities exist, the conflict between the Business School tradition in finance and the Economics Department tradition has been and still remains intense.

The hope that studying finance might open the way to successful stock market speculation served to keep up interest in the field even before the modern scientific foundations were laid in the 1950s. The first systematic collection of stock market prices, in fact, was compiled under the auspices of the Alfred Cowles Foundation in the 1930s. Cowles himself had a lifelong enthusiasm for the stock market, dimmed only slightly by the catastrophic crash of 1929. Cowles is perhaps better known by academic economists these days as the sponsor of the Cowles Foundation, currently an adjunct of the Yale Economics Department and the source of much fundamental research on econometrics in the 1940s and '50s. Cowles' indexes of stock prices have long since been superseded by much more detailed and computerized databases, such as those of the Center for Research in Security Prices at the University of Chicago. And to those computer databases, in turn, goes much of the credit for stimulating the empirical research in finance that has given the field its distinctive flavor.

Even before these new computerized indexes came into widespread use in the early 1960s, however, the mechanical approach to above-normal investment returns was already being seriously challenged. That challenge was being delivered, curiously enough, not by economists, but by statisticians like M.G. Kendall and my colleague Harry Roberts—who argued that stock prices were essentially random walks. That implied, among other things, that the record of past stock prices, however rich in "patterns" it might appear, had no predictive power for future stock prices and returns.

By the late 1960s, however, the evidence was clear that stock prices were not random walks by the strictest definition of that term. Some elements of predictability *could* be detected particularly in long-run returns. The issue of whether publicly available information could be used for successful stock market speculation had to be rephrased—a task in which my colleague Eugene Fama played the leading role—as whether the observed departures from randomness in the time series of returns on common stocks represented true profit opportunities after transaction costs and after appropriate compensation for changes in risk over time. With that shift in focus from returns to cost- and risk-adjusted returns, the Efficient Markets debate was no longer a matter of statistics, but one of economics.

This tieback to economics helps explain why the Efficient Market Hypothesis of finance remains as strong as ever despite the steady drumbeat of empirical studies directed against it. Suppose you find some mechanical rule that seems to earn above normal returns—and with thousands of researchers spinning through the mountains of tapes of past data, anomalies, like the currently fashionable "momentum effects," are bound to keep turning up. Then imitators will enter and compete away those above-normal returns exactly as in any other setting in economics. Above-normal profits, wherever they are found, inevitably carry with them the seeds of their own decay.

THE MODIGLIANI-MILLER PROPOSITIONS

Still other pillars on which the field of finance rests are the Modigliani-Miller Propositions on capital structure. Here, the tensions between the micro normative and the macro normative approaches were evident from the outset, as is clear from the very title of the first M&M paper, "The Cost of Capital, Corporation Finance and the Theory of Investment." The theme of that paper, and indeed of the whole field of corporate finance at the time, was capital budgeting. The micro normative wing was concerned with the "cost of capital," in the sense of the optimal "cut off" rate for investment when the firm can finance the project either with debt or equity or some combination of both. The macro normative or economics wing sought to express the aggregate demand for investment by corporations as a function of the cost of capital that firms were actually using as their optimal cutoffs, rather than just the rate of interest on long-term government bonds. The M&M analysis provided answers that left both wings of the profession dissatisfied. At the macro normative level, the M&M measure of the cost of capital for aggregate investment functions never really caught on, and, indeed, the very notion of estimating aggregate demand functions for investment has long since been abandoned by macro economists. At the micro level, the M&M proportions implied that the choice of financing instrument was irrelevant for the optimal cut-off. That cut-off depended solely on the risk (or "risk-class") of the investment regardless of how it was financed, hardly a happy position for professors of finance to explain to their students being trained presumably in the art of selecting optimal capital structures.

Faced with the unpleasant action-consequences of the M&M model at the micro level, the tendency of many at first was to dismiss the assumptions underlying M&M's then-novel arbitrage proof as unrealistic. The assumptions underlying the CAPM, of course, are equally or even more implausible, as noted earlier, but the profession seemed far more willing to accept Friedman's "the assumptions don't matter" position for the CAPM than for the M&M Propositions. The likely reason is that the second blade of the Friedman positivism slogan—what *does* count is the descriptive power of the model itself—was not followed up. Tests by the hundreds of the CAPM filled the literature. But direct calibration tests of the M&M Propositions and their implications did not exist.

One fundamental difficulty of testing the M&M Propositions showed up in the initial M&M paper itself. The capital structure proposition says that if you could find two firms whose underlying earnings were identical, then so would be their market values, regardless of how much of the capital structure took the form of equity as opposed to debt. But how do you find two companies whose earnings are identical? M&M tried using industry as a way of holding earnings constant, but that sort of filter was far too crude to be decisive. Attempts to exploit the power of the CAPM were no more successful. How do you compute a ß for the underlying real assets?

One way to avoid the difficulty of not having two identical firms, of course, is to see what happens when the *same* firm changes its capital structure. If a firm borrows and uses the proceeds to pay its shareholders a huge dividend or to buy back shares, does the value of the firm increase? Many studies have suggested that they do. But the interpretation of those results faces a hopeless identification problem. The firm, after all, never issues a press release saying we are just conducting a purely scientific investigation of the M&M Propositions. The market, which is forward looking, has every reason to believe that these capital structure decisions are conveying management's views about changes in the firm's prospects for the future. These confounding "information effects," present in every dividend and capital structure decision, render indecisive all tests based on specific corporate actions.

Nor can we hope to refute the M&M Propositions indirectly by calling attention to the multitude of new securities and of variations on old securities that are introduced year after year. The M&M Propositions say only that no gains could be earned from such innovations if the market were in fact "complete." But the new securities in question may well be serving to complete the market, earning a first-mover's profit to the particular innovation. Only those in Wall Street know how hard it is these days to come by those innovator's profits.

If all this seems reminiscent of the Efficient Markets Hypothesis, that is no accident. The M&M Propositions are also ways of saying that there are no free lunches. Firms cannot hope to gain by issuing what looks like low-cost debt rather than high-cost equity. They just make the higher cost equity even higher. And if any substantial number of firms, at the same time, sought to replace what they think is their high-cost equity with low-cost debt (even tax-advantaged debt), then the interest costs of debt would rise and the required yields on equity would fall until the perceived incentives to change capital structures (or dividend policies for that matter) were eliminated. The M&M Propositions, in short, like the Efficient Markets Hypothesis, are about *equilibrium* in the capital markets—what equilibrium looks like and what forces are set in motion once it is disturbed. And that is why neither the Efficient Markets Hypothesis nor the Modigliani-Miller propositions have ever set well with those in the profession who see finance as essentially a branch of management science.

Fortunately, however, recent developments in finance, also recognized by the Nobel Committee, suggest that the conflict between the two traditions in finance, the Business School stream and the Economics Department stream, may be on the way to reconciliation.

Options mean that, for the first time in its close to 50-year history, the field of finance can be built, or as I will argue be rebuilt, on the basis of "observable" magnitudes. When it comes to capital budgeting, for example, the decision impact of what have come to be called "real" options is substantially greater than that of variations in the cost of capital.

OPTIONS

That new development, of course, is the field of options, whose pioneers, recently honored by the Nobel Committee, were Robert Merton and Myron Scholes (with the late Fischer Black everywhere acknowledged as the third pivotal figure). Because the intellectual achievement of their work has been memorialized over and over this past year—and rightly so—I will not seek to review it here. Instead, in keeping with my theme today, I want to focus on what options mean for the history of finance.

Options mean, among other things, that for the first time in its close to 50-year history, the field of finance can be built, or as I will argue be rebuilt on the basis of "observable" magnitudes. 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.

To say that option prices are based on observables is not strictly true, of course. The option price in the Black-Scholes-Merton formula depends on the current market value of the underlying share, the striking price, the time to maturity of the contract, and the risk-free rate of interest, all of which are observable either exactly or very closely. But the option price depends also, and very critically, on the *variance* of the distribution of returns on the underlying share, which is not directly observable; it must be estimated. Still, as Fischer Black always reminded us, estimating variances is orders of magnitude easier than estimating the means or expected returns that are central to the models of Markowitz, Sharpe, or Modigliani-Miller. The precision of an estimate of the variance can be increased, as noted earlier, by cutting time into smaller and smaller units—from weeks to days to hours to minutes. For means, however, the precision of estimate can be increased only by lengthening the sample period, giving rise to the well-known dilemma that by the time a high degree of precision in estimating the mean from past data has been achieved, the mean itself has almost surely shifted.

