Advances in Business and Management Forecasting

Emerald Book Chapter: An Evaluation of Financial Analysts and Naïve Methods in Forecasting Long-Term Earnings

Michael Lacina, B. Brian Lee, Randall Zhaohui Xu

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:

$$\mathrm{ANEPS}_{t+\tau} = \mathrm{ANEPS}_{t+s} \times (1 + \mathrm{LTG})^{\tau - s}$$

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:

$$\mathrm{RW}_{t+\tau} = \mathrm{EPS}_{t-1}$$

where EPS_{t-1} is last year's actual EPS, and $\tau = 0, \ldots, 4$.

The RWGDP model is specified as follows:

$$\mathrm{RWGDP}_{t+\tau} = \mathrm{EPS}_{t-1}(1 + g)^{\tau+1}$$

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:

$$\mathrm{RWLTG}_{t+\tau} = \mathrm{EPS}_{t-1}(1 + \mathrm{LTG})^{\tau+1}$$

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:

$$\frac{(|\text{EPS}_{t+\tau} - \text{ANEPS}_{t+\tau}| - |\text{EPS}_{t+\tau} - \text{STATEPS}_{t+\tau}|)}{|\text{EPS}_{t+\tau}|}$$

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):

$$\frac{\text{EPS}_{t+\tau}}{\text{EPS}_{t-1}} - \frac{\text{STATEPS}_{t+\tau}}{\text{EPS}_{t-1}} = \alpha + \beta\left(\frac{\text{ANEPS}_{t+\tau}}{\text{EPS}_{t-1}} - \frac{\text{STATEPS}_{t+\tau}}{\text{EPS}_{t-1}}\right) + \varepsilon_{t+\tau} \quad (3)$$

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 $(|\mathrm{EPS}_{t+\tau}-\mathrm{ANEPS}_{t+\tau}$ (or $\mathrm{STATEPS}_{t+\tau})|)/P_{t-1}$, 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 $(|\mathrm{EPS}_{t+\tau}-\mathrm{ANEPS}_{t+\tau}$ (or $\mathrm{STATEPS}_{t+\tau})|)/|\mathrm{ANEPS}_{t+\tau}$ (or $\mathrm{STATEPS}_{t+\tau})|$, 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 $(|\mathrm{EPS}_{t+\tau}-\mathrm{ANEPS}_{t+\tau}|-|\mathrm{EPS}_{t+\tau}-\mathrm{STATEPS}_{t+\tau}|)/P_{t-1},$ 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 $(|\mathrm{EPS}_{t+\tau}-\mathrm{ANEPS}_{t+\tau}|-|\mathrm{EPS}_{t+\tau}-\mathrm{STATEPS}_{t+\tau}|)/|\mathrm{EPS}_{t+\tau}|$, 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:

$$\frac{\text{EPS}_{t+\tau}}{\text{EPS}_{t-1}} - \frac{\text{STATEPS}_{t+\tau}}{\text{EPS}_{t-1}} = \alpha + \beta\left(\frac{\text{ANEPS}_{t+\tau}}{\text{EPS}_{t-1}} - \frac{\text{STATEPS}_{t+\tau}}{\text{EPS}_{t-1}}\right) + \varepsilon_{t+\tau}$$

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 subsamples. 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:

$$\begin{aligned} SP_{i,t} &= R_{i,t} - \left(R_{f,t} + \beta_i (R_{m,t} - R_{f,t})\right), \text{ and} \\ SP_i &= \frac{1}{T} \sum_{t=1}^{T} SP_{i,t} \qquad i = 1, \ldots 10. \end{aligned} \tag{1}$$

[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

$$\begin{aligned} SP_{i,t}^{t+j} &= R_{i,t}^{t+j} - \left(R_{f,t} + \beta_i (R_{m,t} - R_{f,t})\right), \text{ and} \\ SP_{i}^{t+j} &= \frac{1}{T}\sum_{t=1}^{T} SP_{i,t}^{t+j}, \end{aligned} \tag{9}$$

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 over-estimate 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.

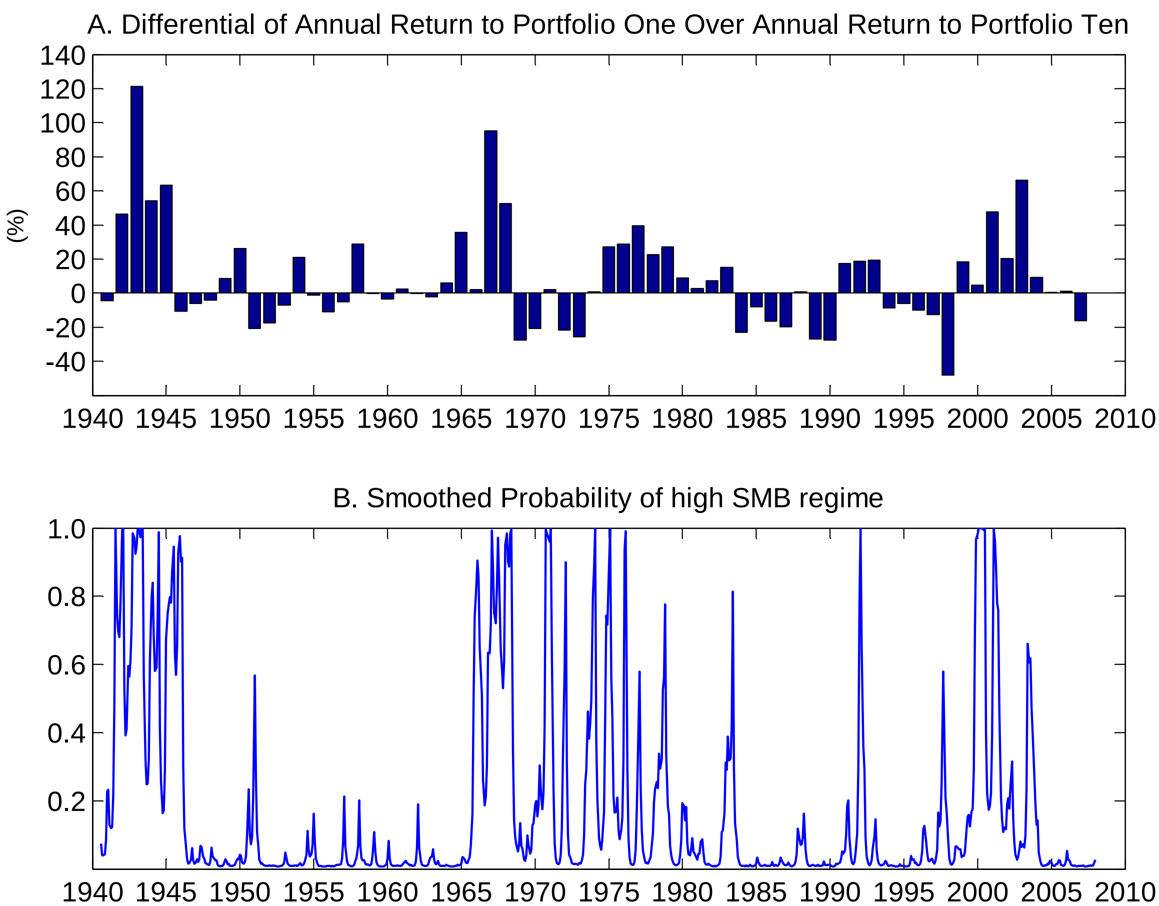


Figure 2: Three different dummy variables indicates three different economic environments. The first row includes the GDP growth rate of the U.S. and the business cycle dummy. The second row presents the CRSP monthly return and the market trend dummy variable derived from the smoothed probability of 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.

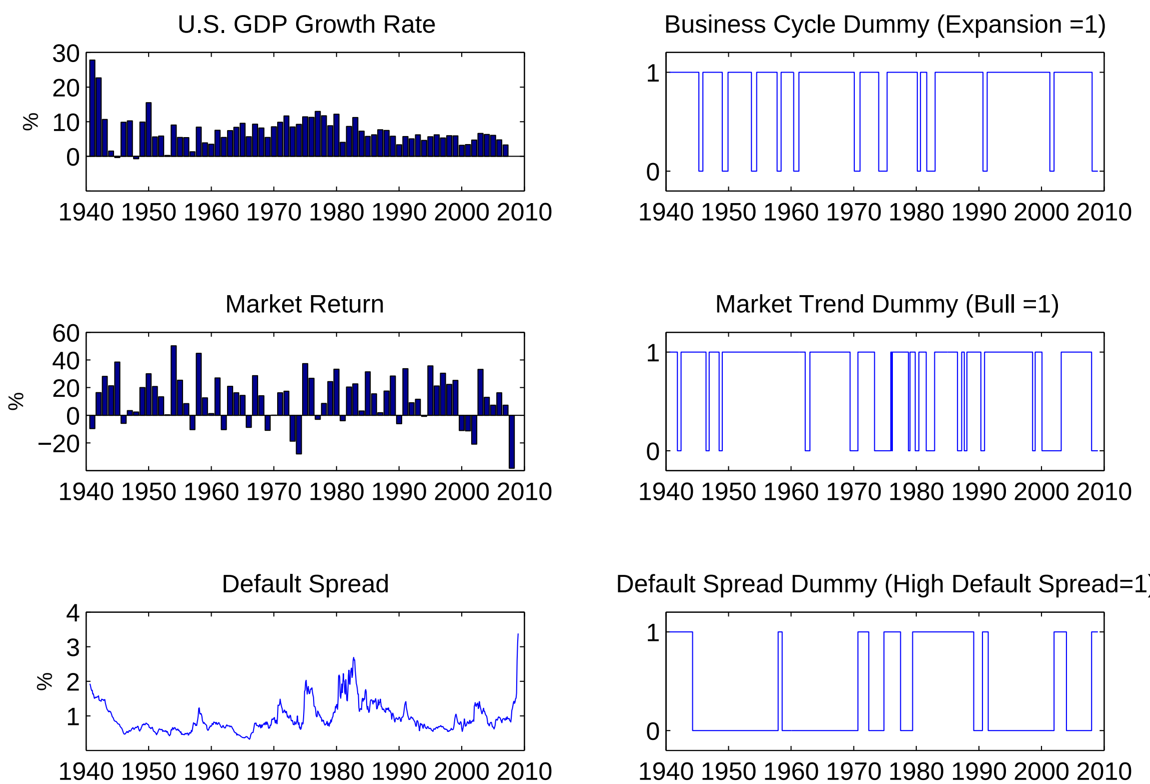


Table 1: **Returns on Size Portfolios and Size Premiums in Reference to CAPM**

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.

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."

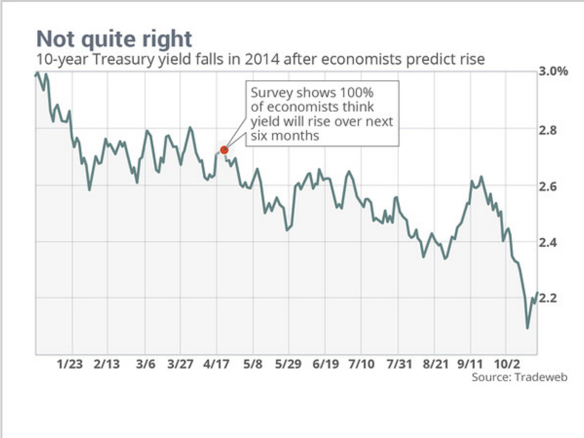
The April 22 report was based on a Bloomberg survey of 67 economists, all of whom expected the 10-year Treasury note 10_YEAR, +0.34% yield — which closed at 2.73% that day — to rise over the following half year.

"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.

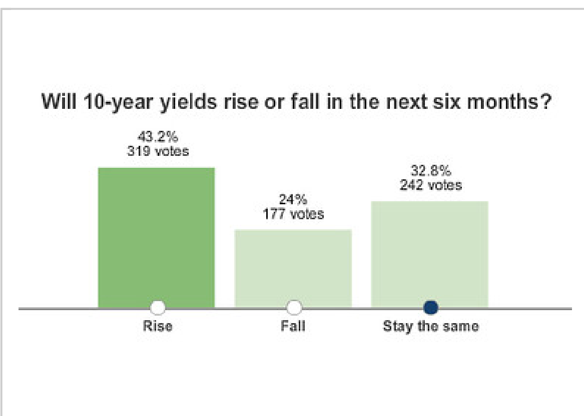
"There's an inherent bias out there that you can only get validation that the economy is improving if rates go up," said George Goncalves, head of interest-rate strategy at Nomura Securities. He was among the strategists saying in the spring that yields would keeping falling.