Having a base in observable quantities—or virtually observable quantities—on which to value securities might seem at first sight to have benefited primarily the management science stream in finance. And, indeed, recent years have seen the birth of a new and rapidly growing specialty area within the profession, that of financial engineering (with the recent establishment of a journal with that name a clear sign that the field is here to stay). The financial engineers have already reduced the original Black-Scholes-Merton formula to model-T status. Nor has the micro normative field of *corporate* finance been left out. When it comes to capital budgeting, long a major focus of that field, the decision impact of what have come to be called "real" options—even simple ones like the right to close down a mine when the output price falls and reopen it when it rises—is substantially greater than that of variations in the cost of capital.

The options revolution, if I may call it that, is also transforming the macro normative or economics stream in finance. The hint of things to come in that regard was prefigured in the title of the original Black-Scholes paper itself, "The Pricing of Options and Corporate Liabilities." The latter phrase was added to the title precisely to convince the editors of the *Journal of Political Economy*—about as economicsy a journal as you can get—that the original (rejected) version of their paper was not just a technical *tour de force* in mathematical statistics, but an advance with wide applicability for the study of market prices.

And indeed, the Black-Scholes analysis showed, among other things, how options serve to "complete the market" for securities by eliminating or at least substantially weakening the constraints on high leverage obtainable with ordinary securities. The Black-Scholes demonstration that the shares in highly leveraged corporations are really call options also serves in effect to complete the M & M model of the pricing of corporate equities subject to the prior claims of the debt holders. But we can go even further. *Every* security can be thought of as a package of component Arrow-Debreu state-price options, just as every physical object is a package of component atoms and molecules.

But I propose to speculate no further about these and other exciting prospects for the future. Let me close rather with the question I raised in the beginning: what would I advise a young member of the German Finance Association to specialize in?

What would *I* specialize in if I were starting over and entering the field today?

Well, I certainly wouldn't go into asset pricing or corporate finance. Research in those subfields has already reached the phase of rapidly diminishing returns. Agency theory, I would argue, is best left to the legal profession and behavioral finance is best left to the psychologists. So at the risk of sounding a bit like the character in the movie "The Graduate," I reduce my advice to a single word: options. When it comes to research potential, options have much to offer both the management-science business-school wing within the profession *and* the economics wing. In fact, so vast are the research opportunities for both wings that the field is surely due for a total reconstruction as profound as that following the original breakthrough by Harry Markowitz in 1953.

The shift towards options in the center of gravity of finance that I foresee should be particularly welcomed by the members of the German Finance Association. I can remember when research in finance in Germany was just beginning and tended to consist of copies of American studies using German data. But when it comes to a relatively new area like options, we all stand roughly equal at the starting line. And it's an area in which the rigorous and mathematical German academic training may even offer a comparative advantage.

It is no accident, I believe, that the Deutsche Termin Borse (or Eurex, as it has now become after merging with the Swiss exchange) has taken the high-tech road to a leading position among the world's future exchanges only eight years after a great conference in Frankfurt where Hartmut Schmidt, Fischer Black, and I sought to persuade the German financial establishment that allowing futures and options trading would not threaten the German economy. Hardware and electronic trading were the key to DTB's success; but I see no reason why the German scholarly community can't duplicate that success on the more abstract side of research in finance as well.

Whether they can should be clear by the time of your 25th Annual Meeting. I'm only sorry I won't be able to see that happy occasion.

■ MERTON MILLER

was Robert R. McCormick Distinguished Service Professor Emeritus at the University of Chicago's Graduate School of Business. He was awarded the Nobel Prize in economics in 1990.

Journal of Applied Corporate Finance (ISSN 1078-1196 [print], ISSN 1745-6622 [online]) is published quarterly on behalf of Morgan Stanley by Blackwell Publishing, with offices at 350 Main Street, Malden, MA 02148, USA, and PO Box 1354, 9600 Garsington Road, Oxford OX4 2XG, UK. Call US: (800) 835-6770, UK: +44 1865 778315; fax US: (781) 388-8232, UK: +44 1865 471775, or e-mail: subscrip@bos.blackwellpublishing.com.

Information For Subscribers For new orders, renewals, sample copy requests, claims, changes of address, and all other subscription correspondence, please contact the Customer Service Department at your nearest Blackwell office.

Subscription Rates for Volume 17 (four issues) Institutional Premium Rate* The Americas† $330, Rest of World £201; Commercial Company Premium Rate, The Americas $440, Rest of World £268; Individual Rate, The Americas $95, Rest of World £70, €105‡; Students**, The Americas $50, Rest of World £28, €42.

*Includes print plus premium online access to the current and all available backfiles. Print and online-only rates are also available (see below).

†Customers in Canada should add 7% GST or provide evidence of entitlement to exemption

‡Customers in the UK should add VAT at 5%; customers in the EU should also add VAT at 5%, or provide a VAT registration number or evidence of entitlement to exemption

** Students must present a copy of their student ID card to receive this rate.

For more information about Blackwell Publishing journals, including online access information, terms and conditions, and other pricing options, please visit www.blackwellpublishing.com or contact our customer service department, tel: (800) 835-6770 or +44 1865 778315 (UK office).

Back Issues Back issues are available from the publisher at the current single-issue rate.

Mailing *Journal of Applied Corporate Finance* is mailed Standard Rate. Mailing to rest of world by DHL Smart & Global Mail. Canadian mail is sent by Canadian publications mail agreement number 40573520. **Postmaster** Send all address changes to Journal of Applied Corporate Finance, Blackwell Publishing Inc., Journals Subscription Department, 350 Main St., Malden, MA 02148-5020.

Journal of Applied Corporate Finance is available online through Synergy, Blackwell's online journal service which allows you to:

- Browse tables of contents and abstracts from over 290 professional, science, social science, and medical journals
- Create your own Personal Homepage from which you can access your personal subscriptions, set up e-mail table of contents alerts and run saved searches
- Perform detailed searches across our database of titles and save the search criteria for future use
- Link to and from bibliographic databases such as ISI.

Sign up for free today at http://www.blackwell-synergy.com.

MARKETS

More Good Inflation News for Investors

Nov. 17, 2014 12:13 a.m. ET

A big reason stocks are doing so well now is something investors don't talk about a lot: exceptionally weak inflation.

Markets have gotten more good news on that front lately, as economists have been cutting their inflation forecasts even as the U.S. economy has strengthened.

Goldman Sachs, which had expected inflation to rise gradually, now says it will stay flat at just 1.5% a year until the end of 2015. Bank of America Merrill Lynch, which had forecast inflation at 1.6% for 2015, now predicts 1.4%.

Both are talking about core PCE inflation, the Federal Reserve's preferred gauge, which is based on consumer spending with volatile energy and food costs removed.

The Wall Street consensus for that inflation measure is about 1.8% in 2015, below the Fed's 2% target.

"That is very positive for stocks and for bonds," said investment strategist Edgar Peters of First Quadrant LP, which manages about $23 billion in Pasadena, Calif. Tame inflation "is high on the list" of the reasons stocks have been so strong, he said.

A November Merrill Lynch survey of professional bond investors showed less than 5% worried about inflation, down from more than 15% in September.

Low inflation helps stocks by holding down market interest rates. That means lower mortgage rates, credit-card bills and corporate financing costs. All of this supports consumer spending and corporate profits. Low interest rates also mean low bond yields, which make stocks look better than bonds to some investors.

In addition, weak inflation makes it less urgent for the Federal Reserve to raise its target interest rates. The longer the Fed waits, and the slower it goes once it starts raising rates, the happier stock and bond investors will be.

Until recently, Merrill was forecasting that the Fed would start raising rates in June. It has changed that to September. Goldman also says September, with a rising possibility it could happen later.

In Merrill's November survey, 59% of bond investors thought the Fed would start in the second half of 2015, up from 49% in September.

Low inflation "gives the Fed more time" before it has to act, said Jim McDonald, chief investment strategist at Northern Trust, which oversees $923 billion in Chicago. "It is very important" for stock prices, he said. He, too, forecasts that the Fed won't raise rates until the latter part of next year.

All of this helps explain why stocks have done better than expected this year. Money managers have warned of a 10% pullback at some point, together with weak overall stock gains. Instead, the S&P 500 is up 10% this year, not counting dividends, which is above the historical average of about 7%. It closed Friday at its 41st record of the year. Neither it nor the Dow Jones Industrial Average has pulled back as much as 10%. The Dow is up 6% for 2014, 18 points from another record of its own.

Low inflation also helps explain why the experts have been so wrong in forecasting bond yields. They widely expected the yield of the 10-year Treasury note to be 3.25% or more by now. Instead, it was at 2.32% Friday, down from 3% at the start of the year.

Economists cite many reasons for low inflation, and most of them come down to this: Economic growth is slow world-wide. Weak growth holds down oil and commodity prices. It restricts world-wide wage gains. All of that means lower inflation everywhere. And because the U.S. is an island of relative prosperity, the dollar is strong, which reduces import costs.

"Import prices have been falling for three months," said Michelle Meyer, senior U.S. economist at Bank of America Merrill Lynch. Low inflation "will be a supportive environment for the stock market, I think."

What could go wrong?

Inflation could get hotter. With U.S. job creation rising and unemployment falling, stronger wage gains seem likely at some point, which would push inflation higher.

Alternatively, global growth could get weak enough to hurt U.S. growth and job creation. That could squeeze corporate profits, which would hurt stocks even if inflation and interest rates stayed tame.

Complicating matters, stocks aren't cheap. The S&P 500 traded Friday at more than 19 times its companies' net profits for the past 12 months, well above its 15.5 historical average, according to Birinyi Associates. If the economic outlook turns cloudy, it won't take much for people to start selling stocks again.

Or, the Fed could decide to be more aggressive about raising rates than some investors expect. Fed Chairwoman Janet Yellen has warned that low interest rates can cause lending standards to decline, fueling financial-market bubbles. That kind of concern could make the Fed raise rates even if inflation remains low.

None of that seems to be looming now. In the late 1990s, Mr. Peters of First Quadrant was one of those warning that stocks were overdue for a severe decline. Today, he says he wouldn't be surprised to see stock volatility, even a 10% decline. But as long as inflation is low and the economy avoids recession, he adds, it is hard to see what would cause a bear market, meaning a drop of 20% or more.

"There really isn't any reason for the market to have a huge correction at this stage in the cycle," Mr. Peters said.

Is There a Link Between GDP Growth and Equity Returns? | May 2010

MSCI Barra RESEARCH BULLETIN

Introduction

A recurring question in finance concerns the relationship between economic growth and stock market return. Recently, for example, some emerging market countries have experienced spectacular growth, and many institutional investors wonder if they should assign a higher weight to these countries (based on gross domestic product [GDP] rather than market capitalization). These investors hope that this higher weight will be justified by a subsequent higher return.

This question is not new; "supply-side" models have been developed to explain and forecast stock market returns based on macroeconomic performance. These models are based on the theory that equity returns have their roots in the productivity of the underlying real economy and long term returns cannot exceed or fall short of the growth rate of the underlying economy.

In this research bulletin, we empirically test the steps leading from GDP growth to stock returns. We use long-term MSCI equity index data and macroeconomic data to conduct this analysis.

Mechanics of Supply-Side Models

Supply-side models assume that GDP growth of the underlying economy flows to shareholders in three steps. First, it transforms into corporate profit growth; second, the aggregate earnings growth translates into earnings per share (EPS) growth, and finally EPS growth translates into stock price increases.

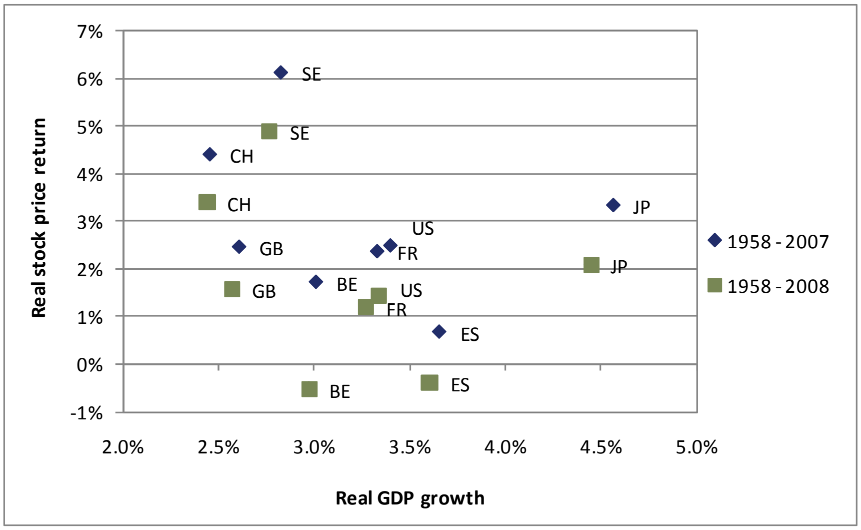
If we further assume that:

- the share of company profits in the total economy remains constant;
- investors have a claim on a constant proportion of those profits;
- valuation ratios are constant;
- the country's stock market only lists domestic companies;
- the country's economy is closed,

then we would expect an exact match between real price increase and real GDP growth. *This theory is simple and makes intuitive sense. But is it true in practice?*

Several studies (Dimson et al. [2002], Ritter [2005]) have examined whether countries with higher long-run real GDP growth also had higher long-run real stock market return. The surprising result was contrary to expectations -- the correlation between stock returns and economic growth across countries can be negative! Our own analysis confirms this empirical finding: Exhibit 1 plots stock returns versus GDP growth for eight developed markets between 1958 and 2008 and also shows negative correlation. Note, however, that these tests are dependent on the starting and ending point of the period analyzed; by changing the period by only one year to 1958-2007, we get very different results (although the observed correlation in this example is still negative). For example, the annualized return for Belgium is changed from 1.7% to -0.5%.

Exhibit 1: Annual real GDP growth versus annual real stock returns, 1958 – 2007 and 1958 – 2008



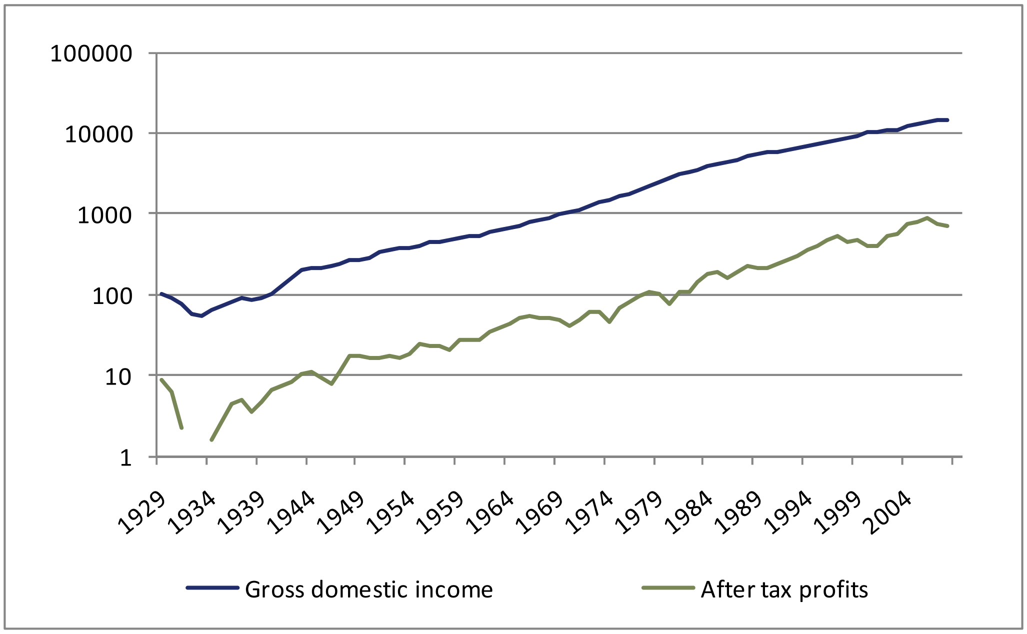
Source: MSCI Barra, IMF, OECD. Growth rates are annualized.

How can we reconcile these empirical findings with the theoretical argument? We will examine the steps leading from GDP growth to stock market performance and show that many assumptions of supply-side models can be challenged and need to be refined.