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

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

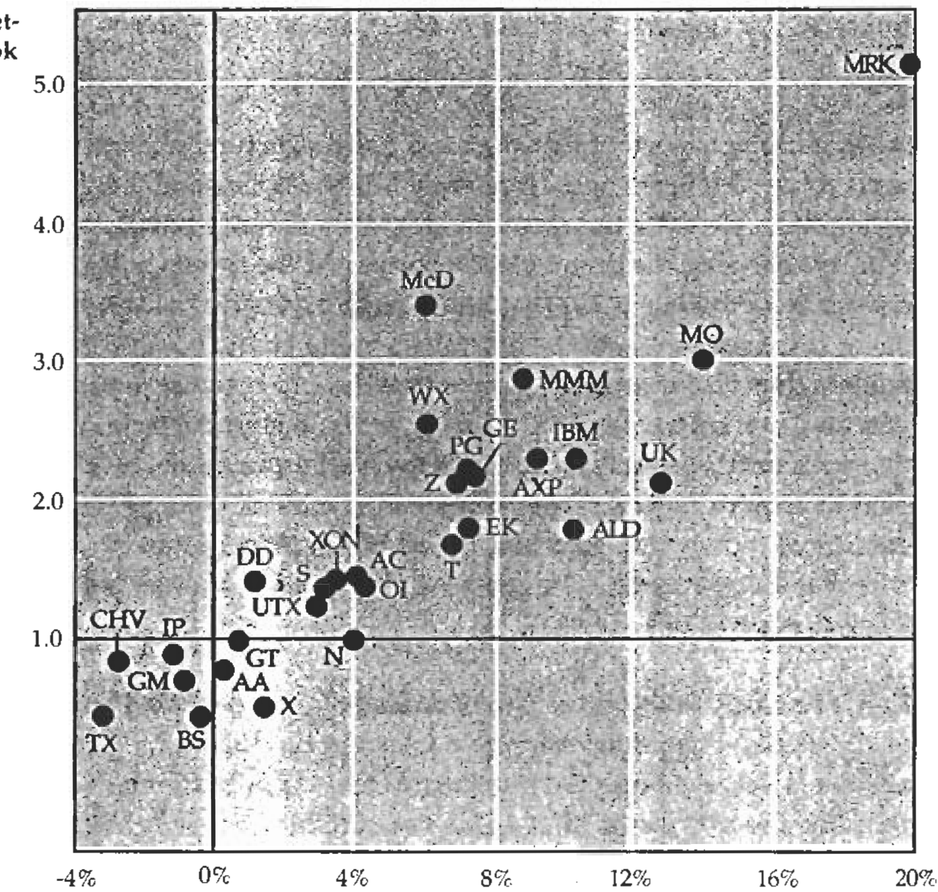
2) Poor balance sheet management, and

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

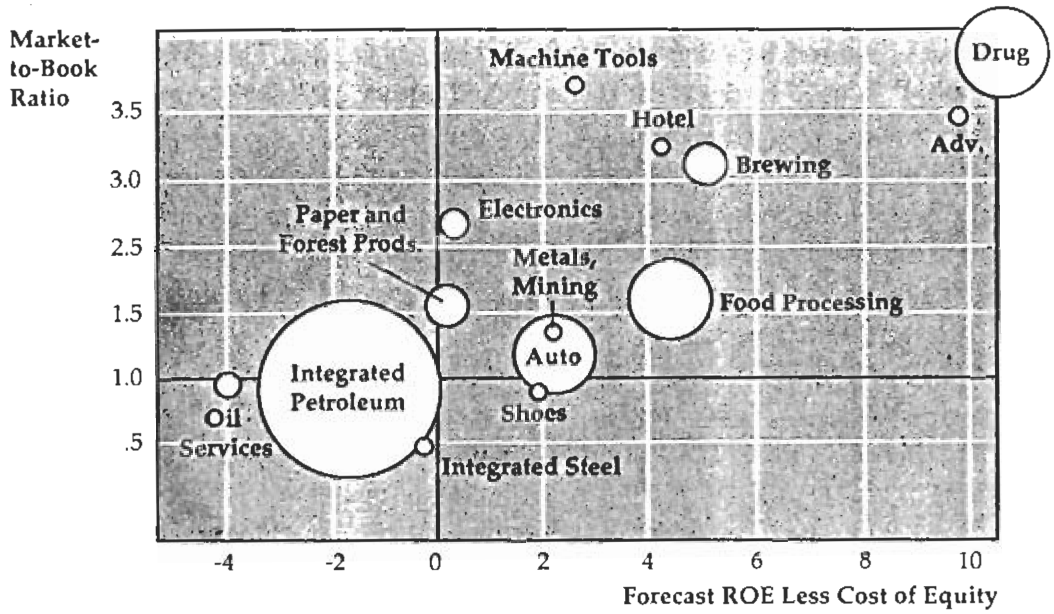
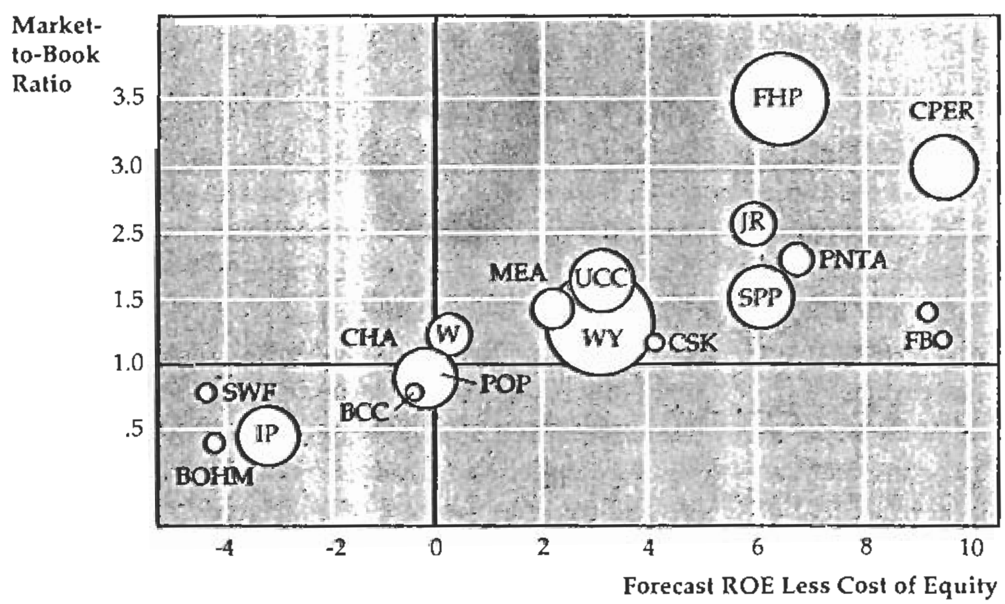


Exhibit III: Profitability of Paper and Forest Products Companies – Spring 1986



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

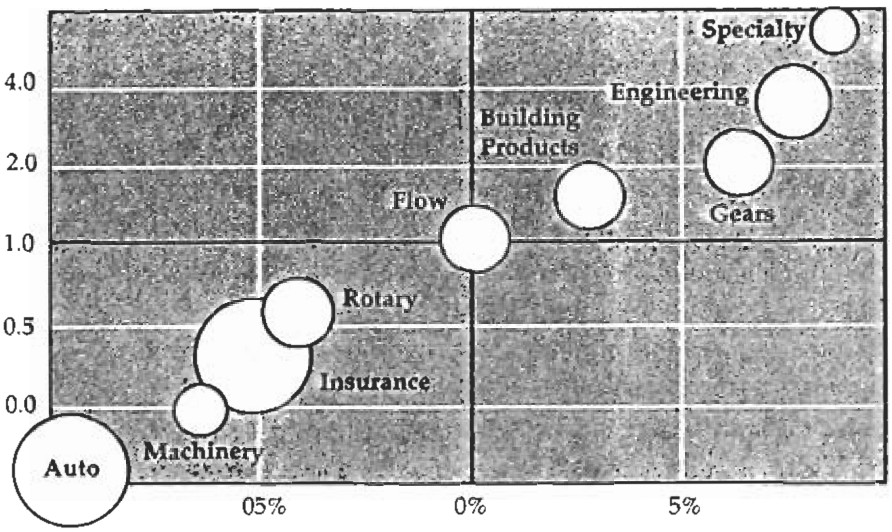
The wide variation in industry and company profitability also occurs within a typical diversified company's portfolio of businesses. Within a company, however, the capital allocation discipline provided by creditors and investors is replaced by management policies and strategies, which can significantly 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.

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	Standard deviation	% real return on a relatively riskless security: Mean	Standard deviation	% risk premium: Mean	Standard deviation	% real return on S&P 500: Mean	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).

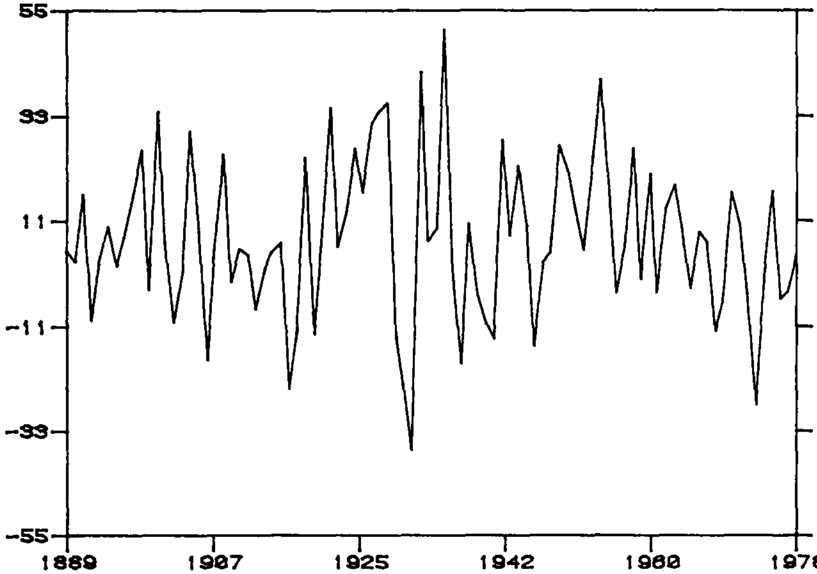


Fig. 1. Real annual return on S&P 500, 1889–1978 (percent).

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).

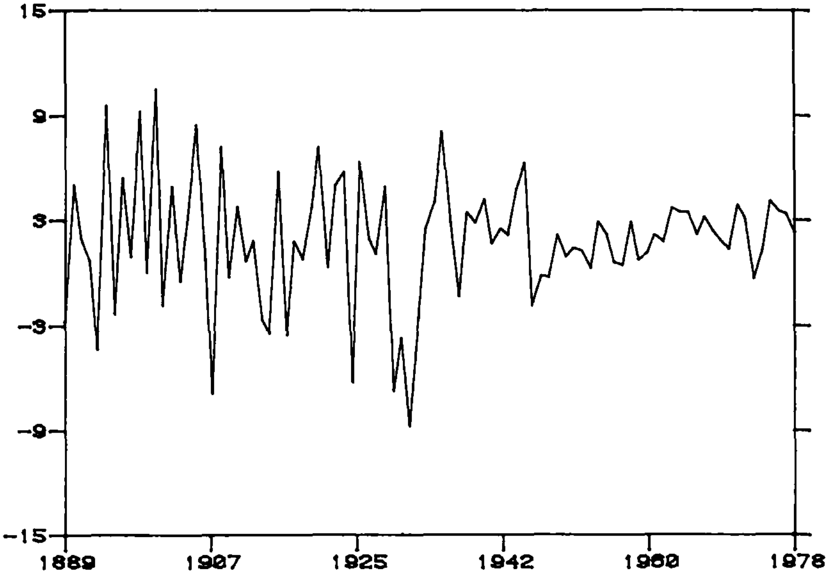


Fig. 2. Growth rate of real per capita consumption, 1889–1978 (percent).

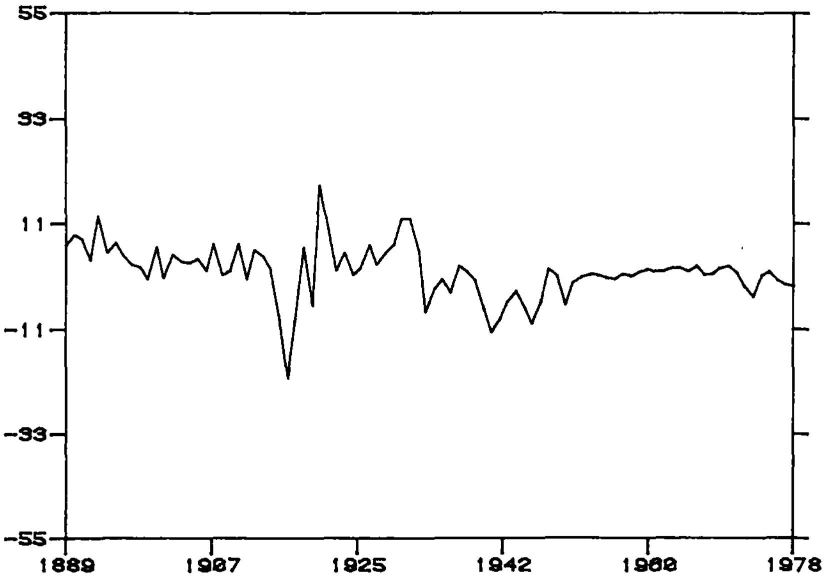


Fig. 3. Real annual return on a relatively riskless security, 1889–1978 (percent).

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 $x_{t+1} \in \{\lambda_1, \ldots, \lambda_n\}$ 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 $i, j = 1, \ldots, n$ 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 $y_s = y_t \cdot x_{t+1} \cdot \cdots \cdot x_s$, 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-

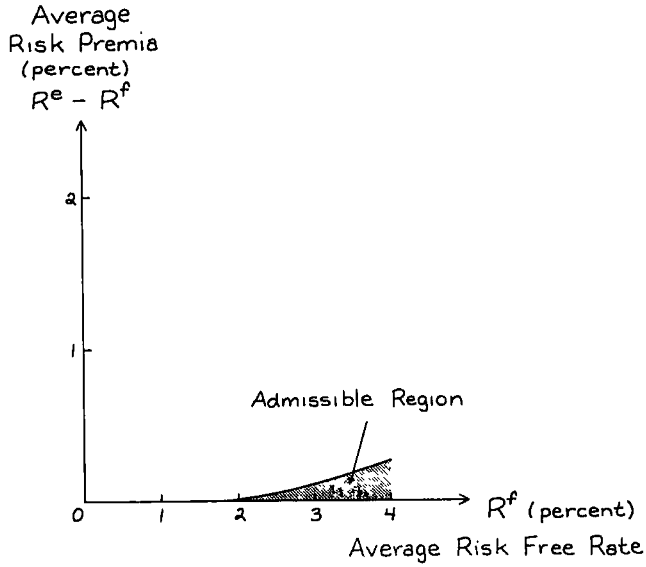


Fig. 4. Set of admissible average equity risk premia and real returns.

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 $\dot{R}^f_i$. 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.

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 AACCOUNT

*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.

MOODY'S
INVESTORS SERVICE

SECTOR IN-DEPTH
10 MARCH 2015

Rate this Research

ANALYST CONTACTS

Jim Hempstead 212-553-4318
Associate Managing Director
james.hempstead@moodys.com

Ryan Wobbrock 212-553-7104
AVP-Analyst
ryan.wobbrock@moodys.com

Jeffrey F. Cassella 212-553-1665
AVP-Analyst
jeffrey.cassella@moodys.com

Lesley Ritter 212-553-1607
Analyst
lesley.ritter@moodys.com

Jairo Chung 212-553-5123
Analyst
jairo.chung@moodys.com

Natividad Martel 212-553-4561
VP-Senior Analyst
natividad.martel@moodys.com

Susana Vivares 212-553-4694
VP-Senior Analyst
susana.vivares@moodys.com

Toby Shea 212-553-1779
VP-Senior Analyst
toby.shea@moodys.com

Swami Venkataraman, CFA 212-553-7950
VP-Sr Credit Officer
swami.venkat@moodys.com

US Regulated Utilities

Lower Authorized Equity Returns Will Not Hurt Near-Term Credit Profiles

The credit profiles of US regulated utilities will remain intact over the next few years despite our expectation that regulators will continue to trim the sector's profitability by lowering its authorized returns on equity (ROE). Persistently low interest rates and a comprehensive suite of cost recovery mechanisms ensure a low business risk profile for utilities, prompting regulators to scrutinise their profitability, which is defined as the ratio of net income to book equity. We view cash flow measures as a more important rating driver than authorized ROEs, and we note that regulators can lower authorized ROEs without hurting cash flow, for instance by targeting depreciation, or through special rate structures. Regulators can also adjust a utility's equity capitalization in its rate base. All else being equal, we think most utilities would prefer a thicker equity base and a lower authorized ROE over a small equity layer and a high authorized ROE.

» **More timely cost recovery helps offset falling ROEs**. Regulators continue to permit a robust suite of mechanisms that enable utilities to recoup prudently incurred operating costs, including capital investments such as environment related or infrastructure hardening expenditures. Strong cost recovery is credit positive because it ensures a stable financial profile. Despite lower authorized ROEs, we see the sector maintaining a ratio of Funds From Operations (FFO) to debt near 20%, a level that continues to support strong investment-grade ratings.

» **Utilities' cash flow is somewhat insulated from lower ROEs**. Net income represents about 30% - 40% of utilities' cash flow, so lower authorized returns won't necessarily affect cash flow or key financial credit ratios, especially when the denominator (equity) is rising. Regulators set the equity layer when capitalizing rate base, and the equity layer multiplied by the authorized ROE drives the annual revenue requirements. Across the sector, the ratio of equity to total assets has remained flat in the 30% range since 2007.

» **Utilities' actual financial performance remains stable**. Earned ROEs, which typically lag authorized ROEs, have not fallen as much as authorized returns in recent years. Since 2007, vertically integrated utilities, transmission and distribution only utilities, and natural gas local distribution companies have maintained steady earned ROE's in the 9% - 10% range. Holding companies with primarily regulated businesses also earned ROEs of around 9% - 10%, while returns for holding companies with diversified operations, namely unregulated generation, have fallen from 11% (over the past seven year average) to around 9% today.

Robust Suite of Cost Recovery Mechanisms Is Credit Positive

Over the past few years, the US regulatory environment has been very supportive of utilities. We think this is partly because regulators acknowledge that utility infrastructure needs a material amount of ongoing investment for maintenance, refurbishment and renovation. Utilities have also been able to garner support from both politicians and regulators for prudent investment in these critical assets because it helps create jobs, spurring economic growth. We also think regulators prefer to regulate financially healthy utilities.

Across the US, we continue to see regulators approving mechanisms that allow for more timely recovery of costs, a material credit positive. These mechanisms, which keep utilities' business risk profile low compared to most industrial corporate sectors, include: formulaic rate structures; special purpose trackers or riders; decoupling programs (which delink volumes from revenue); the use of future test years or other pre-approval arrangements. We also see a sustained increase in the frequency of rate case filings.

A supportive regulatory environment translates into a more transparent and stable financial profile, which in turn results in reasonably unfettered access to capital markets - for both debt and equity. Today, we think utilities enjoy an attractive set of market conditions that will remain in place over the next few years. By themselves, neither a slow (but steady) decline in authorized profitability, nor a material revision in equity market valuation multiples, will derail the stable credit profile of US regulated utilities.

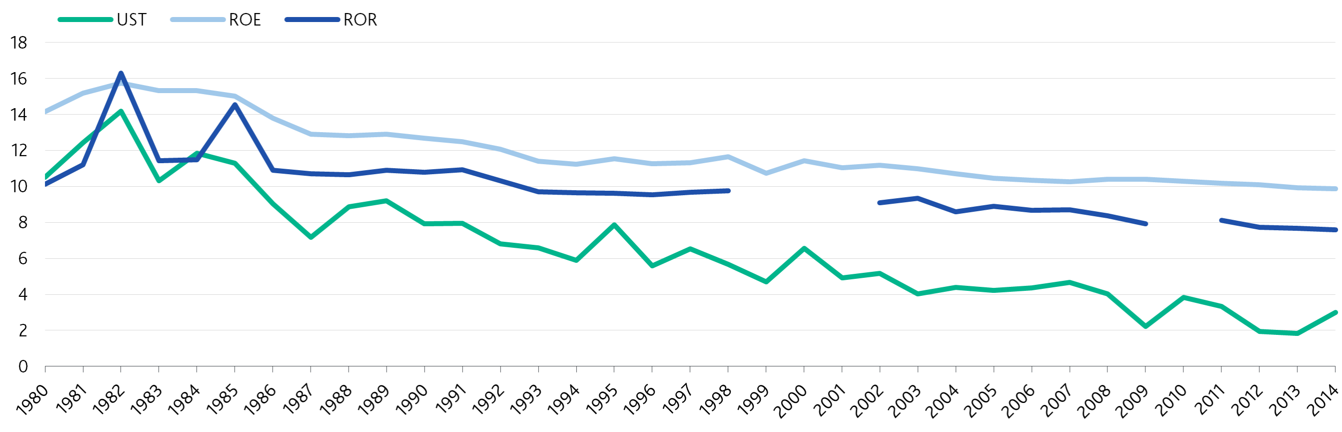
Cost recovery will help offset falling ROEs

Robust cost recovery mechanisms will help ensure that US regulated utilities' credit quality remains intact over the next few years. As a result, falling authorized ROEs are not a material credit driver at this time, but rather reflect regulators' struggle to justify the cost of capital gap between the industry's authorized ROEs and persistently low interest rates. We also see utilities struggling to defend this gap, while at the same time recovering the vast majority of their costs and investments through a variety of rate mechanisms.

In the table below, we show the US Treasury 10-year yield, which has steadily fallen from the 5% range in the summer of 2007 to the 2% range today. US utilities benefit from these lower interest rates because they borrow approximately $50 billion a year. For some utilities, a lower cost of debt translates directly into a higher return on equity, as long as their rate structure includes an embedded weighted average cost of capital (and the utilities can stay out of a general rate case proceeding).

Exhibit 1

Regulators hold up their end of the bargain by limiting reduction in return on equity (ROE) and overall rate of return (ROR) when compared with the decline in US Treasury 10-year yields



SOURCE: SNL Financial, LP, Moody's

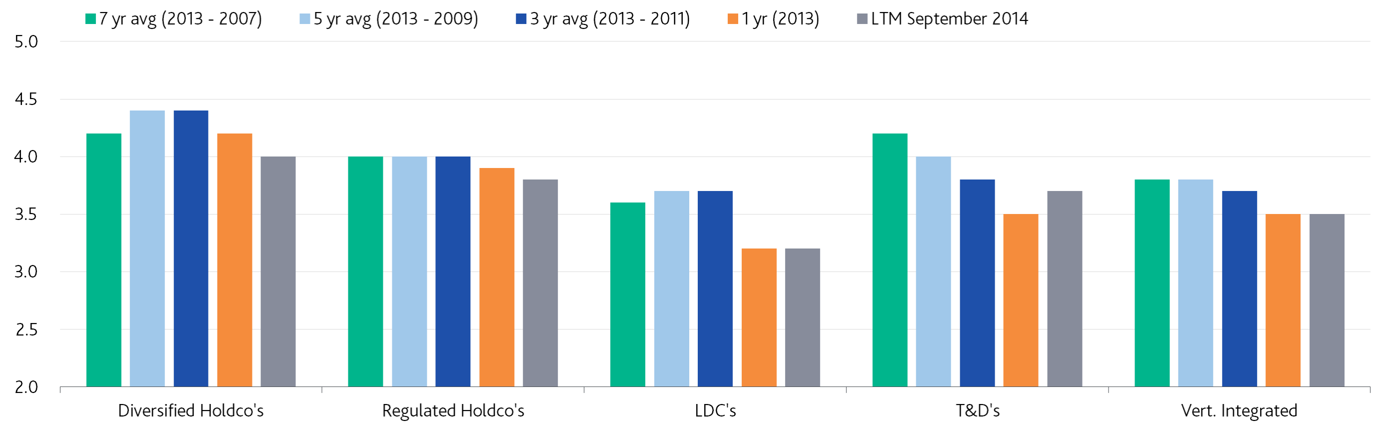
As utilities increasingly secure more up-front assurance for cost recovery in their rate proceedings, we think regulators will increasingly view the sector as less risky. The combination of low capital costs, high equity market valuation multiples (which are better than or on par with the broader market despite the regulated utilities' low risk profile), and a transparent assurance of cost recovery tend to support the case for lower authorized returns, although because utilities will argue they should rise, or at least stay unchanged.

One of the arguments for keeping authorized ROEs steady is that lowering them would make utilities less attractive to providers of capital. Utility holding companies assert that they would rather invest in higher risk-adjusted opportunities than in a regulated utility with sub-par return prospects. We see a risk that this argument could lead to a more contentious regulatory environment, a material credit negative. We do not think this scenario will develop over the next few years.

Our default and recovery data provides strong evidence that regulated utilities are indeed less risky (from the perspective of a probability of default and expected loss given default, as defined by Moody's) than their non-financial corporate peers. On a global basis, we nonetheless see a material amount of capital looking for regulated utility investment opportunities, and the same is true in the US despite, despite a lower authorized return. This is partly because investors can use holding company leverage to increase their actual equity returns, by borrowing capital at today's low interest rates and investing in the equity of a regulated utility.

Despite the reduction in authorized ROEs, US utilities are thankful to their regulators for the robust suite of timely cost recovery mechanisms which allow them to recoup prudently incurred operating costs such as fuel, as well as some investment expenses. These recovery mechanisms drive a stable and transparent dividend policy, which translates into historically very high equity multiples. Moreover, cost recovery helps keep the sector's overall financial profile stable, thereby supporting strong investment-grade ratings.

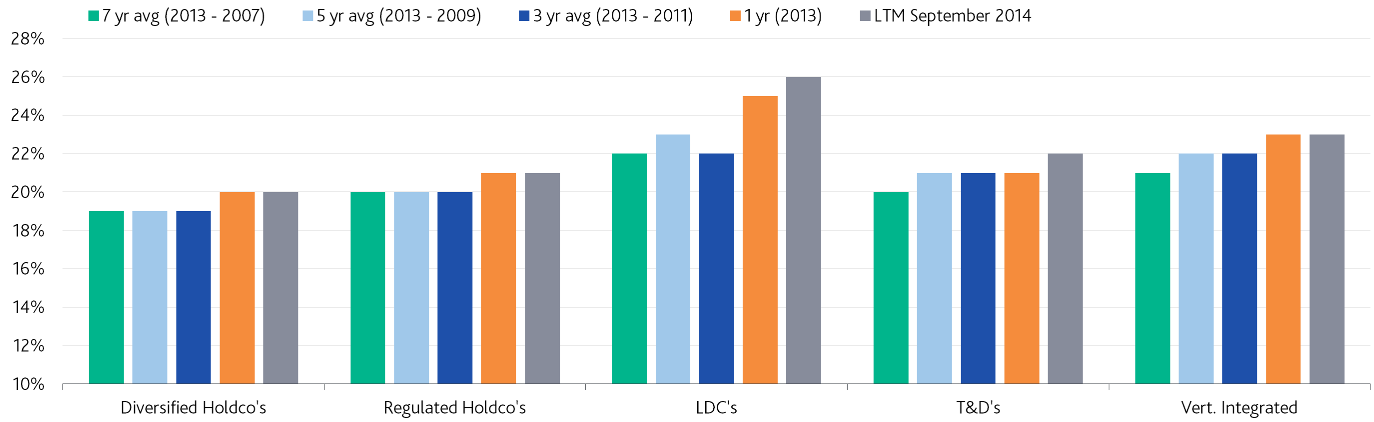
Exhibit 2
With better recovery mechanisms, the ratio of debt-to-EBITDA can rise, modestly, without negatively impacting credit profiles



SOURCE: Company filings; Moody's

Exhibit 3

The ratio of Funds From Operations to debt is rising, a material credit positive,
but the rise is partly funded by bonus depreciation and deferred taxes, which will eventually reverse



SOURCE: Company filings; Moody's

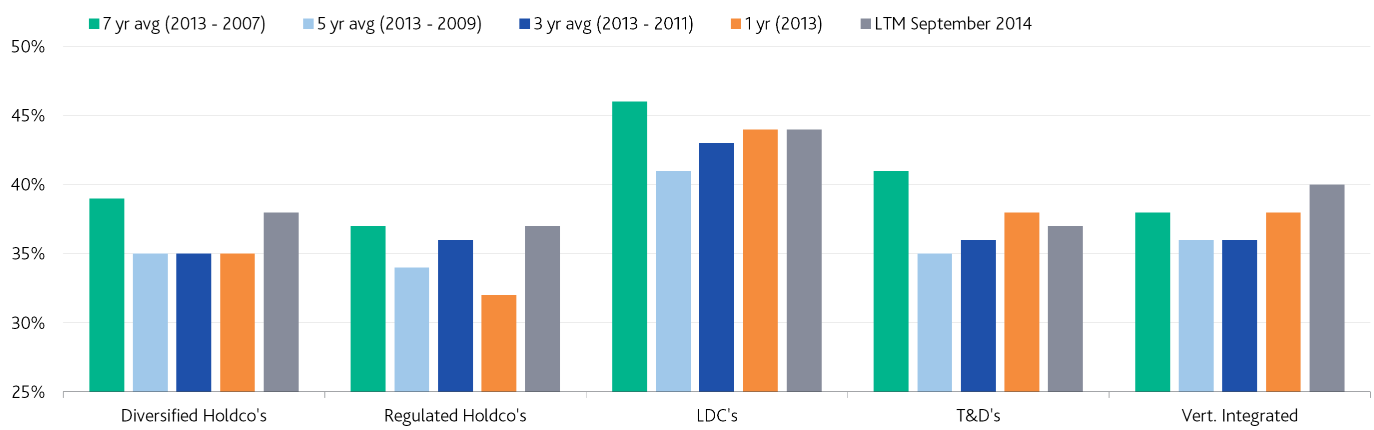
Utilities' cash flow is somewhat insulated from declining ROEs

Across all our utility group sub-sectors (see Appendix), net income - the numerator in the calculation of ROE – accounts for between 30% - 40% of cash flow. While net income is important, cash flow exerts a much greater influence over creditworthiness. This is primarily because cash flow takes into account depreciation and amortization expenses, along with other deferred tax adjustments. We note that deferred taxes have risen over the past few years, in part due to bonus depreciation elections, which will eventually reverse. From a credit perspective, there is a difference between the nominal amount of net income, which goes into cash flow, and the relationship of net income to book equity (a measure of profitability).

In the chart below, we highlight the ratio of net income to cash flow from operations (CFO) for our selected peer groups. Across all of the sectors, the longer term historical average of net income to CFO has fallen compared with the late 2000s, but has been rising over the more recent past. This is partly a function of deferred taxes, which have become a larger component of CFO over the past decade.

Exhibit 4

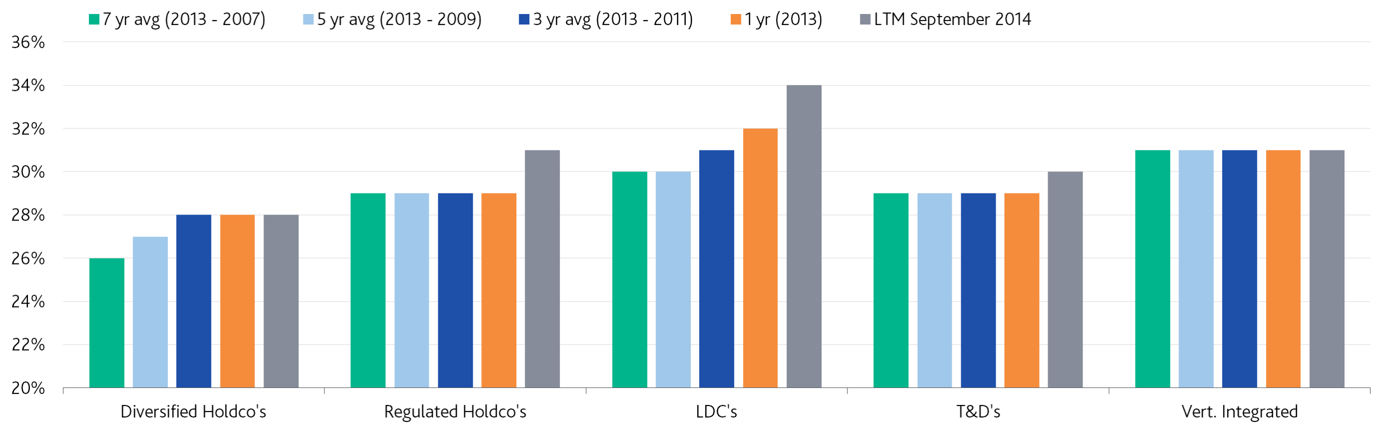
Net income as a % of cash flow from operations has been steadily rising (since 2011)



SOURCE: Company filings, Moody's

We can also envisage scenarios where regulators seek to achieve a reduction in authorized ROEs without harming credit profiles by focusing on utilities' equity layer. In the chart below, we illustrate median equity as a percentage of total assets for our selected peer groups. In our illustration, utilities will benefit from acquisition related goodwill on one hand, and impairments on the other.

Exhibit 5
Equity as a % of total assets, not capitalization, includes both goodwill and impairments



SOURCE: Company filings; Moody's

Utilities' actual financial performance remains stable

Earned ROE's, as reported by utilities and adjusted by Moody's, have been relatively flat over the past few years, despite the decline in authorized ROEs. This means utilities are closer to earning their authorized equity returns, which is positive from an equity market valuation perspective.