GDP and Aggregate Earnings

We start by examining the relationship between GDP and aggregate corporate earnings. In Exhibit 2, we use the United States as an example and plot US GDP and corporate earnings (which represent 4-6% of the GDP) from 1929 until 2008. We infer that growth of GDP and aggregate corporate earnings have been remarkably similar throughout the last 80 years, with the exception of 1932 and 1933 when profits were actually negative. This supports the first assumption of supply-side models: over the long run, aggregate corporate earnings tend to grow at the same pace as GDP.

Exhibit 2: Gross domestic product and after-tax corporate profits in the United States, 1929 – 2008



Source: US Department of Commerce, annual data as of 2008. Note that negative values cannot be represented on a log-scale graph.

Aggregate Earnings and EPS

We next examine the theory that aggregate corporate earnings growth translates into EPS growth. This assumption may be somewhat hasty (Bernstein and Arnott [2003]).There is indeed a distinction between growth in aggregate earnings of an economy and the growth in earnings per share to which *current* investors have a claim. These two growth rates do not necessarily match, since there are factors that can dilute aggregate earnings. A portion of GDP growth comes from capital increases, such as new share issuances, rights issues, or IPOs, which increase aggregate earnings but are not accessible to current investors. In fact, investors do not automatically participate in the profits of new companies. When buying shares of new businesses, they have to dilute their holdings in the "old" economy or invest additional capital. This dilution causes the growth in EPS available to current investors to be lower than growth in aggregate earnings. A simple measure of dilution suggested by Bernstein and Arnott is the difference between the growth of the aggregate market capitalization for a market and the performance of the aggregate index for that market. Based on very long term US data, this dilution is estimated to subtract 2% from real GDP growth.

EPS and Stock Prices

The last assumption in the theory that leads from GDP growth to equity performance is that EPS growth translates into stock price increases. This is only true however, if there are no changes in valuations (the price to earnings ratio) as illustrated by the equation below:

$$1 + r = (1 + g_{rEPS})(1 + g_{PE})$$

where r is the price return of the stock, g_{rEPS} is the growth rate in real earnings per share and g_{PE} is the growth rate in the price-to-earnings ratio. Some research claims that there are no reasons for valuations to change over the long term, which supports the supply-side models. However, empirical tests show that valuations have generally expanded over the last 40 years (see 'What Drives Long Term Equity Returns?' MSCI Barra [2010]). This can be explained in several ways, for example, due to different regimes (declining inflation), better market and information efficiency, or improved corporate governance.

Exhibit 3 correlates the historical data for the MSCI developed market countries over the last 40 years. To relate the data to economic growth, the last two columns display the amounts by which EPS and price returns have fallen compared to GDP growth rates.

We find that the mean "slippage" between real GDP growth and EPS growth is 2.3%. On average, stock prices have followed GDP more closely; the mean difference is only 0.3%. This is a consequence of the considerable expansion (2.0%) in the PE ratio during the same period that offset the earnings dilution effect.

Exhibit 3: Real GDP, real earnings per share, real price growth and price-to-earnings growth[1] for selected countries, 1969 – 2009

1969 - 2009	Real GDP growth rates	Real stock price return	Real EPS growth rates	PE change	GDP growth minus stock price return	GDP growth minus EPS growth
Australia	3.1%	0.0%	0.5%	-0.4%	3.1%	2.7%
Norway	3.0%	2.7%	0.9%	1.8%	0.3%	2.1%
Spain	3.0%	-1.4%	n. a.	n. a.	4.5%	n. a.
Canada	2.9%	2.5%	1.3%	1.1%	0.4%	1.6%
United States	2.8%	1.6%	0.0%	1.6%	1.2%	2.8%
Japan	2.8%	1.5%	not meaningful	not meaningful	1.3%	n. a.
Austria	2.6%	0.6%	-1.9%	2.6%	1.9%	4.6%
Netherlands	2.4%	1.9%	-2.6%	4.6%	0.5%	5.1%
France	2.3%	1.7%	n. a.	n. a.	0.6%	n. a.
Belgium	2.3%	0.6%	-2.8%	3.5%	1.7%	5.3%
United Kingdom	2.2%	1.1%	1.6%	-0.6%	1.1%	0.5%
Sweden	2.1%	5.8%	4.4%	1.3%	-3.5%	-2.3%
Italy	2.0%	-1.7%	n. a.	n. a.	3.8%	n. a.
Germany	1.8%	1.6%	-1.1%	2.7%	0.3%	2.9%
Denmark	1.7%	3.6%	1.2%	2.4%	-1.9%	0.5%
Switzerland	1.5%	2.6%	-0.5%	3.1%	-1.1%	2.0%
Average	2.4%	2.0%	0.1%	2.0%	0.3%	2.3%
MSCI ACWI[1]	2.7%	2.1%	0.6%	1.5%	0.6%	2.1%

Source: MSCI Barra, US Department of Agriculture, OECD. Average based on all countries excluding Spain, Japan, France, Italy.

From this data we infer that although the average long term equity performance was similar to GDP growth, this was due to the increasing valuations offsetting the dilution effect.Variance among countries is striking. In one extreme case, the EPS of the MSCI Sweden Index has grown 2.3% faster than Sweden's GDP and the index itself has performed 3.5% better than the GDP. At the other extreme, the MSCI Spain Index grew 4.5% slower than Spain's GDP.

International Considerations and Other Arguments

The prior examples suggest there may be complications in the simple model that has GDP mechanically flowing through to stock returns.

For example, part of the difference among countries may be explained by the different level of openness of the economies, and by the disparities in the proportion of listed companies. Indeed, a company's profit can be earned outside the country in which it is listed. As economic globalization continues, more firms operate in several locations throughout the world.

[1] The price return, EPS growth rate, and PE change for the MSCI All Country World Index (ACWI)I is based on a combination of MSCI World Index data prior to December 31, 1987, and MSCI ACWI data after that date. Similarly, real GDP growth is based on summing GDPs of countries included in the MSCI World Index prior to December 31, 1987, and in MSCI ACWI after that date.

Consequently, parts of the production process for these multinational firms are not reflected in the country's GDP. This can create a discrepancy between the company's performance and the local economy. On the other hand, the company's revenues and share price largely depend on the global GDP growth, as an increasing proportion of its products is sold abroad.

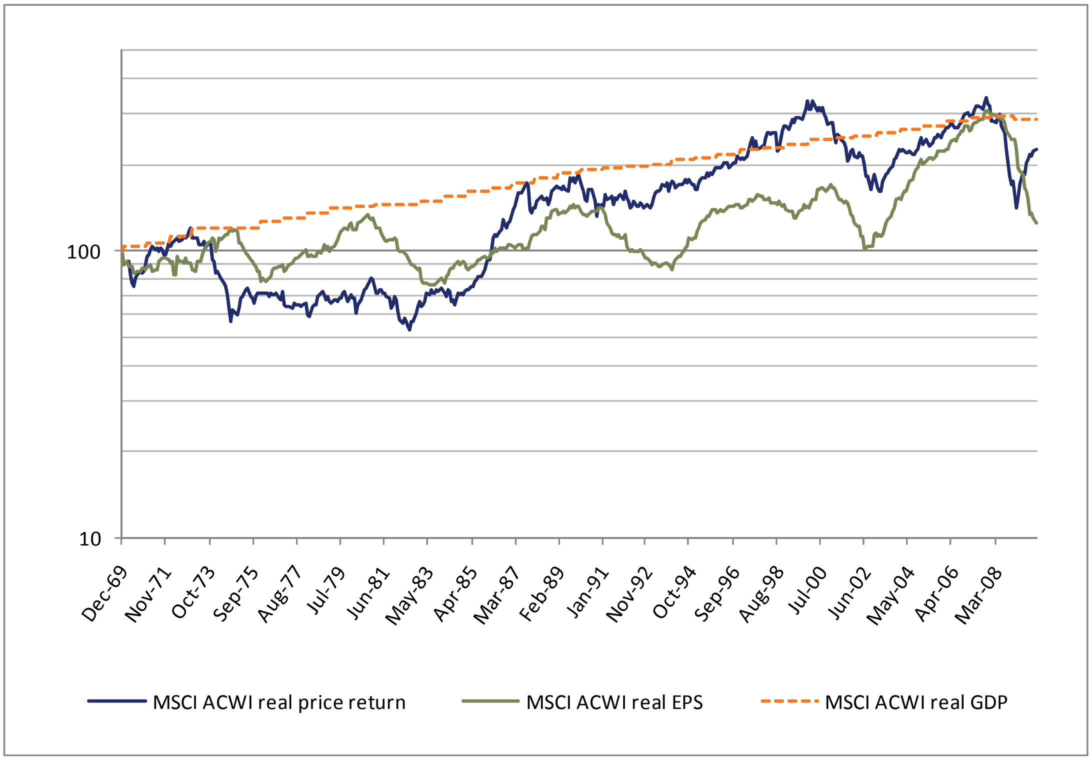
This decoupling effect is amplified because the biggest firms in each country, and consequently in each country index, tend to be multinational companies. This decoupling between company listing and company contribution to GDP may disappear if we consider an aggregate of countries. Indeed, by taking a large set of countries (ideally the whole global economy), the majority of production – even those of multinational firms – will become domestic and contribute to the aggregate GDP. When comparing the growth of this aggregate GDP to the performance of the aggregate stock market of the same set of countries, the distorting effect of companies listed in one country and producing in another can be almost totally discarded.