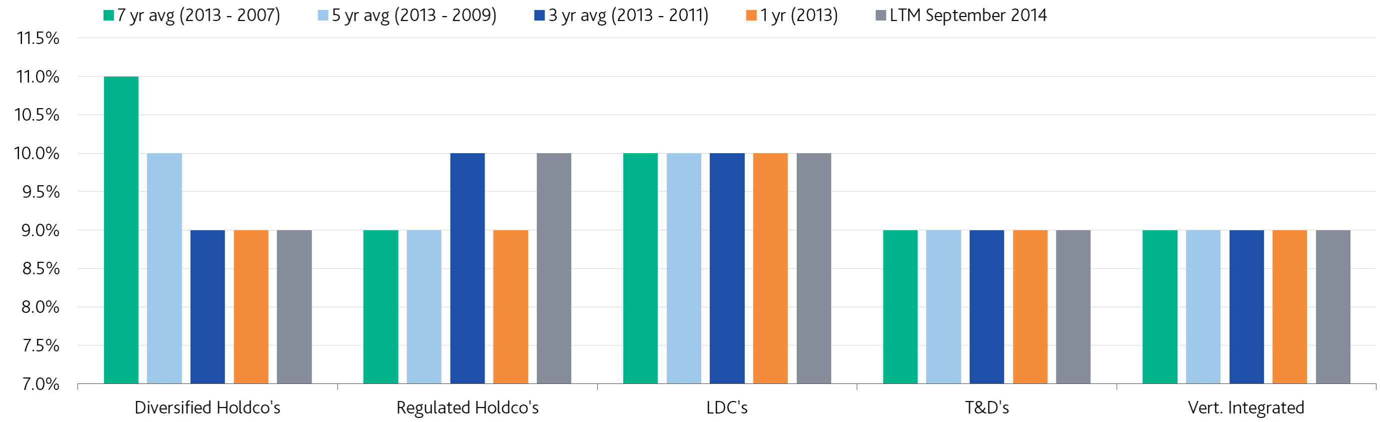
The authorized ROE is a popular focal point in many regulatory rate case proceedings. In addition, many regulatory jurisdictions look to established precedents that rely on various methodologies to determine an appropriate ROE, such as the capital asset pricing model or discounted cash flow analysis. In some jurisdictions where formulaic based rate structures point to lower ROEs for a longer projected period of time, regulators are incorporating a view that today's interest rate environment is "artificially" being held low.

Regardless, we think interest rates will go up, eventually. When they do, we also think authorized ROEs will trend up as well. However, just as authorized ROEs declined in a lagging fashion when compared to falling interest rates, we expect authorized ROEs to rise in a lagging fashion when interest rates rise.

Depending on alternative sources of risk-adjusted capital investment opportunities, this could spell trouble for utilities. For now, utilities can enjoy their (historically) high equity valuations, in terms of dividend yield and price-earnings ratios.

Exhibit 6

GAAP adjusted earned ROE's are relatively flat across all sub-sectors except Holding Companies with Diversified Operations, while the lower-risk LDC sector is outperforming



NOTE: GAAP adjusted ROE, not regulated ROE, does not adjust for goodwill or impairments.

Source: Company filings; Moody's

Appendix

Exhibit 7

Utilities with the highest earned ROEs (ranked by 7-year average)

Company Name	Sector	Rating	1-year average (2013) ROE	3-year average (2013 - 2011) ROE	5-year average (2013 - 2009) ROE	7-year average (2013 - 2007) ROE
CenterPoint Energy Houston Electric, LLC	T&D	A3	33%	32%	25%	23%
Questar Corporation	Holdco - Primarily Regulated	A2	14%	18%	20%	20%
AEP Texas Central Company	T&D	Baa1	14%	28%	22%	20%
Exelon Corporation	Holdco - Diversified	Baa2	7%	10%	14%	17%
CenterPoint Energy, Inc.	Holdco - Primarily Regulated	Baa1	7%	16%	15%	17%
Ohio Edison Company	T&D	Baa1	23%	18%	17%	16%
Public Service Enterprise Group	Holdco - Diversified	Baa2	11%	12%	14%	15%
Dayton Power & Light Company	T&D	Baa3	7%	9%	13%	15%
Dominion Resources Inc.	Holdco - Diversified	Baa2	13%	9%	12%	15%
Southern California Gas Company	LDC	A1	14%	13%	14%	15%
PECO Energy Company	T&D	A2	12%	12%	12%	14%
PPL Corporation	Holdco - Diversified	Baa3	9%	12%	11%	14%
UGI Utilities, Inc.	LDC	A2	15%	13%	13%	13%
Entergy Corporation	Holdco - Diversified	Baa3	7%	11%	12%	13%
Cleco Corporation	Holdco - Primarily Regulated	Baa1	10%	12%	13%	13%
Alabama Gas Corporation	LDC	A2	4%	11%	12%	13%
Entergy New Orleans, Inc.	Vertically Integrated Utility	Ba2	5%	10%	11%	12%
Entergy Gulf States Louisiana, LLC	Vertically Integrated Utility	Baa1	11%	13%	12%	12%
Piedmont Natural Gas Company, Inc.	LDC	A2	11%	11%	12%	12%
Ohio Power Company	T&D	Baa1	25%	14%	13%	12%
Southern Company (The)	Holdco - Primarily Regulated	Baa1	9%	11%	11%	12%
Georgia Power Company	Vertically Integrated Utility	A3	12%	12%	12%	12%
Alabama Power Company	Vertically Integrated Utility	A1	12%	12%	12%	12%
Southern California Edison Company	Vertically Integrated Utility	A2	8%	12%	12%	12%
NextEra Energy, Inc.	Holdco - Diversified	Baa1	10%	11%	11%	12%
Wisconsin Energy Corporation	Holdco - Primarily Regulated	A2	13%	13%	12%	12%
West Penn Power Company	T&D	Baa1	17%	13%	12%	12%
San Diego Gas & Electric Company	Vertically Integrated Utility	A1	9%	10%	11%	12%
Interstate Power and Light Company	Vertically Integrated Utility	A3	10%	9%	9%	12%

NOTE: GAAP adjusted ROE, not regulated ROE, does not adjust for goodwill or impairments.

SOURCE: Moody's; company filings

Exhibit 8

Highest (over 30%) and lowest (less than 20%) equity level as a % of total assets (ranked by 7-year average) [NOTE: Book equity is not adjusted for goodwill or impairments]

Company Name	Sector	Rating	1-year average (2013)	3-year average (2013 - 2011)	5-year average (2013 - 2009)	7-year average (2013 - 2007)
Duke Energy Ohio, Inc.	T&D	Baa1	48%	47%	48%	50%
Yankee Gas Services Company	LDC	Baa1	41%	42%	43%	43%
Texas-New Mexico Power Company	T&D	Baa1	43%	43%	43%	43%
Oncor Electric Delivery Company LLC	T&D	Baa1	40%	41%	41%	43%
Dayton Power & Light Company	T&D	Baa3	37%	38%	39%	40%
Pennsylvania Power Company	T&D	Baa1	25%	30%	34%	40%
Black Hills Power, Inc.	Vertically Integrated Utility	A3	38%	38%	37%	38%
ALLETE, Inc.	Vertically Integrated Utility	A3	38%	37%	37%	38%
Central Maine Power Company	T&D	A3	39%	38%	38%	38%
MGE Energy, Inc.	Holdco - Primarily Regulated	NR	39%	37%	38%	38%
Duke Energy Corporation	Holdco - Primarily Regulated	A3	36%	36%	37%	38%
Jersey Central Power & Light Company	T&D	Baa2	32%	33%	36%	38%
Oklahoma Gas & Electric Company	Vertically Integrated Utility	A1	36%	37%	37%	37%
Public Service Company of Colorado	Vertically Integrated Utility	A3	37%	37%	37%	37%
Virginia Electric and Power Company	Vertically Integrated Utility	A2	37%	37%	37%	35%
Wisconsin Public Service Corporation	Vertically Integrated Utility	A1	34%	34%	34%	35%
PacifiCorp	Vertically Integrated Utility	A3	36%	35%	35%	35%
UGI Utilities, Inc.	LDC	A2	35%	34%	34%	34%
Cleco Corporation	Holdco - Primarily Regulated	Baa1	37%	36%	34%	34%
Empire District Electric Company (The)	Vertically Integrated Utility	Baa1	35%	34%	34%	34%
Great Plains Energy Incorporated	Holdco - Primarily Regulated	Baa2	35%	35%	34%	34%
Nevada Power Company	Vertically Integrated Utility	Baa1	32%	33%	33%	33%
Tampa Electric Company	Vertically Integrated Utility	A2	34%	33%	33%	33%
Wisconsin Power and Light Company	Vertically Integrated Utility	A1	34%	33%	32%	33%
Questar Corporation	Holdco - Primarily Regulated	A2	29%	28%	31%	33%
Duke Energy Kentucky, Inc.	Vertically Integrated Utility	Baa1	31%	30%	33%	33%
Florida Power & Light Company	Vertically Integrated Utility	A1	36%	35%	34%	33%
Alabama Gas Corporation	LDC	A2	59%	40%	35%	33%
El Paso Electric Company	Vertically Integrated Utility	Baa1	34%	32%	32%	33%
IDACORP, Inc.	Holdco - Primarily Regulated	Baa1	34%	33%	33%	33%
PPL Electric Utilities Corporation	Vertically Integrated Utility	Baa1	34%	34%	34%	33%
Commonwealth Edison Company	T&D	Baa1	31%	32%	32%	33%
Georgia Power Company	Vertically Integrated Utility	A3	33%	33%	33%	33%
CMS Energy Corporation	Holdco - Primarily Regulated	Baa2	20%	19%	18%	18%
Hawaiian Electric Industries, Inc.	Holdco - Diversified		17%	16%	16%	16%
CenterPoint Energy, Inc.	Holdco - Primarily Regulated	Baa1	20%	19%	17%	15%
CenterPoint Energy Houston Electric, LLC	T&D	A3	9%	15%	15%	15%
AEP Texas Central Company	T&D	Baa1	13%	15%	14%	13%

SOURCE: Moody's; company filings

Exhibit 9

Highest (over 30%) and lowest (less than 15%) ratio of FFO to debt (ranked by 7-year average)

Company Name	Sector	Rating	1-year average (2013)	3-year average (2013 - 2011)	5-year average (2013 - 2009)	7-year average (2013 - 2007)
Dayton Power & Light Company	T&D	Baa3	32%	34%	42%	42%
Questar Corporation	Holdco - Primarily Regulated	A2	29%	30%	31%	42%
Pennsylvania Power Company	T&D	Baa1	30%	34%	32%	37%
Exelon Corporation	Holdco - Diversified	Baa2	28%	34%	37%	37%
Alabama Gas Corporation	LDC	A2	23%	27%	32%	36%
Florida Power & Light Company	Vertically Integrated Utility	A1	34%	35%	35%	35%
Southern California Gas Company	LDC	A1	42%	37%	35%	34%
Southern California Edison Company	Vertically Integrated Utility	A2	32%	33%	35%	32%
Madison Gas and Electric Company	Vertically Integrated Utility	A1	39%	35%	34%	31%
PECO Energy Company	T&D	A2	29%	31%	33%	31%
Dominion Resources Inc.	Holdco - Diversified	Baa2	16%	17%	16%	14%
Entergy Texas, Inc.	Vertically Integrated Utility	Baa3	15%	14%	12%	14%
Monongahela Power Company	T&D	Baa2	13%	16%	15%	14%
CMS Energy Corporation	Holdco - Primarily Regulated	Baa2	18%	16%	15%	14%
Appalachian Power Company	Vertically Integrated Utility	Baa1	15%	13%	14%	14%
Pennsylvania Electric Company	T&D	Baa2	15%	14%	12%	13%
NiSource Inc.	Holdco - Diversified	Baa2	15%	14%	14%	13%
Puget Energy, Inc.	Vertically Integrated Utility	Baa3	14%	12%	12%	13%
Toledo Edison Company	T&D	Baa3	10%	10%	8%	13%
Cleveland Electric Illuminating Company	T&D	Baa3	11%	11%	12%	13%
AEP Texas Central Company	T&D	Baa1	14%	15%	13%	12%

SOURCE: Moody's; company filings

Exhibit 10

Highest (over 4.5x) and lowest (less than 3.0x) ratio of debt to EBITDA (ranked by 1-year average, 2013, to focus on more recent performance)

Company Name	Sector	Rating	1-year average (2013)	3-year average (2013 - 2011)	5-year average (2013 - 2009)	7-year average (2013 - 2007)
Berkshire Hathaway Energy Company	Holdco - Diversified	A3	7.1	5.8	5.6	5.3
FirstEnergy Corp.	Holdco - Diversified	Baa3	6.0	5.2	4.8	4.4
Wisconsin Electric Power Company	Vertically Integrated Utility	A1	5.9	6.1	5.6	5.0
Entergy Texas, Inc.	Vertically Integrated Utility	Baa3	5.8	6.1	6.2	6.1
Monongahela Power Company	T&D	Baa2	5.6	5.2	5.7	6.0
NiSource Inc.	Holdco - Diversified	Baa2	5.2	5.5	5.4	5.5
PPL Corporation	Holdco - Diversified	Baa3	5.1	4.9	5.1	4.6
Appalachian Power Company	Vertically Integrated Utility	Baa1	5.0	5.0	5.2	5.4
Progress Energy, Inc.	Holdco - Primarily Regulated	Baa1	4.9	5.6	5.1	4.9
Puget Energy, Inc.	Vertically Integrated Utility	Baa3	4.9	5.6	5.9	5.6
Cleveland Electric Illuminating Company	T&D	Baa3	4.9	5.2	4.7	4.2
Northwest Natural Gas Company	LDC	A3	4.8	4.8	4.5	4.2
Jersey Central Power & Light Company	T&D	Baa2	4.7	5.5	4.2	3.6
NorthWestern Corporation	Vertically Integrated Utility	A3	4.7	4.5	4.4	4.3
Pepco Holdings, Inc.	Holdco - Primarily Regulated	Baa3	4.7	5.1	5.2	5.2
Laclede Gas Company	LDC	A3	4.7	5.5	5.3	5.6
Atlantic City Electric Company	T&D	Baa2	4.7	4.9	4.8	4.7
Nevada Power Company	Vertically Integrated Utility	Baa1	4.6	4.6	4.9	5.0
Black Hills Power, Inc.	Vertically Integrated Utility	A3	2.9	3.2	3.8	3.6
Virginia Electric and Power Company	Vertically Integrated Utility	A2	2.9	3.1	3.4	3.4
Duke Energy Kentucky, Inc.	Vertically Integrated Utility	Baa1	2.9	3.3	3.3	3.4
Texas-New Mexico Power Company	T&D	Baa1	2.9	2.9	3.2	3.3
Oklahoma Gas & Electric Company	Vertically Integrated Utility	A1	2.9	2.9	2.9	3.0
Cleco Power LLC	Vertically Integrated Utility	A3	2.9	3.2	3.6	3.7
Consumers Energy Company	Vertically Integrated Utility	A1	2.9	3.1	3.3	3.5
Alabama Power Company	Vertically Integrated Utility	A1	2.8	2.9	3.0	3.1
Public Service Electric and Gas Company	T&D	A2	2.8	3.0	3.2	3.3
Alabama Gas Corporation	LDC	A2	2.8	2.7	2.5	2.4
Pinnacle West Capital Corporation	Holdco - Primarily Regulated	Baa1	2.8	3.1	3.3	3.6
Cleco Corporation	Holdco - Primarily Regulated	Baa1	2.8	2.9	3.4	3.6
PECO Energy Company	T&D	A2	2.8	3.0	2.6	2.6
Northern States Power Company (Wisconsin)	Vertically Integrated Utility	A2	2.8	2.9	2.8	2.8
Duke Energy Carolinas, LLC	Vertically Integrated Utility	A1	2.8	3.1	3.2	3.1
UGI Utilities, Inc.	LDC	A2	2.7	3.0	3.1	3.3
Exelon Corporation	Holdco - Diversified	Baa2	2.7	2.8	2.5	2.5
West Penn Power Company	T&D	Baa1	2.7	3.3	3.3	3.4
Questar Corporation	Holdco - Primarily Regulated	A2	2.7	2.8	2.7	2.3
Tampa Electric Company	Vertically Integrated Utility	A2	2.6	2.7	2.8	2.9
Arizona Public Service Company	Vertically Integrated Utility	A3	2.6	2.9	3.1	3.3
New York State Electric and Gas Corporation	T&D	A3	2.6	2.9	3.2	4.3
Dayton Power & Light Company	T&D	Baa3	2.5	2.2	2.0	1.9
Florida Power & Light Company	Vertically Integrated Utility	A1	2.4	2.7	2.6	2.6
Ohio Power Company	T&D	Baa1	2.4	2.8	3.1	3.3
Madison Gas and Electric Company	Vertically Integrated Utility	A1	2.4	2.8	2.8	2.9
Pennsylvania Power Company	T&D	Baa1	2.4	2.3	2.4	2.2
MGE Energy, Inc.	Holdco - Primarily Regulated	NR	2.3	2.7	2.9	3.1
Rochester Gas & Electric Corporation	T&D	Baa1	2.3	2.9	3.0	3.5
Public Service Enterprise Group Incorporated	Holdco - Diversified	Baa2	2.3	2.3	2.3	2.4
NSTAR Electric Company	T&D	A2	2.2	2.6	2.7	2.8
Southern California Gas Company	LDC	A1	2.2	2.5	2.4	2.5
Mississippi Power Company	Vertically Integrated Utility	Baa1	(3.2)	3.5	3.4	3.1

Exhibit 11

List of Companies (NOTE: in our appendix tables, we exclude utilities with private ratings)

Company Name	Sector	Rating
Berkshire Hathaway Energy Company	Holdco - Diversified	A3
Black Hills Corporation	Holdco - Diversified	Baa1
Dominion Resources Inc.	Holdco - Diversified	Baa2
DTE Energy Company	Holdco - Diversified	A3
Entergy Corporation	Holdco - Diversified	Baa3
Exelon Corporation	Holdco - Diversified	Baa2
FirstEnergy Corp.	Holdco - Diversified	Baa3
Hawaiian Electric Industries, Inc.	Holdco - Diversified	NR
Integrys Energy Group, Inc.	Holdco - Diversified	A3
NextEra Energy, Inc.	Holdco - Diversified	Baa1
NiSource Inc.	Holdco - Diversified	Baa2
PPL Corporation	Holdco - Diversified	Baa3
Public Service Enterprise Group Incorporated	Holdco - Diversified	Baa2
Sempra Energy	Holdco - Diversified	Baa1
Alliant Energy Corporation	Holdco - Primarily Regulated	A3
Ameren Corporation	Holdco - Primarily Regulated	Baa2
American Electric Power Company, Inc.	Holdco - Primarily Regulated	Baa1
CenterPoint Energy, Inc.	Holdco - Primarily Regulated	Baa1
Cleco Corporation	Holdco - Primarily Regulated	Baa1
CMS Energy Corporation	Holdco - Primarily Regulated	Baa2
Consolidated Edison, Inc.	Holdco - Primarily Regulated	A3
Duke Energy Corporation	Holdco - Primarily Regulated	A3
Edison International	Holdco - Primarily Regulated	A3
Great Plains Energy Incorporated	Holdco - Primarily Regulated	Baa2
IDACORP, Inc.	Holdco - Primarily Regulated	Baa1
MGE Energy, Inc.	Holdco - Primarily Regulated	NR
Northeast Utilities	Holdco - Primarily Regulated	Baa1
Pepco Holdings, Inc.	Holdco - Primarily Regulated	Baa3
PG&E Corporation	Holdco - Primarily Regulated	Baa1
Pinnacle West Capital Corporation	Holdco - Primarily Regulated	Baa1
PNM Resources, Inc.	Holdco - Primarily Regulated	Baa3
Progress Energy, Inc.	Holdco - Primarily Regulated	Baa1
Questar Corporation	Holdco - Primarily Regulated	A2
SCANA Corporation	Holdco - Primarily Regulated	Baa3
Southern Company (The)	Holdco - Primarily Regulated	Baa1
Wisconsin Energy Corporation	Holdco - Primarily Regulated	A2
Xcel Energy Inc.	Holdco - Primarily Regulated	A3
Alabama Gas Corporation	LDC	A2
Atmos Energy Corporation	LDC	A2
DTE Gas Company	LDC	Aa3
Laclede Gas Company	LDC	A3
New Jersey Natural Gas Company	LDC	Aa2
Northern Natural Gas Company [Private]	LDC	A2
Northwest Natural Gas Company	LDC	A3
Piedmont Natural Gas Company, Inc.	LDC	A2
South Jersey Gas Company	LDC	A2
Southern California Gas Company	LDC	A1
Southwest Gas Corporation	LDC	A3
UGI Utilities, Inc.	LDC	A2
Washington Gas Light Company	LDC	A1
Wisconsin Gas LLC [Private]	LDC	A1
Yankee Gas Services Company	LDC	Baa1
AEP Texas Central Company	T&D	Baa1
AEP Texas North Company	T&D	Baa1
Atlantic City Electric Company	T&D	Baa2

Baltimore Gas and Electric Company	T&D	A3
CenterPoint Energy Houston Electric, LLC	T&D	A3
Central Hudson Gas & Electric Corporation	T&D	A2
Central Maine Power Company	T&D	A3
Cleveland Electric Illuminating Company (The)	T&D	Baa3
Commonwealth Edison Company	T&D	Baa1
Connecticut Light and Power Company	T&D	Baa1
Consolidated Edison Company of New York, Inc.	T&D	A2
Dayton Power & Light Company	T&D	Baa3
Delmarva Power & Light Company	T&D	Baa1
Duke Energy Ohio, Inc.	T&D	Baa1
Jersey Central Power & Light Company	T&D	Baa2
Metropolitan Edison Company	T&D	Baa1
Monongahela Power Company	T&D	Baa2
New York State Electric and Gas Corporation	T&D	A3
NSTAR Electric Company	T&D	A2
Ohio Edison Company	T&D	Baa1
Ohio Power Company	T&D	Baa1
Oncor Electric Delivery Company LLC	T&D	Baa1
Orange and Rockland Utilities, Inc.	T&D	A3
PECO Energy Company	T&D	A2
Pennsylvania Electric Company	T&D	Baa2
Pennsylvania Power Company	T&D	Baa1
Potomac Edison Company (The)	T&D	Baa2
Potomac Electric Power Company	T&D	Baa1
Public Service Electric and Gas Company	T&D	A2
Rochester Gas & Electric Corporation	T&D	Baa1
Texas-New Mexico Power Company	T&D	Baa1
Toledo Edison Company	T&D	Baa3
West Penn Power Company	T&D	Baa1
Western Massachusetts Electric Company	T&D	A3
Alabama Power Company	Vertically Integrated Utility	A1
ALLETE, Inc.	Vertically Integrated Utility	A3
Appalachian Power Company	Vertically Integrated Utility	Baa1
Arizona Public Service Company	Vertically Integrated Utility	A3
Avista Corp.	Vertically Integrated Utility	Baa1
Black Hills Power, Inc.	Vertically Integrated Utility	A3
Cleco Power LLC	Vertically Integrated Utility	A3
Consumers Energy Company	Vertically Integrated Utility	A1
DTE Electric Company	Vertically Integrated Utility	A2
Duke Energy Carolinas, LLC	Vertically Integrated Utility	A1
Duke Energy Florida, Inc.	Vertically Integrated Utility	A3
Duke Energy Kentucky, Inc.	Vertically Integrated Utility	Baa1
Duke Energy Progress, Inc.	Vertically Integrated Utility	A1
El Paso Electric Company	Vertically Integrated Utility	Baa1
Empire District Electric Company (The)	Vertically Integrated Utility	Baa1
Entergy Arkansas, Inc.	Vertically Integrated Utility	Baa2
Entergy Gulf States Louisiana, LLC	Vertically Integrated Utility	Baa1
Entergy Louisiana, LLC	Vertically Integrated Utility	Baa1
Entergy Mississippi, Inc.	Vertically Integrated Utility	Baa2
Entergy New Orleans, Inc.	Vertically Integrated Utility	Ba2
Entergy Texas, Inc.	Vertically Integrated Utility	Baa3
Florida Power & Light Company	Vertically Integrated Utility	A1
Georgia Power Company	Vertically Integrated Utility	A3
Gulf Power Company	Vertically Integrated Utility	A2
Hawaiian Electric Company, Inc.	Vertically Integrated Utility	Baa1
Idaho Power Company	Vertically Integrated Utility	A3
Indiana Michigan Power Company	Vertically Integrated Utility	Baa1
Interstate Power and Light Company	Vertically Integrated Utility	A3
Kansas City Power & Light Company	Vertically Integrated Utility	Baa1
Kentucky Power Company	Vertically Integrated Utility	Baa2

Madison Gas and Electric Company	Vertically Integrated Utility	A1
MidAmerican Energy Company	Vertically Integrated Utility	A1
Mississippi Power Company	Vertically Integrated Utility	Baa1
Nevada Power Company	Vertically Integrated Utility	Baa1
Northern States Power Company (Minnesota)	Vertically Integrated Utility	A2
Northern States Power Company (Wisconsin)	Vertically Integrated Utility	A2
NorthWestern Corporation	Vertically Integrated Utility	A3
Oklahoma Gas & Electric Company	Vertically Integrated Utility	A1
Pacific Gas & Electric Company	Vertically Integrated Utility	A3
PacifiCorp	Vertically Integrated Utility	A3
Portland General Electric Company	Vertically Integrated Utility	A3
PPL Electric Utilities Corporation	Vertically Integrated Utility	Baa1
Public Service Company of Colorado	Vertically Integrated Utility	A3
Public Service Company of New Hampshire	Vertically Integrated Utility	Baa1
Public Service Company of New Mexico	Vertically Integrated Utility	Baa2
Public Service Company of Oklahoma	Vertically Integrated Utility	A3
Puget Energy, Inc.	Vertically Integrated Utility	Baa3
Puget Sound Energy, Inc.	Vertically Integrated Utility	Baa1
San Diego Gas & Electric Company	Vertically Integrated Utility	A1
Sierra Pacific Power Company	Vertically Integrated Utility	Baa1
South Carolina Electric & Gas Company	Vertically Integrated Utility	Baa2
Southern California Edison Company	Vertically Integrated Utility	A2
Southwestern Electric Power Company	Vertically Integrated Utility	Baa2
Southwestern Public Service Company	Vertically Integrated Utility	Baa1
Tampa Electric Company	Vertically Integrated Utility	A2
Tucson Electric Power Company	Vertically Integrated Utility	Baa1
Union Electric Company	Vertically Integrated Utility	Baa1
Virginia Electric and Power Company	Vertically Integrated Utility	A2
Wisconsin Electric Power Company	Vertically Integrated Utility	A1
Wisconsin Power and Light Company	Vertically Integrated Utility	A1
Wisconsin Public Service Corporation	Vertically Integrated Utility	A1

ANALYST CONTACTS

NEW YORK	212-553-1653
Mihoko Manabe *Senior Vice President* mihoko.manabe@moodys.com	212-553-1942
Michael Haggarty *Associate Managing Director* michael.haggerty@moodys.com	212-553-7172
William L. Hess *Managing Director - Utilities* william.hess@moodys.com	212-553-3837

MOODY'S
INVESTORS SERVICE

SURVEY OF
PROFESSIONAL FORECASTERS

Release Date: February 10, 2017

FIRST QUARTER 2017

Brighter Outlook for Growth and Labor Markets over the Next Three Years

The U.S. economy looks stronger now than it did three months ago, according to 42 forecasters surveyed by the Federal Reserve Bank of Philadelphia. The forecasters predict real GDP will grow at an annual rate of 2.2 percent this quarter and 2.3 percent next quarter. On an annual-average over annual-average basis, the forecasters predict real GDP growing 2.3 percent in 2017, 2.4 percent in 2018, and 2.6 percent in 2019. The forecasts for 2017, 2018, and 2019 are higher than the estimates of three months ago. For 2020, real GDP is estimated to grow 2.1 percent.

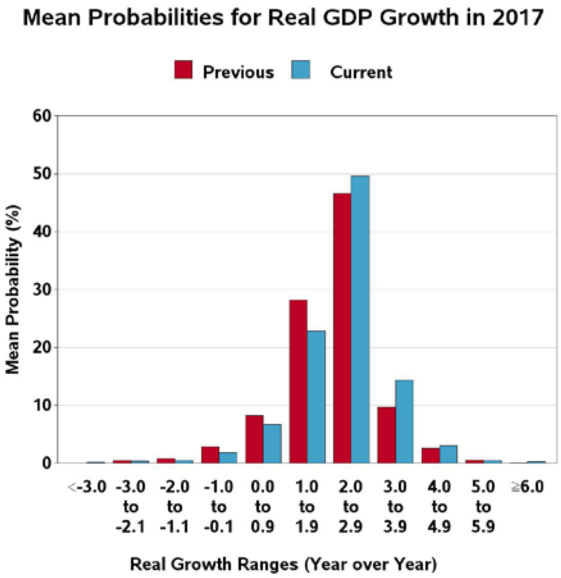
A brighter outlook for the labor market accompanies the outlook for stronger output growth. The forecasters predict that the unemployment rate will average 4.6 percent in 2017, 4.5 percent in 2018 and 2019, and 4.6 percent in 2020. The projections for 2017, 2018, and 2019 are below those of the last survey, indicating a brighter outlook for unemployment.

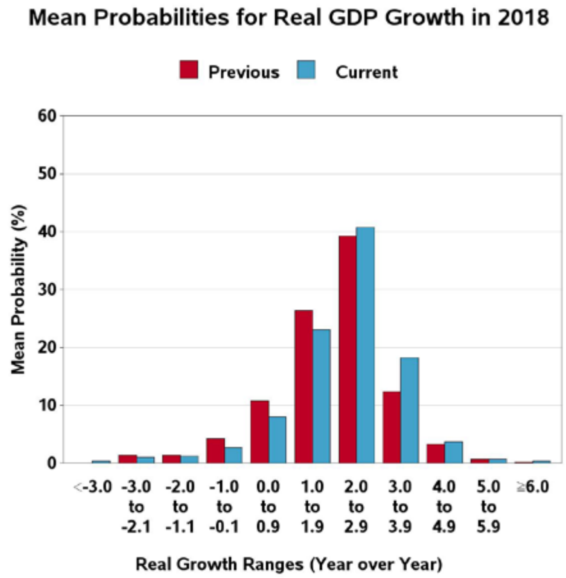
The panelists also predict an improvement in the employment outlook for 2017. The forecasters' projections for the annual-average level of nonfarm payroll employment suggest job gains at a monthly rate of 180,300 in 2017, up from the previous estimate of 173,600. (These annual-average estimates are computed as the year-to-year change in the annual-average level of nonfarm payroll employment, converted to a monthly rate.)