In Exhibit 4, we investigate this idea by looking at global equity returns as represented by a combination[2] of the MSCI All Country World Index (ACWI) and the MSCI World Index, and comparing them to the GDP growth of countries included in the same indices. The countries included in this combined index are a good approximation of the global economy. Although it only included 16 developed market countries in 1969 (US, Canada, Japan, Australia, and countries from Europe), those countries represented 78% percent of the global economic production, as measured by their real GDP. The coverage ratio jumped above 80% in 1988, when emerging markets are included in the combined index, and reached 93% in 2009.

Using this aggregation, we see that long term trends in real GDP and equity prices are more similar for global equities than for most individual markets. The annual real GDP growth rate of the MSCI World and MSCI ACWI countries between 1969 and 2009 was 2.7% and real price return was 2.1%. However, the dilution effect is still observable as real EPS grew at a 0.6% annual pace -- the wedge between GDP growth and EPS growth was 2.1% over the last 40 years, but real stock price lagged GDP growth by only 0.6%. This can be attributed to the extreme expansion in the PE ratio during the long bull market of the 1980s.

[2] Global equity return calculation is based on a combination of MSCI World Index returns prior to January 1, 1988, and MSCI ACWI returns after that date.

Exhibit 4: MSCI ACWI[3] real price return, real EPS and real GDP growth, 1969 – 2009

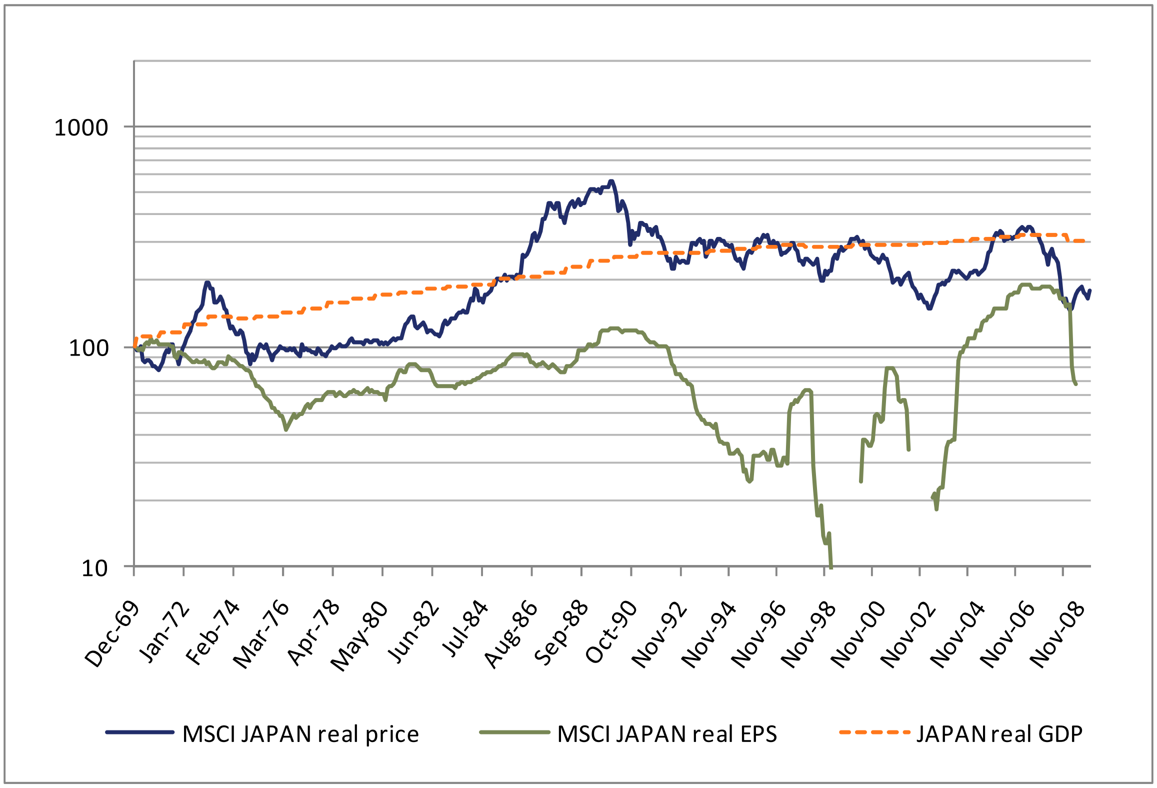


Source: MSCI Barra, US Department of Agriculture, data as of December 2009. Real GDP growth is shown as a chain-linked index to avoid the distorting effect of changes in the country composition of the corresponding global equity indices (MSCI World before January 1, 1988 and MSCI ACWI after that date). Real index and per share data is obtained by deflating by the global GDP deflator.

An additional argument by Siegel (1998) to explain the lack of observable correlation between GDP growth and stock returns is that expected economic growth is already impounded into the prices, thus lowering future returns. As shown in Exhibit 5, Japan is an example of this effect. We see that growth expectations were overly optimistic and 20 years of future growth were already discounted in the 1980s when stock prices grew faster than GDP. In the last two decades, equity performance was negative, while the GDP continued to grow.

[3] MSCI ACWI is replaced by the MSCI World Index prior to January 1, 1988.

Exhibit 5: MSCI Japan Index real price return, real EPS and real GDP growth, in JPY, 1969 – 2009



Source: MSCI Barra, US Department of Agriculture. Note that negative values cannot be represented on a log-scale graph.

Conclusions

We may intuitively think of stock returns as a result of the underlying real economy growth. However, we have observed that long term real earnings growth fell behind long term GDP growth in many countries over the observed period.

Several factors may explain this discrepancy. First, in today's integrated world we need to look at global rather than local markets. Second, a significant part of economic growth comes from new enterprises and not the high growth of existing ones; this leads to a dilution of GDP growth before it reaches shareholders. Lastly, expected economic growth may be built into the prices and thus reduce future realized returns.

In their refined version, supply-side models tie a country's stock returns to its GDP growth, but they do not suggest a perfect match between the two variables. Instead, they view real GDP growth as a cap on long-run stock returns, as other factors dilute GDP before it reaches shareholders.

However, the empirical analysis of the presumed link between GDP and stock growth has certain limitations. Although we use a relatively long-term international equity data set, the analysis results are dependent on the start and end dates of the time series, since the economy and stocks follow cyclical patterns. Another issue concerns the role of investors' expectations. If expectation of future GDP growth is entirely built into today's valuations, stock price movements will tend to precede developments in the underlying economy. A deeper analysis is needed to test for a lag between the two time series.

References

William J. Bernstein and Robert D. Arnott, Earnings growth: the two percent dilution, Financial Analyst Journal, September/October 2003, 47 – 55

Bradford Cornell, Economic Growth and Equity Investing, Financial Analyst Journal, Volume 66, 2010

Jeffrey J. Diermeier, Roger G. Ibbotson and Laurence B. Siegel, The Supply of Capital Market Returns, Financial Analyst Journal, March/April 1984, 74 - 80

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

Jay R. Ritter, Economic growth and equity returns, Pasific-Basin Finance Journal 13 (2005) 489 – 503

Jeremy J. Siegel, Stocks for the Long Run, 1998, Second edition, McGraw-Hill

What Drives Long Term Equity Returns?, MSCI Barra Research Bulletin, January 2010

Contact Information

clientservice@mscibarra.com

Americas	
Americas	1.888.588.4567 (toll free)
Atlanta	+ 1.404.551.3212
Boston	+ 1.617.532.0920
Chicago	+ 1.312.675.0545
Montreal	+ 1.514.847.7506
Monterrey	+ 52.81.1253.4020
New York	+ 1.212.804.3901
San Francisco	+ 1.415.836.8800
Sao Paulo	+ 55.11.3706.1360
Stamford	+1.203.325.5630
Toronto	+ 1.416.628.1007

Europe, Middle East & Africa	
Amsterdam	+ 31.20.462.1382
Cape Town	+ 27.21.673.0100
Frankfurt	+ 49.69.133.859.00
Geneva	+ 41.22.817.9777
London	+ 44.20.7618.2222
Madrid	+ 34.91.700.7275
Milan	+ 39.02.5849.0415
Paris	0800.91.59.17 (toll free)
Zurich	+ 41.44.220.9300

Asia Pacific	
China North	10800.852.1032 (toll free)
China South	10800.152.1032 (toll free)
Hong Kong	+ 852.2844.9333
Seoul	+ 827.0768.88984
Singapore	800.852.3749 (toll free)
Sydney	+ 61.2.9033.9333
Tokyo	+ 81.3.5226.8222

www.mscibarra.com

Notice and Disclaimer

About MSCI Barra

MSCI Barra is a leading provider of investment decision support tools to investment institutions worldwide. MSCI Barra products include indices and portfolio risk and performance analytics for use in managing equity, fixed income and multi-asset class portfolios.

The company's flagship products are the MSCI International Equity Indices, which include over 120,000 indices calculated daily across more than 70 countries, and the Barra risk models and portfolio analytics, which cover 59 equity and 48 fixed income markets. MSCI Barra is headquartered in New York, with research and commercial offices around the world.

What Drives Long-Term Equity Returns? | January 2010

MSCI Barra RESEARCH BULLETIN

Introduction

In this Research Bulletin, we analyze long run returns of international equity markets using historical data spanning the 1975 - 2009 period. We decompose these returns into components and analyze their evolution over time.

This topic has been studied in the past. For example, Ibbotson and Chen (2003) provide a good overview of various decomposition methods and apply them to the US market. However, in our study we use a similar method and present the results using an international view.

Decomposition of the MSCI World Index

We decompose the equity total return (geometric average) into inflation, dividends, and real capital gain. The real capital gain is further broken down into real book value (r.BV) growth and growth in the price to book (PB) ratio. By using book value rather than earnings, we avoid periods with negative earnings where decomposition would not be meaningful. This method is summarized by the following formula:

$$TotalReturn = Inflation + g(PB) + g(r.BV) + DivIncome + Res$$

Residual interactions (*Res*) account for the geometric interaction between the various components when they are compounded over several periods. This term is small compared to the other four. For simplicity, this study ignores the effect of the exchange rates.

First, we decompose the MSCI World Index gross returns from the viewpoint of a US-based investor. The performance is expressed in US Dollars and we measure inflation by US domestic inflation. The results are presented in Exhibit 1.

Exhibit 1: Components of the MSCI World Index gross returns and their volatilities, 1975-2009 and subperiods

Period	1975 - 2009	1975 - 1979	1980 - 1989	1990 - 1999	2000 - 2009	volatility 1975 - 2009
Gross Index Return (USD)	11.1%	16.0%	19.9%	12.0%	-0.2%	14.9%
Inflation (USD)	4.2%	8.1%	5.1%	2.9%	2.6%	1.3%
Price to Book Growth	1.5%	2.3%	8.0%	5.0%	-8.3%	14.0%
Real Book Value Growth	2.1%	0.2%	2.1%	1.4%	3.8%	5.6%
Dividend Income	2.9%	4.6%	3.6%	2.1%	2.2%	0.4%
Residual Interactions	0.4%	0.7%	1.2%	0.5%	-0.5%	0.3%

Source: MSCI Barra and OECD (inflation data); annualized values. Data as of September 30, 2009.

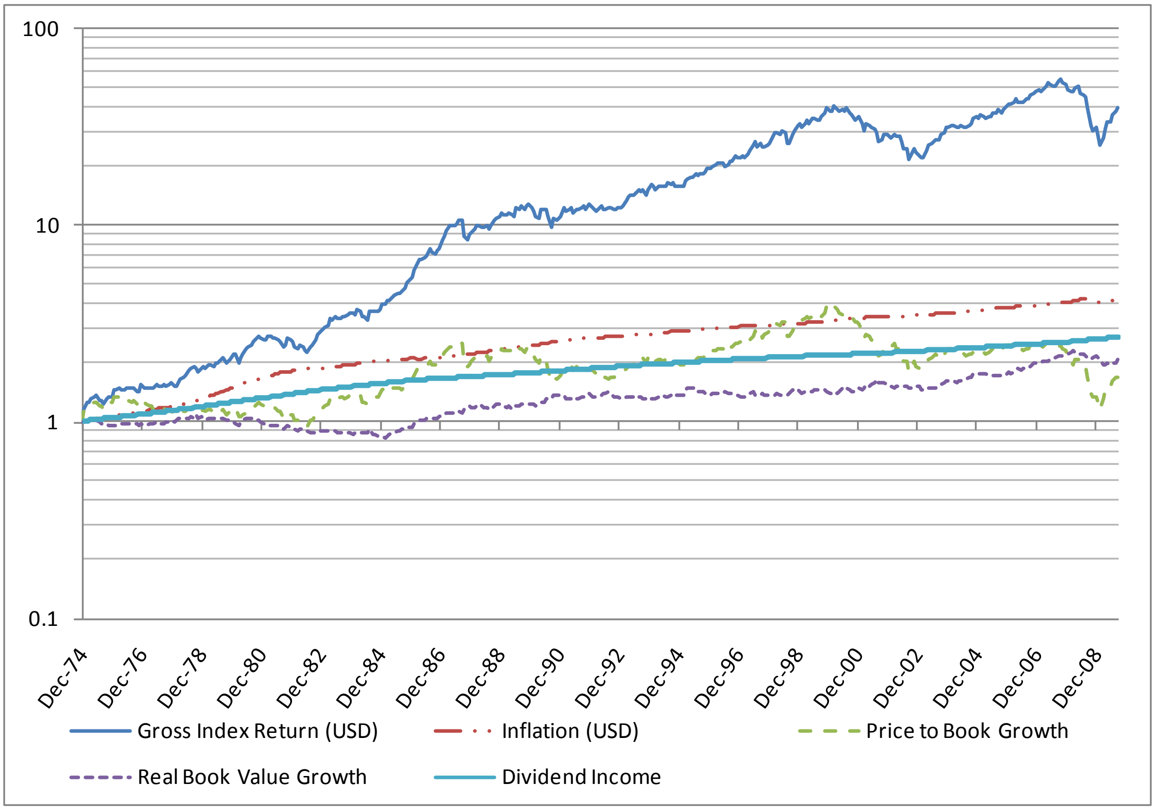
The MSCI World Index annualized gross index return for the total 35-year time span was 11.0%. The biggest component of this return was inflation at 4.2%, contributing more than one third of the total return. Other important components were dividend income (2.9%), emphasizing the importance of dividend reinvestment in long-term investing, and real book value growth (2.0%). Price to book growth contributed the least (1.5%).

When looking at the sub-period breakdown of the return components, interesting patterns emerge. Dividend income was on a downward trend, declining from 4.6% in the 1970s to 2.2% in the current decade. The relatively small effect of the valuation (PB) change in the long run hides a very volatile history: in the last three decades, it was the most important component of equity returns, expanding annually by 8% in the 80s, 5.0% in the 1990s and shrinking by 8.4% in the last decade.

This behavior can also be seen in Exhibit 2, which shows the cumulative contribution of the different return components over time. While inflation, dividend income, and book value present steady growth (barring a slight decline in real book value growth in the early 1980s), the price to book value component represents the source of volatility in the overall equity return.

This observation is also confirmed by the last column of Exhibit 1, where we see the annualized volatilities of the different return components for the complete period. Indeed, the volatility of the PB growth component is 14.0%, just slightly below the overall volatility of 14.9%.

Exhibit 2: Cumulative return of the components of the MSCI World Index (gross), 1975-2009



Source: MSCI Barra and OECD (inflation data). Data as of September 30, 2009.