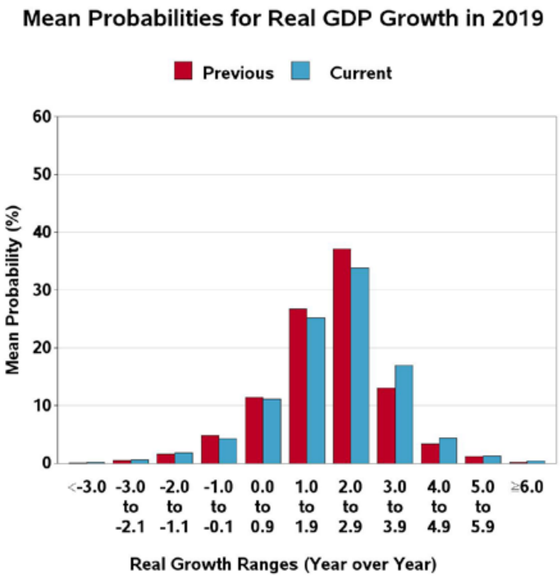
Median Forecasts for Selected Variables in the Current and Previous Surveys

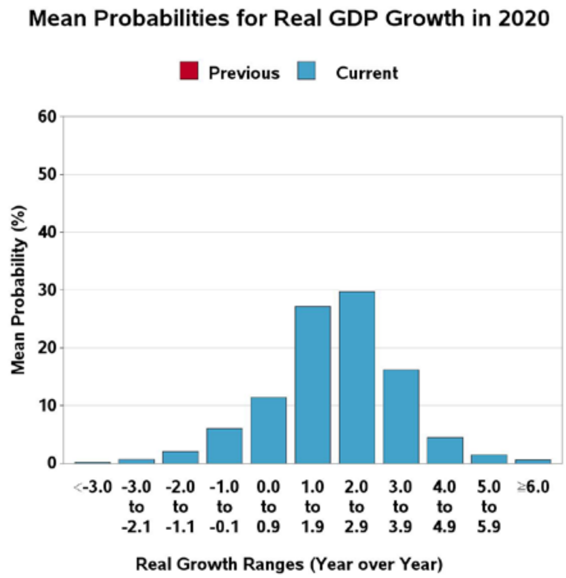
	Real GDP (%)		Unemployment Rate (%)		Payrolls (000s/month)	
	Previous	New	Previous	New	Previous	New
Quarterly data:						
2017:Q1	2.2	2.2	4.8	4.7	161.0	184.3
2017:Q2	2.2	2.3	4.7	4.6	179.2	167.0
2017:Q3	2.2	2.4	4.7	4.6	166.2	168.9
2017:Q4	2.2	2.4	4.7	4.5	166.0	160.3
2018:Q1	N.A.	2.2	N.A.	4.5	N.A.	157.6
Annual data (projections are based on annual-average levels):						
2017	2.2	2.3	4.7	4.6	173.6	180.3
2018	2.1	2.4	4.6	4.5	N.A.	164.5
2019	2.1	2.6	4.7	4.5	N.A.	N.A.
2020	N.A.	2.1	N.A.	4.6	N.A.	N.A.

The charts below provide some insight into the degree of uncertainty the forecasters have about their projections for the rate of growth in the annual-average level of real GDP. Each chart (except the one for 2020) presents the forecasters' previous and current estimates of the probability that growth will fall into each of 11 ranges. The charts show the forecasters have revised upward their estimates of the probability that real GDP growth will be above 3.0 percent in 2017, 2018, and 2019.



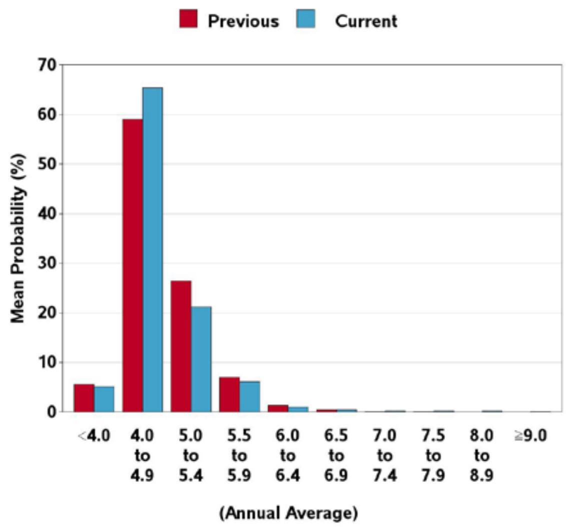




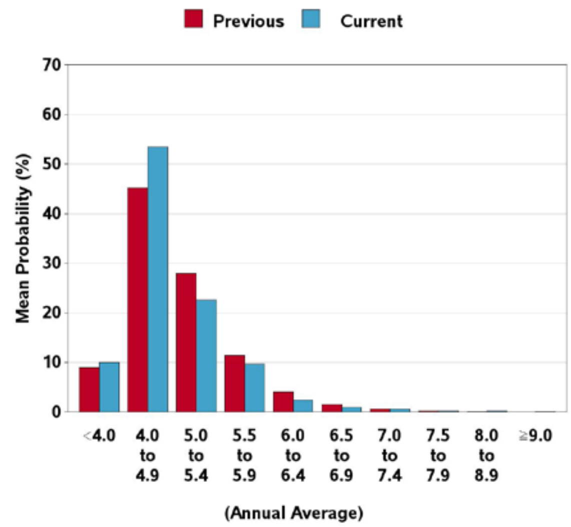


The forecasters’ density projections for unemployment, shown below, shed light on uncertainty about the labor market over the next four years. Each chart presents the forecasters’ current estimates of the probability that unemployment will fall into each of 10 ranges. The charts show the panelists are raising their density estimates over the next three years at the range of 4.0 percent to 4.9 percent of unemployment outcomes.

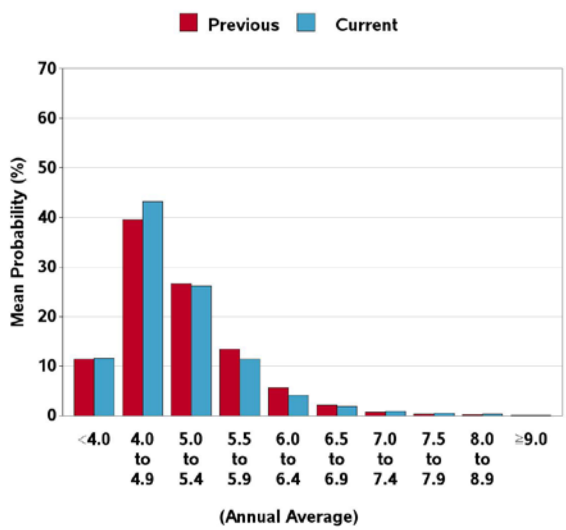
Mean Probabilities for Unemployment Rate in 2017



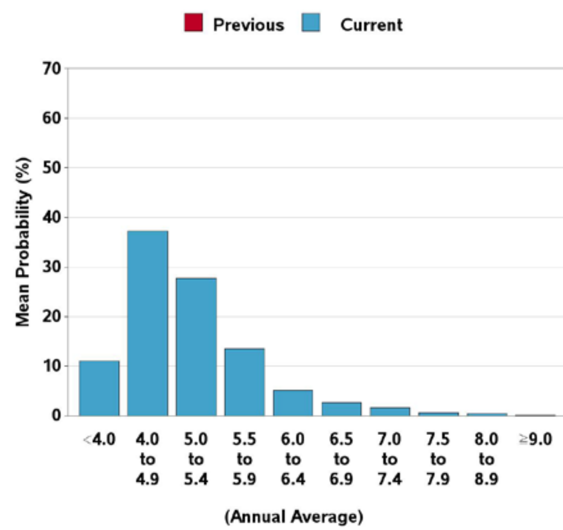
Mean Probabilities for Unemployment Rate in 2018



Mean Probabilities for Unemployment Rate in 2019



Mean Probabilities for Unemployment Rate in 2020



Forecasters See Higher Inflation

The forecasters expect higher headline CPI inflation in 2017 and 2018 than they predicted three months ago. Measured on a fourth-quarter over fourth-quarter basis, headline CPI inflation is expected to average 2.4 percent in 2017 and 2.3 percent in 2018, up from 2.2 percent in both 2017 and 2018 in the last survey. The forecasters have also revised upward slightly their projections for headline PCE inflation in 2017 to 2.0 percent, up from 1.9 percent in the survey of three months ago.

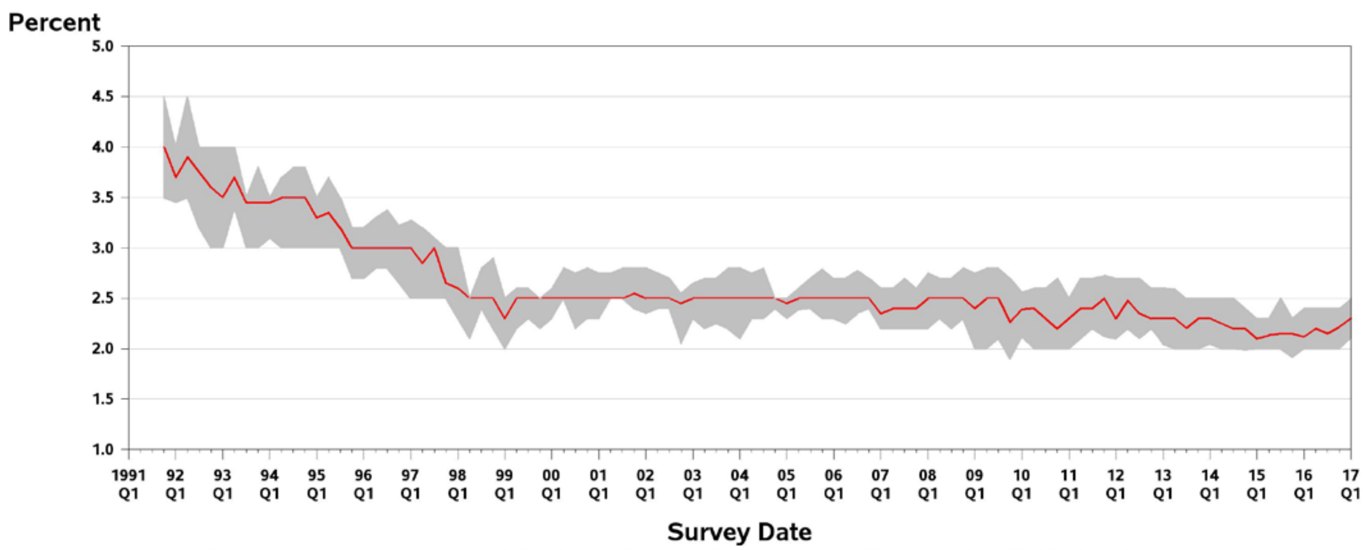
Over the next 10 years, 2017 to 2026, the forecasters expect headline CPI inflation to average 2.30 percent at an annual rate. The corresponding estimate for 10-year annual-average PCE inflation is 2.10 percent.

Median Short-Run and Long-Run Projections for Inflation (Annualized Percentage Points)

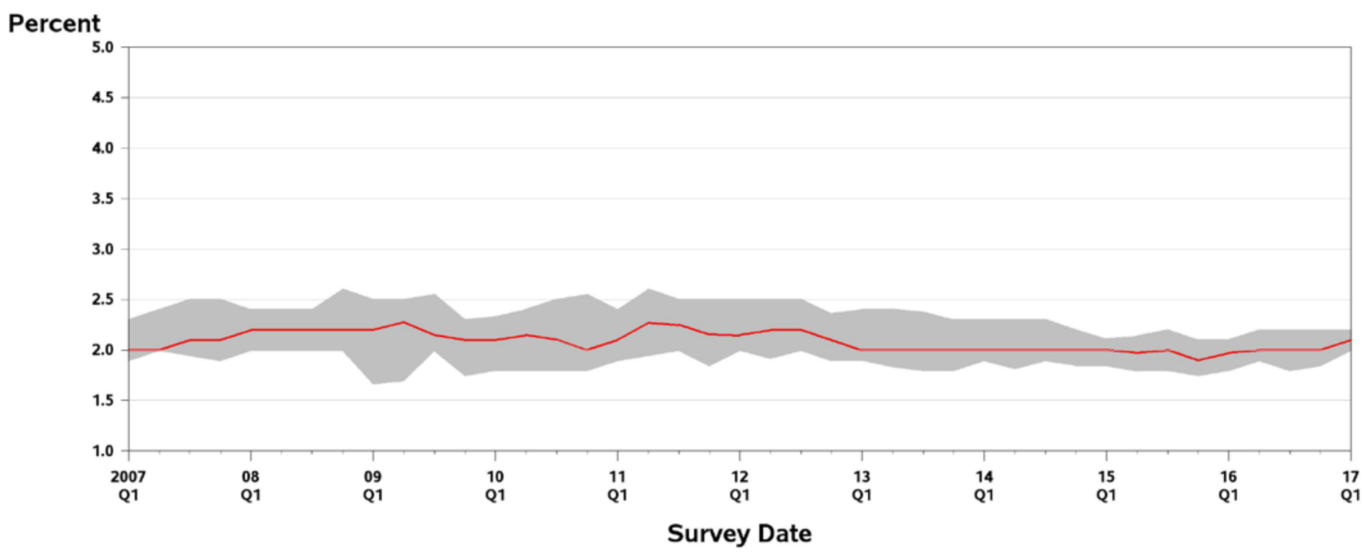
	Headline CPI		Core CPI		Headline PCE		Core PCE	
	Previous	Current	Previous	Current	Previous	Current	Previous	Current
Quarterly								
2017:Q1	2.2	2.5	2.2	2.4	1.8	2.0	1.8	1.8
2017:Q2	2.2	2.3	2.2	2.2	1.9	2.0	1.8	1.9
2017:Q3	2.2	2.3	2.2	2.1	1.9	2.0	1.9	1.9
2017:Q4	2.2	2.5	2.2	2.2	2.0	2.1	1.9	1.9
2018:Q1	N.A.	2.4	N.A.	2.3	N.A.	2.1	N.A.	2.0
Q4/Q4 Annual Averages								
2017	2.2	2.4	2.2	2.2	1.9	2.0	1.9	1.9
2018	2.2	2.3	2.2	2.3	2.0	2.0	1.9	2.0
2019	N.A.	2.3	N.A.	2.2	N.A.	2.0	N.A.	2.0
Long-Term Annual Averages								
2016-2020	2.13	N.A.	N.A.	N.A.	1.90	N.A.	N.A.	N.A.
2017-2021	N.A.	2.30	N.A.	N.A.	N.A.	2.03	N.A.	N.A.
2016-2025	2.22	N.A.	N.A.	N.A.	2.00	N.A.	N.A.	N.A.
2017-2026	N.A.	2.30	N.A.	N.A.	N.A.	2.10	N.A.	N.A.

The charts below show the median projections (the red line) and the associated interquartile ranges (gray areas around the red line) for the projections for 10-year annual-average CPI and PCE inflation. The top panel shows a higher level of the long-term projection for CPI inflation, at 2.3 percent. The bottom panel depicts the higher 10-year forecast for PCE inflation, at 2.1 percent.

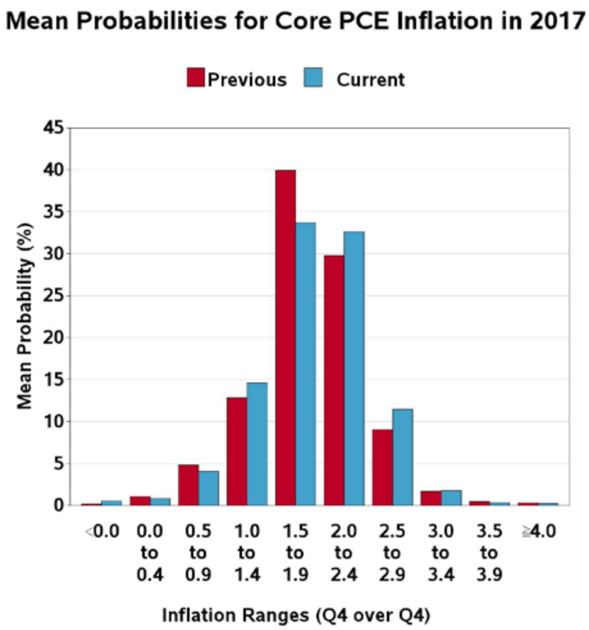
Projections for the 10-Year Annual-Average Rate of CPI Inflation
(Median and Interquartile Range)

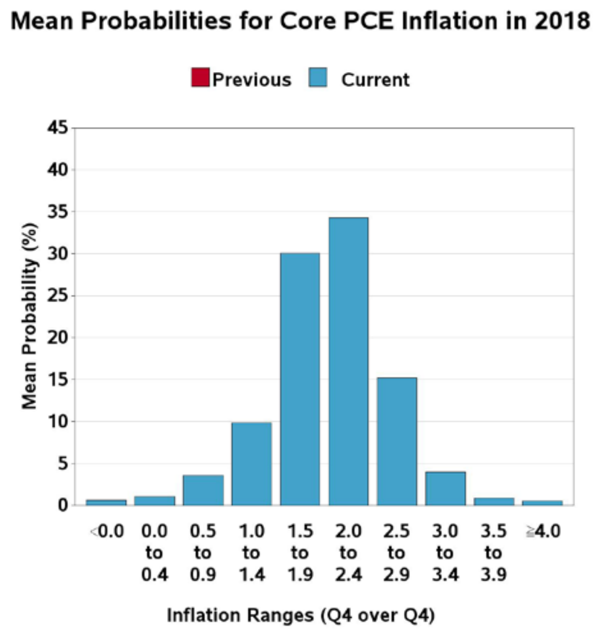


Projections for the 10-Year Annual-Average Rate of PCE Inflation
(Median and Interquartile Range)



The figures below show the probabilities that the forecasters are assigning to the possibility that fourth-quarter over fourth-quarter core PCE inflation in 2017 and 2018 will fall into each of 10 ranges. For 2017, the forecasters have increased the probability that core PCE inflation will be above 2.0 percent, compared with their estimates in the survey of three months ago.





Lower Risk of a Negative Quarter

The forecasters have revised downward the chance of a contraction in real GDP in any of the next four quarters. For the current quarter, the forecasters predict a 7.7 percent chance of negative growth, down from 14.0 percent in the survey of three months ago. The panelists have also made downward revisions to their forecasts for the next three quarters in 2017.

Risk of a Negative Quarter (%)
Survey Means

Quarterly data:	Previous	New
2017:Q1	14.0	7.7
2017:Q2	15.0	11.2
2017:Q3	16.5	14.6
2017:Q4	18.9	16.2
2018:Q1	N.A.	17.7

Forecasters State Their Views on Home Price Growth over the Next Two Years
In this survey, a special question asked panelists to provide their forecasts for fourth-quarter over fourth-quarter growth in house prices, as measured by a number of alternative indices. The panelists were allowed to choose their measure from a list of indices or to write in their own index. For each index of their choosing, the panelists provided forecasts for growth in 2017 and 2018.

Eighteen panelists answered the special question. Some panelists provided projections for more than one index. The table below provides a summary of the forecasters' responses. The number of responses (N) is low for each index. The median estimates for the seven house-price indices listed in the table below range from 3.9 percent to 5.4 percent in 2017 and from 3.8 percent to 4.7 percent in 2018.

Projections for Growth in Various Indices of House Prices
Q4/Q4, Percentage Points

	2017 (Q4/Q4 Percent Change)			2018 (Q4/Q4 Percent Change)		
Index	N	Mean	Median	N	Mean	Median
S&P CoreLogic Case-Shiller: U.S. National	7	3.8	5.0	7	3.6	4.3
S&P CoreLogic Case-Shiller: Composite 10	3	3.9	4.0	3	4.0	3.8
S&P CoreLogic Case-Shiller: Composite 20	4	4.0	3.9	4	3.8	4.0
FHFA: U.S. Total	7	5.4	5.4	7	4.4	4.5
FHFA: Purchase Only	6	4.5	4.9	6	4.0	4.4
CoreLogic: National HPI, incl. Distressed Sales (Single Family Combined)	3	4.4	4.5	3	4.1	4.0
NAR Median: Total Existing	1	5.4	5.4	1	4.7	4.7

Forecasters See Higher Long-Run Growth in Output and Productivity and in Returns to Financial Assets
In our first-quarter surveys, the forecasters provide their long-run projections for an expanded set of variables, including growth in output and productivity, as well as returns on financial assets.

As the table below shows, the forecasters have increased their estimates for the annual-average rate of growth in real GDP over the next 10 years. Currently, the forecasters expect real GDP to grow at an annual-average rate of 2.45 percent over the next 10 years, up from their projection of 2.28 percent in the first-quarter survey of 2016. Ten-year annual average productivity growth is now expected to average 1.60 percent, up from 1.40 percent.

Upward revisions to the return on the financial assets accompany the current outlook. The forecasters see the S&P 500 returning an annual-average 6.00 percent per year over the next 10 years, up from 5.37 percent in last year's first-quarter survey. The forecasters expect the rate on 10-year Treasuries to average 3.86 percent over the next 10 years, up from 3.39 percent in last year's first-quarter survey. Three-month Treasury bills will return an annual-average 2.50 percent per year over the next 10 years, unchanged from last year's survey.

Median Long-Term (10-Year) Forecasts (%)

	First Quarter 2016	*Current Survey*
Real GDP Growth	2.28	2.45
Productivity Growth	1.40	1.60
Stock Returns (S&P 500)	5.37	6.00
Rate on 10-Year Treasury Bonds	3.39	3.86
Bill Returns (3-Month)	2.50	2.50

Technical Notes

Moody's Aaa and Baa Historical Rates
The historical values of Moody's Aaa and Baa rates are proprietary and, therefore, not available in the data files on the Bank's website or on the tables that accompany the survey's complete write-up in the PDF.

New File Format Coming
On May 12, 2017, the survey's data files on the Bank's website will be changed to a .xlsx extension instead of .xls.

The Federal Reserve Bank of Philadelphia thanks the following forecasters for their participation in recent surveys:

Lewis Alexander, Nomura Securities; **Scott Anderson**, Bank of the West (BNP Paribas Group); **Robert J. Barbera**, Johns Hopkins University Center for Financial Economics; **Peter Bernstein**, RCF Economic and Financial Consulting, Inc.; **Christine Chmura**, **Ph.D.**, and **Xiaobing Shuai**, **Ph.D.**, Chmura Economics & Analytics; **Gary Ciminero**, **CFA**, GLC Financial Economics; **Nathaniel Curtis**, Navigant Consulting; **Gregory Daco**, Oxford Economics USA, Inc.; **Rajeev Dhawan**, Georgia State University; **Robert Dietz**, National Association of Home Builders; **Gabriel Ehrlich**, **Daniil Manaenkov**, **Ben Meiselman**, and **Aditi Thapar**, RSQE, University of Michigan; **Michael R. Englund**, Action Economics, LLC; **J.D. Foster**, U.S. Chamber of Commerce; **Michael Gapen**, Barclays Capital; **James Glassman**, JPMorgan Chase & Co.; **Jan Hatzius**, Goldman Sachs; **Keith Hembre**, Nuveen Asset Management; **Peter Hooper**, Deutsche Bank Securities, Inc.; **IHS Markit**; **Sam Kahan**, Kahan Consulting Ltd. (ACT Research LLC); **N. Karp**, BBVA Research USA; **Walter Kemmsies**, Jones Lang LaSalle; **Jack Kleinhenz**, Kleinhenz & Associates, Inc.; **Thomas Lam**, RHB Securities Singapore Pte. Ltd.; **L. Douglas Lee**, Economics from Washington; **John Lonski**, Moody's Capital Markets Group; **Macroeconomic Advisers**, **LLC**; **R. Anthony Metz**, Pareto Optimal Economics; **Michael Moran**, Daiwa Capital Markets America; **Joel L. Naroff**, Naroff Economic Advisors; **Mark Nielson, Ph.D.**, MacroEcon Global Advisors; **Luca Noto**, Anima Sgr; **Brendon Ogmundson**, BC Real Estate Association; **Tom Porcelli**, RBC Capital Markets; **Arun Raha** and **Maira Trimble**, Eaton Corporation; **Philip Rothman**, East Carolina University; **Chris Rupkey**, MUFG Union Bank; **John Silvia**, Wells Fargo; **Allen Sinai**, Decision Economics, Inc.; **Sean M. Snaith**, **Ph.D.**, University of Central Florida; **Constantine G. Soras, Ph.D.**, CGS Economic Consulting; **Stephen Stanley**, Amherst Pierpont Securities; **Charles Steindel**, Ramapo College of New Jersey; **Susan M. Sterne**, Economic Analysis Associates, Inc.; **James Sweeney**, Credit Suisse; **Thomas Kevin Swift**, American Chemistry Council; **Richard Yamarone**, Bloomberg, LP; **Mark Zandi**, Moody's Analytics; **Ellen Zentner**, Morgan Stanley.

This is a partial list of participants. We also thank those who wish to remain anonymous.

SUMMARY TABLE
SURVEY OF PROFESSIONAL FORECASTERS
MAJOR MACROECONOMIC INDICATORS

	2017 Q1	2017 Q2	2017 Q3	2017 Q4	2018 Q1	2017	2018	2019	2020
						(YEAR-OVER-YEAR)			
PERCENT GROWTH AT ANNUAL RATES									
1. REAL GDP (BILLIONS, CHAIN WEIGHTED)	2.2	2.3	2.4	2.4	2.2	2.3	2.4	2.6	2.1
2. GDP PRICE INDEX (PERCENT CHANGE)	2.1	2.1	2.0	2.3	2.0	2.0	2.2	N.A.	N.A.
3. NOMINAL GDP ($ BILLIONS)	4.2	4.3	4.5	4.7	4.7	4.4	4.6	N.A.	N.A.
4. NONFARM PAYROLL EMPLOYMENT									
(PERCENT CHANGE)	1.5	1.4	1.4	1.3	1.3	1.5	1.3	N.A.	N.A.
(AVG MONTHLY CHANGE)	184.3	167.0	168.9	160.3	157.6	180.3	164.5	N.A.	N.A.
VARIABLES IN LEVELS									
5. UNEMPLOYMENT RATE (PERCENT)	4.7	4.6	4.6	4.5	4.5	4.6	4.5	4.5	4.6
6. 3-MONTH TREASURY BILL (PERCENT)	0.6	0.8	1.0	1.1	1.3	0.9	1.6	2.2	2.6
7. 10-YEAR TREASURY BOND (PERCENT)	2.5	2.6	2.7	2.8	2.9	2.6	3.0	3.4	3.6

	2017 Q1	2017 Q2	2017 Q3	2017 Q4	2018 Q1	2017	2018	2019
						(Q4-OVER-Q4)		
INFLATION INDICATORS								
8. CPI (ANNUAL RATE)	2.5	2.3	2.3	2.5	2.4	2.4	2.3	2.3
9. CORE CPI (ANNUAL RATE)	2.4	2.2	2.1	2.2	2.3	2.2	2.3	2.2
10. PCE (ANNUAL RATE)	2.0	2.0	2.0	2.1	2.1	2.0	2.0	2.0
11. CORE PCE (ANNUAL RATE)	1.8	1.9	1.9	1.9	2.0	1.9	2.0	2.0

THE FIGURES ON EACH LINE ARE MEDIANS OF 42 INDIVIDUAL FORECASTERS.

SOURCE: RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA.
SURVEY OF PROFESSIONAL FORECASTERS, FIRST QUARTER 2017.

SURVEY OF PROFESSIONAL FORECASTERS

First Quarter 2017

Tables

Note: Data in these tables listed as "actual" are the data that were available to the forecasters when they were sent the survey questionnaire on January 27, 2017; the tables do not reflect subsequent revisions to the data. All forecasts were received on or before February 7, 2017.

TABLE ONE
MAJOR MACROECONOMIC INDICATORS
MEDIANS OF FORECASTER PREDICTIONS

	NUMBER OF FORECASTERS	ACTUAL 2016 Q4	FORECAST 2017 Q1	2017 Q2	2017 Q3	2017 Q4	2018 Q1	ACTUAL 2016 ANNUAL	FORECAST 2017 ANNUAL	2018 ANNUAL	2019 ANNUAL	2020 ANNUAL
1. GROSS DOMESTIC PRODUCT (GDP) ($ BILLIONS)	40	18861	19057	19261	19476	19700	19928	18567	19378	20265	N.A.	N.A.
2. GDP PRICE INDEX (2009=100)	40	112.24	112.82	113.40	113.97	114.61	115.19	111.45	113.69	116.18	N.A.	N.A.
3. CORPORATE PROFITS AFTER TAXES ($ BILLIONS)	20	N.A.	1598.9	1622.5	1649.6	1678.0	1698.0	N.A.	1631.2	1727.7	N.A.	N.A.
4. UNEMPLOYMENT RATE (PERCENT)	39	4.7	4.7	4.6	4.6	4.5	4.5	4.9	4.6	4.5	4.5	4.6
5. NONFARM PAYROLL EMPLOYMENT (THOUSANDS)	35	145131	145684	146185	146692	147172	147645	144314	146477	148451	N.A.	N.A.
6. INDUSTRIAL PRODUCTION (2012=100)	38	104.2	104.6	105.2	105.8	106.3	106.9	104.2	105.5	107.8	N.A.	N.A.
7. NEW PRIVATE HOUSING STARTS (ANNUAL RATE, MILLIONS)	36	1.22	1.22	1.25	1.27	1.28	1.30	1.17	1.26	1.33	N.A.	N.A.
8. 3-MONTH TREASURY BILL RATE (PERCENT)	36	0.43	0.58	0.75	0.96	1.13	1.32	0.32	0.86	1.56	2.22	2.64
9. MOODY'S AAA CORP BOND YIELD * (PERCENT)	23	N.A.	4.00	4.10	4.18	4.40	4.55	N.A.	4.17	4.63	N.A.	N.A.
10. MOODY'S BAA CORP BOND YIELD * (PERCENT)	24	N.A.	4.81	4.96	5.10	5.33	5.38	N.A.	5.04	5.55	N.A.	N.A.
11. 10-YEAR TREASURY BOND YIELD (PERCENT)	39	2.13	2.47	2.60	2.66	2.80	2.87	1.84	2.63	3.00	3.40	3.60
12. REAL GDP (BILLIONS, CHAIN WEIGHTED)	40	16805	16896	16994	17093	17194	17290	16660	17043	17450	17902	18282
13. TOTAL CONSUMPTION EXPENDITURE (BILLIONS, CHAIN WEIGHTED)	38	11640.4	11712.5	11785.0	11858.6	11931.3	12009.1	11514.9	11823.1	12120.1	N.A.	N.A.
14. NONRESIDENTIAL FIXED INVESTMENT (BILLIONS, CHAIN WEIGHTED)	36	2205.5	2226.3	2246.1	2266.0	2290.9	2313.9	2190.7	2256.5	2343.5	N.A.	N.A.
15. RESIDENTIAL FIXED INVESTMENT (BILLIONS, CHAIN WEIGHTED)	36	596.8	604.4	611.5	619.2	626.2	634.3	592.2	614.9	647.9	N.A.	N.A.
16. FEDERAL GOVERNMENT C & I (BILLIONS, CHAIN WEIGHTED)	36	1121.1	1123.8	1126.2	1129.2	1132.7	1138.7	1120.5	1128.2	1145.8	N.A.	N.A.
17. STATE AND LOCAL GOVT C & I (BILLIONS, CHAIN WEIGHTED)	35	1792.0	1796.5	1802.6	1808.5	1814.4	1821.2	1786.6	1804.8	1831.5	N.A.	N.A.
18. CHANGE IN PRIVATE INVENTORIES (BILLIONS, CHAIN WEIGHTED)	35	48.7	40.0	42.0	47.0	47.8	47.9	21.8	45.0	46.6	N.A.	N.A.
19. NET EXPORTS (BILLIONS, CHAIN WEIGHTED)	37	-599.6	-606.6	-620.1	-628.9	-641.9	-660.9	-561.7	-623.3	-668.8	N.A.	N.A.

* THE HISTORICAL VALUES OF MOODY'S AAA AND BAA RATES ARE PROPRIETARY AND THEREFORE NOT AVAILABLE TO THE GENERAL PUBLIC.

SOURCE: RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA. SURVEY OF PROFESSIONAL FORECASTERS, FIRST QUARTER 2017.