Decomposition of regional returns

We now apply the same decomposition method to the gross returns of five regional and country indices, expressed in their home currency[1]: MSCI USA, MSCI Japan, MSCI Europe, MSCI Australia, and MSCI UK. The results are presented in Exhibit 3.

[1] Before the inception of Euro in 1999, we use DEM and German inflation for Europe.

Exhibit 3: Components of regional gross index returns and their volatilities, 1975-2009 and sub-periods

	Period	1975 - 2009	1975 - 1979	1980 - 1989	1990 - 1999	2000 - 2009	volatility 1975 - 2009
MSCI USA	Gross Index Return (USD)	**11.4%**	**13.3%**	**17.1%**	**19.0%**	**-1.9%**	**15.4%**
	Inflation (USD)	4.2%	8.1%	5.1%	2.9%	2.6%	1.3%
	Price to Book Growth	1.7%	0.7%	6.0%	10.4%	-9.9%	15.6%
	Real Book Value Growth	1.8%	-0.7%	0.6%	2.2%	4.2%	4.5%
	Dividend Income	3.2%	4.8%	4.6%	2.5%	1.8%	0.4%
	Residual Interactions	0.5%	0.4%	0.9%	1.0%	-0.6%	0.4%
MSCI Europe	Gross Index Return (EUR/DEM)	**10.7%**	**11.2%**	**18.3%**	**16.1%**	**-2.0%**	**16.6%**
	Inflation (EUR/DEM)	2.7%	4.1%	2.8%	2.6%	2.1%	1.0%
	Price to Book Growth	2.3%	3.2%	7.9%	8.2%	-9.2%	16.1%
	Real Book Value Growth	1.7%	-1.7%	2.3%	2.0%	2.6%	5.7%
	Dividend Income	3.6%	5.4%	4.2%	2.7%	3.0%	0.6%
	Residual Interactions	0.4%	0.3%	1.0%	0.8%	-0.5%	0.3%
MSCI Japan	Gross Index Return (JPY)	**5.2%**	**13.5%**	**22.3%**	**-4.0%**	**-4.7%**	**18.3%**
	Inflation (JPY)	1.8%	6.6%	2.3%	1.1%	-0.2%	1.9%
	Price to Book Growth	-0.8%	3.6%	9.7%	-6.6%	-6.9%	18.9%
	Real Book Value Growth	2.9%	0.4%	7.7%	0.9%	1.4%	5.2%
	Dividend Income	1.3%	2.4%	1.2%	0.8%	1.3%	0.4%
	Residual Interactions	0.1%	0.5%	1.4%	-0.2%	-0.2%	0.4%
MSCI Australia	Gross Index Return (AUD)	**14.3%**	**25.8%**	**17.8%**	**10.6%**	**9.1%**	**18.4%**
	Inflation (AUD)	5.5%	11.1%	8.3%	2.3%	3.2%	1.3%
	Price to Book Growth	2.7%	10.5%	1.0%	5.3%	-2.0%	19.6%
	Real Book Value Growth	1.2%	-2.6%	3.2%	-1.2%	3.7%	5.9%
	Dividend Income	4.3%	5.2%	4.4%	4.0%	4.1%	0.6%
	Residual Interactions	0.7%	1.6%	0.9%	0.3%	0.2%	0.8%
MSCI UK	Gross Index Return (GBP)	**15.4%**	**34.6%**	**23.2%**	**14.2%**	**0.8%**	**19.9%**
	Inflation (GBP)	5.4%	15.4%	6.5%	3.1%	1.9%	2.3%
	Price to Book Growth	4.2%	14.6%	8.2%	7.7%	-7.5%	20.4%
	Real Book Value Growth	0.8%	-3.9%	2.1%	-0.4%	3.4%	7.3%
	Dividend Income	4.1%	5.8%	4.8%	3.3%	3.5%	0.5%
	Residual Interactions	0.8%	2.6%	1.7%	0.5%	-0.4%	1.2%

Source: MSCI Barra, OECD (inflation). AUD inflation is based on Australian Bureau of Statistics data[2]. Data as of September 30, 2009.

We observe similar trends for the US and Europe: the first three periods saw high total returns whereas the last decade had a decline. Valuation ratios showed considerable growth in the 1980s and 1990s for both regions, and inflation was lower in Europe than in the US.

These dynamics were significantly different in Japan. First, during this 35-year period, the annualized performance of the MSCI Japan Index was approximately half that of the other two regions, even after accounting for inflation. Notably, the last two decades in Japan were marked by a continued underperformance, mainly due to the shrinking valuation ratios after the burst of the Japanese bubble. Second, dividend income was less than half of that in the other regions and was not the most important component of the total return after inflation.

Australia and the UK generally outperformed the other regions during the 1975-2009 period in local currency terms. This outperformance is mainly due to their higher inflation rates and dividend yield. The first five-year subperiod (1975-1979) saw exceptional gross returns in both countries (25.8% for the MSCI Australia Index and 34.8% for the MSCI UK Index) due to annual inflation and PB growth rates above 10%. It is also interesting to note that Australia had a positive

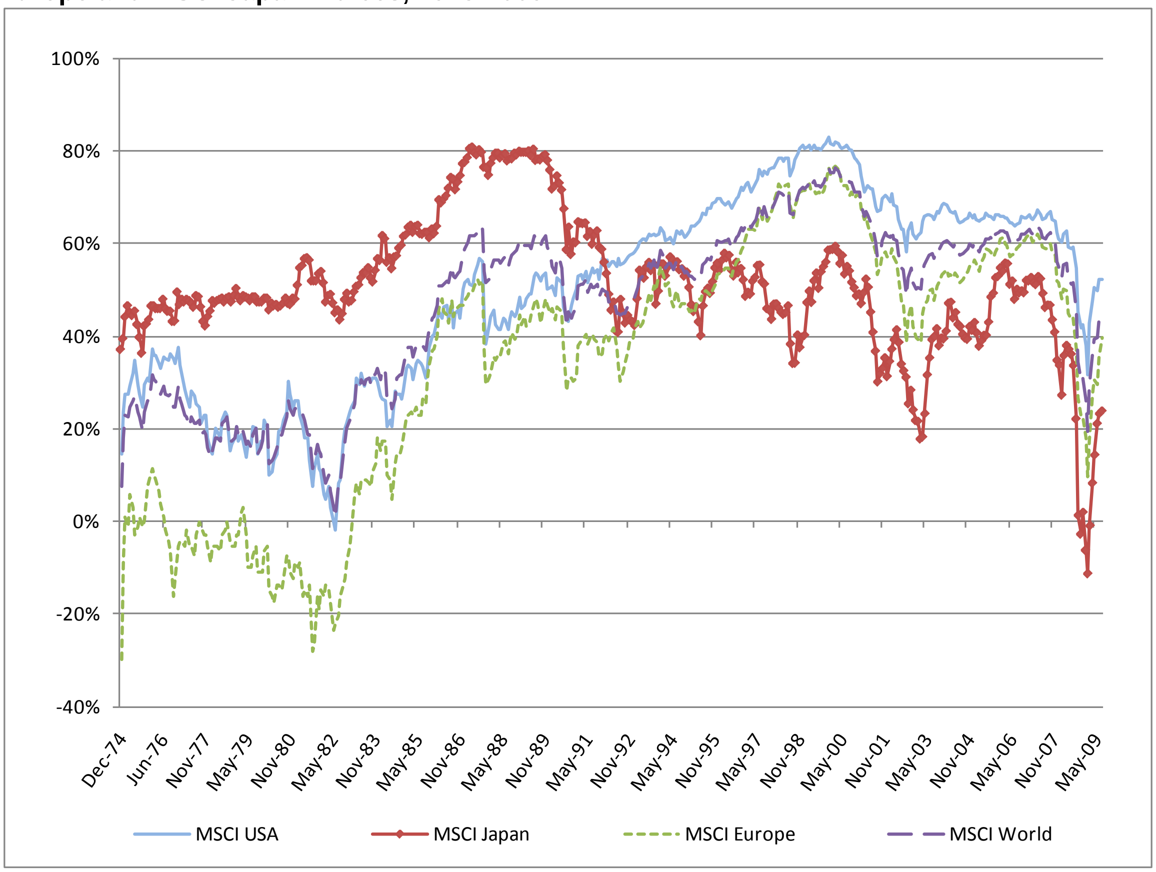
[2] ABS publishes quarterly CPI data. We used linear interpolation to generate monthly series. Note that this process also lowers the volatility of the inflation component.

annualized gross performance of 9.1% in the last decade, due to a relatively high dividend income and a relatively small decline in the PB ratio.

Decomposing price into book value and expectations of excess returns

Next, we take a closer look at the evolution of the price component of the regional indices. To do this, we decompose the price index level. We look at the book value per share, which we assume to be the liquidation value of the companies represented by the index. We also look at the difference between the price and the book value per share, which we attribute to expectations of future excess returns (returns above the return on equity— see Ohlson 1995 for the derivation of this result)[3]. Mathematically, the fraction of the book value component in the price is simply 1/PB, whereas the remaining fraction, 1-1/PB, represents the expectations of excess returns. Exhibit 4 shows the evolution of the latter for the MSCI World, MSCI USA, MSCI Europe and MSCI Japan price indices.

Exhibit 4: Fraction of expectations of excess returns in the MSCI World, MSCI USA, MSCI Europe and MSCI Japan Indices, 1975-2009



Source: MSCI Barra. Data as of September 30, 2009

We observe similar trends throughout the history for the MSCI World, MSCI USA, and to a lesser extent MSCI Europe Indices. From the mid 1970s, expectations of excess returns have been on an increasing trend. They stabilized in the 1980s at around 40-50%. Extreme events (for example, the dot-com bubble and the latest financial crisis) caused expectations of excess

[3] Note that one limitation of this analysis is its reliance on an accounting (as opposed to economic) measure to derive expectations of excess returns.

returns to drop to very low, even negative values, but these recovered to the pre-crisis levels relatively quickly.

These dynamics are again different in Japan. In Japan, expectations of excess returns started off at a higher level in the mid 1970s and reached a peak earlier than the other regions, at the top of the asset bubble of the 1980s. Afterwards, expectations were on a downward trend, and generally stayed below the levels of the other regions. After the dot-com bubble, Japan started to move in parallel with the other regions.

We can infer from this graph that over time, differences in expectations of excess returns have shrunk significantly among the different regions.

Conclusions

We decomposed long run returns of major equity markets into several components. The analysis showed that after inflation, dividend income was the most important part of equity returns for the majority of markets. Growth in real book value had a low, but steady contribution to performance. Changes in valuation tended to smooth out in the long run, but had important implications to equity investing in the short run.

We also analyzed how expectations of future excess returns – directly related to the price to book ratio - have evolved over time for different regions. After the continuing expansion in the 1980s and 1990s, these expectations have stabilized at historically high levels, quickly recovering from their lows in the 2009 due to the financial crisis. At the same time, differences in expectations of excess returns have shrunk significantly among the different regions.

References

Roger G. Ibbotson and Peng Chen (2002), Long-Run Stock Returns: Participating in the Real Economy, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=274150

James A. Olhson (1995), Earnings, Book values, and Dividends in Equity Valuation, Contemporary Accounting Research Vol. 11. No. 2. pp 661-687, http://www.lingnan.net/cferm/files/earnings,bookvalues,and%20dividends%20in%20equity%20valuation.PDF

Contact Information

clientservice@mscibarra.com

Americas

Americas	1.888.588.4567 (toll free)
Atlanta	+ 1.404.551.3212
Boston	+ 1.617.532.0920
Chicago	+ 1.312.675.0545
Montreal	+ 1.514.847.7506
New York	+ 1.212.804.3901
San Francisco	+ 1.415.576.2323
Sao Paulo	+ 55.11.3706.1360
Stamford	+1.203.325.5630
Toronto	+ 1.416.628.1007

Europe, Middle East & Africa

Amsterdam	+ 31.20.462.1382
Cape Town	+ 27.21.673.0110
Frankfurt	+ 49.69.133.859.00
Geneva	+ 41.22.817.9777
London	+ 44.20.7618.2222
Madrid	+ 34.91.700.7275
Milan	+ 39.02.5849.0415
Paris	0800.91.59.17 (toll free)
Zurich	+ 41.44.220.9300

Asia Pacific

China North	10800.852.1032 (toll free)
China South	10800.152.1032 (toll free)
Hong Kong	+ 852.2844.9333
Seoul	+ 822.2054.8538
Singapore	800.852.3749 (toll free)
Sydney	+ 61.2.9033.9333
Tokyo	+ 81.3.5226.8222

www.mscibarra.com

Notice and Disclaimer

About MSCI Barra

MSCI Barra is a leading provider of investment decision support tools to investment institutions worldwide. MSCI Barra products include indices and portfolio risk and performance analytics for use in managing equity, fixed income and multi-asset class portfolios.

The company's flagship products are the MSCI International Equity Indices, which include over 120,000 indices calculated daily across more than 70 countries, and the Barra risk models and portfolio analytics, which cover 56 equity and 46 fixed income markets. MSCI Barra is headquartered in New York, with research and commercial offices around the world.

Educational Articles

STOCKS

Equity Risk Premiums And Stocks Today

Marek Mscichowski | March 11, 2014

Stocks may appear to be at expensive levels. Looking at Price to Earnings (P/E) multiples of equities and comparing them to their historical averages, however, some commentators (namely, former Federal Reserve Chairman Alan Greenspan and NYU professor Aswath Damodaran) have recently pointed to equity risk premiums as another useful metric for valuing stocks. Unlike P/E multiples, equity premiums take interest rates, some currently at historically low levels historically, into account.

The equity premium is the total expected return (including capital growth and dividends) minus the risk-free rate. The total expected return is currently around 8.5%. The ten-year Treasury yield, an estimate of the risk-free rate, is about 3%. Hence, by our rough arithmetic, the equity premium that compensates investors for the added risk of holding corporate equity over theoretically risk-free U.S. government interest payments is currently about 5.5%.

Historically, the equity premium required by investors has averaged in the range of 3% to 7%. So this premium is about average, while interest rates, in some cases, are at historic lows.

The main reason that interest rates are so low is the Federal Reserve's massive asset-buyback program and abnormally low inflation. Through this lens, the elevated high P/E ratios make more sense, as investors search for returns in a low interest-rate environment. However, the Fed lowered the amount of monthly buybacks by $10 billion, from $85 billion to $75 billion, as 2013 came to a close. It then pared another $10 billion assets in January of this year. The Fed's efforts should eventually increase interest rates, though the timeframe appears to depend on the depth and breadth of an economic recovery. This has lent more urgency to speculation on Fed moves.

If interest rates go up and the required premium stays the same, this will decrease equity prices, all else being equal, as future cash flows are discounted by greater expected total returns. However, Professor Damodaran, who periodically posts his own equity risk premium estimate, argues that over the past decade, estimated returns have circled around the same mean, with equity risk premiums have largely compensated for falling interest rates, which have been in the hands of the Federal Reserve. Still, there are historical precedents for shifts in the total expected return because of either changes in the risk-free rate or equity premiums.

Besides interest rates and required equity premiums, another variable that can affect returns is earnings growth, which ultimately supplies money for returns in the form of dividends and buybacks. In recent years, corporations have been doing well, and the global economy seems to be firming up. Future earnings figures will also affect valuations. Damodaran provides a model (similar to a dividend discount model for a stock) for one to determine the intrinsic value of the S&P 500 Index by providing estimates for the risk-free rate, equity premium, as well as cash returns in the form of buybacks and their assumed growth rates.

What are some possible scenarios and how would they affect investors? Our previous discussion should shed some light. In the worst case scenario, interest rates will grow sharply, while the pace of earnings slow (compared to expectations, at least). This may mean equities are relatively overvalued now. For investors, the best case would be if earnings continue to grow nicely, while interest rates remain subdued. This may mean that the intrinsic value of equities is above the current price. With markets recently reaching all-time highs in some indexes and many stocks trading at premium P/E multiples compared to recent years, looking at the equity risk premium may provide investors with new insights into equity valuation and where stocks can go from here.

Value Line subscribers can compare our total return estimates with current bond yields for an idea of equity risk premium as they differ for each individual stock (In general, riskier stocks require higher premiums). Investors should also focus on our earnings and dividend estimates and projections, when considering if an investment is right for them on a fundamental basis.

At the time of this article's writing, the author did not have positions in any of the companies mentioned.

()