TABLE TWO
MAJOR MACROECONOMIC INDICATORS
PERCENTAGE CHANGES AT ANNUAL RATES

	NUMBER OF FORECASTERS	Q4 2016 TO Q1 2017	Q1 2017 TO Q2 2017	Q2 2017 TO Q3 2017	Q3 2017 TO Q4 2017	Q4 2017 TO Q1 2018	2016 TO 2017	2017 TO 2018	2018 TO 2019	2019 TO 2020
1. GROSS DOMESTIC PRODUCT (GDP) ($ BILLIONS)	40	4.2	4.3	4.5	4.7	4.7	4.4	4.6	N.A.	N.A.
2. GDP PRICE INDEX (2009=100)	40	2.1	2.1	2.0	2.3	2.0	2.0	2.2	N.A.	N.A.
3. CORPORATE PROFITS AFTER TAXES ($ BILLIONS)	20	2.7	6.0	6.9	7.1	4.9	6.1	5.9	N.A.	N.A.
4. UNEMPLOYMENT RATE (PERCENT)	39	0.0	-0.1	-0.0	-0.1	-0.0	-0.2	-0.1	0.0	0.1
5. NONFARM PAYROLL EMPLOYMENT										
(PERCENT CHANGE)	35	1.5	1.4	1.4	1.3	1.3	1.5	1.3	N.A.	N.A.
(AVG MONTHLY CHANGE)	35	184.3	167.0	168.9	160.3	157.6	180.3	164.5	N.A.	N.A.
6. INDUSTRIAL PRODUCTION (2012=100)	38	1.7	2.1	2.2	2.0	2.2	1.3	2.2	N.A.	N.A.
7. NEW PRIVATE HOUSING STARTS (ANNUAL RATE, MILLIONS)	36	1.1	10.2	6.4	5.0	5.0	7.6	5.6	N.A.	N.A.
8. 3-MONTH TREASURY BILL RATE (PERCENT)	36	0.14	0.18	0.20	0.17	0.19	0.54	0.70	0.66	0.43
9. MOODY'S AAA CORP BOND YIELD * (PERCENT)	23	N.A.	0.10	0.08	0.22	0.15	N.A.	0.46	N.A.	N.A.
10. MOODY'S BAA CORP BOND YIELD * (PERCENT)	24	N.A.	0.16	0.14	0.22	0.05	N.A.	0.51	N.A.	N.A.
11. 10-YEAR TREASURY BOND YIELD (PERCENT)	39	0.34	0.13	0.06	0.14	0.07	0.79	0.37	0.40	0.20
12. REAL GDP (BILLIONS, CHAIN WEIGHTED)	40	2.2	2.3	2.4	2.4	2.2	2.3	2.4	2.6	2.1
13. TOTAL CONSUMPTION EXPENDITURE (BILLIONS, CHAIN WEIGHTED)	38	2.5	2.5	2.5	2.5	2.6	2.7	2.5	N.A.	N.A.
14. NONRESIDENTIAL FIXED INVESTMENT (BILLIONS, CHAIN WEIGHTED)	36	3.8	3.6	3.6	4.5	4.1	3.0	3.9	N.A.	N.A.
15. RESIDENTIAL FIXED INVESTMENT (BILLIONS, CHAIN WEIGHTED)	36	5.2	4.8	5.1	4.6	5.3	3.8	5.4	N.A.	N.A.
16. FEDERAL GOVERNMENT C & I (BILLIONS, CHAIN WEIGHTED)	36	0.9	0.9	1.1	1.2	2.2	0.7	1.6	N.A.	N.A.
17. STATE AND LOCAL GOVT C & I (BILLIONS, CHAIN WEIGHTED)	35	1.0	1.4	1.3	1.3	1.5	1.0	1.5	N.A.	N.A.
18. CHANGE IN PRIVATE INVENTORIES (BILLIONS, CHAIN WEIGHTED)	35	-8.7	2.0	5.0	0.8	0.1	23.2	1.6	N.A.	N.A.
19. NET EXPORTS (BILLIONS, CHAIN WEIGHTED)	37	-7.0	-13.5	-8.8	-13.0	-19.0	-61.7	-45.5	N.A.	N.A.

* THE HISTORICAL VALUES OF MOODY'S AAA AND BAA RATES ARE PROPRIETARY AND THEREFORE NOT AVAILABLE TO THE GENERAL PUBLIC.

NOTE: FIGURES FOR UNEMPLOYMENT RATE, 3-MONTH TREASURY BILL RATE, MOODY'S AAA CORPORATE BOND YIELD, MOODY'S BAA CORPORATE BOND YIELD, AND 10-YEAR TREASURY BOND YIELD ARE CHANGES IN THESE RATES, IN PERCENTAGE POINTS. FIGURES FOR CHANGE IN PRIVATE INVENTORIES AND NET EXPORTS ARE CHANGES IN BILLIONS OF CHAIN-WEIGHTED DOLLARS. ALL OTHERS ARE PERCENTAGE CHANGES AT ANNUAL RATES.

SOURCE: RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA. SURVEY OF PROFESSIONAL FORECASTERS, FIRST QUARTER 2017.

TABLE THREE
MAJOR PRICE INDICATORS
MEDIANS OF FORECASTER PREDICTIONS

	NUMBER OF FORECASTERS	ACTUAL	FORECAST(Q/Q)					ACTUAL	FORECAST(Q4/Q4)		
		2016 Q4	2017 Q1	2017 Q2	2017 Q3	2017 Q4	2018 Q1	2016 ANNUAL	2017 ANNUAL	2018 ANNUAL	2019 ANNUAL
1. CONSUMER PRICE INDEX (ANNUAL RATE)	40	3.4	2.5	2.3	2.3	2.5	2.4	1.8	2.4	2.3	2.3
2. CORE CONSUMER PRICE INDEX (ANNUAL RATE)	37	2.0	2.4	2.2	2.1	2.2	2.3	2.2	2.2	2.3	2.2
3. PCE PRICE INDEX (ANNUAL RATE)	37	2.2	2.0	2.0	2.0	2.1	2.1	1.5	2.0	2.0	2.0
4. CORE PCE PRICE INDEX (ANNUAL RATE)	37	1.3	1.8	1.9	1.9	1.9	2.0	1.7	1.9	2.0	2.0

SOURCE: RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA. SURVEY OF PROFESSIONAL FORECASTERS, FIRST QUARTER 2017.

TABLE FOUR
ESTIMATED PROBABILITY OF DECLINE IN REAL GDP

ESTIMATED PROBABILITY (CHANCES IN 100)	Q4 2016 TO Q1 2017	Q1 2017 TO Q2 2017	Q2 2017 TO Q3 2017	Q3 2017 TO Q4 2017	Q4 2017 TO Q1 2018
	NUMBER OF FORECASTERS				
10 OR LESS	32	19	14	11	10
11 TO 20	4	15	19	18	15
21 TO 30	1	3	4	7	11
31 TO 40	0	0	0	1	1
41 TO 50	0	0	0	0	0
51 TO 60	0	0	0	0	0
61 TO 70	0	0	0	0	0
71 TO 80	0	0	0	0	0
81 TO 90	0	0	0	0	0
91 AND OVER	0	0	0	0	0
NOT REPORTING	5	5	5	5	5
MEAN AND MEDIAN					
MEDIAN PROBABILITY	6.73	10.00	15.00	15.00	15.00
MEAN PROBABILITY	7.68	11.21	14.61	16.15	17.73

NOTE: TOTAL NUMBER OF FORECASTERS REPORTING IS 37.
SOURCE: RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA.
SURVEY OF PROFESSIONAL FORECASTERS, FIRST QUARTER 2017.

TABLE FIVE
MEAN PROBABILITIES

MEAN PROBABILITY ATTACHED TO POSSIBLE
CIVILIAN UNEMPLOYMENT RATES:
(ANNUAL AVERAGE)

	2017	2018	2019	2020
9.0 PERCENT OR MORE	0.07	0.09	0.10	0.13
8.0 TO 8.9 PERCENT	0.15	0.18	0.32	0.44
7.5 TO 7.9 PERCENT	0.18	0.24	0.46	0.55
7.0 TO 7.4 PERCENT	0.24	0.58	0.86	1.64
6.5 TO 6.9 PERCENT	0.45	0.85	1.87	2.65
6.0 TO 6.4 PERCENT	1.04	2.41	4.15	5.14
5.5 TO 5.9 PERCENT	6.19	9.66	11.35	13.48
5.0 TO 5.4 PERCENT	21.15	22.56	26.11	27.74
4.0 TO 4.9 PERCENT	65.40	53.50	43.22	37.21
LESS THAN 4.0 PERCENT	5.13	9.94	11.57	11.01

MEAN PROBABILITY ATTACHED TO POSSIBLE
PERCENT CHANGES IN REAL GDP:
(ANNUAL-AVERAGE OVER ANNUAL-AVERAGE)

	2016-2017	2017-2018	2018-2019	2019-2020
6.0 OR MORE	0.23	0.31	0.41	0.64
5.0 TO 5.9	0.46	0.68	1.27	1.45
4.0 TO 4.9	2.98	3.71	4.38	4.52
3.0 TO 3.9	14.38	18.26	16.93	16.16
2.0 TO 2.9	49.66	40.75	33.85	29.75
1.0 TO 1.9	22.87	23.05	25.18	27.12
0.0 TO 0.9	6.67	8.04	11.12	11.38
-1.0 TO -0.1	1.82	2.64	4.32	6.02
-2.0 TO -1.1	0.46	1.21	1.83	2.04
-3.0 TO -2.1	0.33	1.03	0.56	0.71
LESS THAN -3.0	0.13	0.32	0.13	0.20

MEAN PROBABILITY ATTACHED TO POSSIBLE
PERCENT CHANGES IN GDP PRICE INDEX:
(ANNUAL-AVERAGE OVER ANNUAL-AVERAGE)

	2016-2017	2017-2018
4.0 OR MORE	0.46	0.42
3.5 TO 3.9	1.21	1.00
3.0 TO 3.4	3.48	3.97
2.5 TO 2.9	12.30	15.77
2.0 TO 2.4	41.90	38.29
1.5 TO 1.9	26.69	23.98
1.0 TO 1.4	8.66	10.27
0.5 TO 0.9	2.78	3.52
0.0 TO 0.4	1.59	1.65
WILL DECLINE	0.93	1.13

SOURCE: RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA.
SURVEY OF PROFESSIONAL FORECASTERS, FIRST QUARTER 2017.

TABLE SIX
MEAN PROBABILITY OF CORE CPI AND CORE PCE INFLATION (Q4/Q4)

MEAN PROBABILITY ATTACHED TO CORE CPI INFLATION:

	16Q4 TO 17Q4	17Q4 TO 18Q4
4 PERCENT OR MORE	0.42	0.87
3.5 TO 3.9 PERCENT	0.58	1.51
3.0 TO 3.4 PERCENT	3.47	6.12
2.5 TO 2.9 PERCENT	19.73	18.67
2.0 TO 2.4 PERCENT	44.65	41.15
1.5 TO 1.9 PERCENT	23.08	21.74
1.0 TO 1.4 PERCENT	5.51	6.61
0.5 TO 0.9 PERCENT	1.75	2.16
0.0 TO 0.4 PERCENT	0.42	0.65
WILL DECLINE	0.37	0.52

MEAN PROBABILITY ATTACHED TO CORE PCE INFLATION:

	16Q4 TO 17Q4	17Q4 TO 18Q4
4 PERCENT OR MORE	0.23	0.54
3.5 TO 3.9 PERCENT	0.37	0.86
3.0 TO 3.4 PERCENT	1.76	3.99
2.5 TO 2.9 PERCENT	11.42	15.20
2.0 TO 2.4 PERCENT	32.60	34.28
1.5 TO 1.9 PERCENT	33.67	30.04
1.0 TO 1.4 PERCENT	14.58	9.82
0.5 TO 0.9 PERCENT	4.02	3.56
0.0 TO 0.4 PERCENT	0.81	1.05
WILL DECLINE	0.54	0.66

SOURCE: RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA.
SURVEY OF PROFESSIONAL FORECASTERS, FIRST QUARTER 2017.

TABLE SEVEN
LONG-TERM (5-YEAR AND 10-YEAR) FORECASTS

ANNUAL AVERAGE OVER THE NEXT 5 YEARS: 2017-2021

CPI INFLATION RATE		PCE INFLATION RATE	
MINIMUM	2.00	MINIMUM	1.70
LOWER QUARTILE	2.20	LOWER QUARTILE	1.93
MEDIAN	2.30	MEDIAN	2.03
UPPER QUARTILE	2.50	UPPER QUARTILE	2.25
MAXIMUM	3.70	MAXIMUM	3.30
MEAN	2.42	MEAN	2.14
STD. DEVIATION	0.38	STD. DEVIATION	0.33
N	32	N	29
MISSING	10	MISSING	13

ANNUAL AVERAGE OVER THE NEXT 10 YEARS: 2017-2026

CPI INFLATION RATE		PCE INFLATION RATE	
MINIMUM	2.00	MINIMUM	1.70
LOWER QUARTILE	2.11	LOWER QUARTILE	2.00
MEDIAN	2.30	MEDIAN	2.10
UPPER QUARTILE	2.50	UPPER QUARTILE	2.20
MAXIMUM	4.00	MAXIMUM	4.30
MEAN	2.42	MEAN	2.19
STD. DEVIATION	0.44	STD. DEVIATION	0.48
N	31	N	28
MISSING	11	MISSING	14

REAL GDP GROWTH RATE		PRODUCTIVITY GROWTH RATE	
MINIMUM	1.60	MINIMUM	0.50
LOWER QUARTILE	2.13	LOWER QUARTILE	1.10
MEDIAN	2.45	MEDIAN	1.60
UPPER QUARTILE	2.59	UPPER QUARTILE	1.97
MAXIMUM	2.80	MAXIMUM	2.75
MEAN	2.33	MEAN	1.61
STD. DEVIATION	0.34	STD. DEVIATION	0.61
N	28	N	23
MISSING	14	MISSING	19

STOCK RETURNS (S&P 500)		BOND RATE (10-YEAR)		BILL RETURNS (3-MONTH)	
MINIMUM	0.50	MINIMUM	2.00	MINIMUM	1.00
LOWER QUARTILE	4.10	LOWER QUARTILE	3.34	LOWER QUARTILE	2.00
MEDIAN	6.00	MEDIAN	3.86	MEDIAN	2.50
UPPER QUARTILE	6.20	UPPER QUARTILE	4.00	UPPER QUARTILE	2.85
MAXIMUM	10.00	MAXIMUM	5.10	MAXIMUM	3.50
MEAN	5.60	MEAN	3.68	MEAN	2.47
STD. DEVIATION	1.98	STD. DEVIATION	0.68	STD. DEVIATION	0.60
N	19	N	26	N	25
MISSING	23	MISSING	16	MISSING	17

SOURCE: RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA.
SURVEY OF PROFESSIONAL FORECASTERS, FIRST QUARTER 2017.

Journal of Financial Economics 12 (1983) 371–386. North-Holland

ON COMPUTING MEAN RETURNS AND THE SMALL FIRM PREMIUM

Richard ROLL*

University of California, Los Angeles, CA 90024, USA

Received January 1983, final version received June 1983

The mean return computational method has a substantial effect on the estimated small firm premium. The buy-and-hold method, which best mimics actual investment experience, produces an estimated small-firm premium only one-half as large as the arithmetic and re-balanced methods which are often used in empirical studies. Similar biases can be expected in mean returns when securities are classified by any variable related to trading volume.

1. Introduction

There is a potentially serious problem in estimating expected return differences between small and large firms. Even with exactly the same sample observations, the method used to compute sample mean returns can have a substantial effect on the estimates.

With an arithmetic computational method, daily returns on individual stocks are averaged across both firms and days to obtain the mean daily return on an equally-weighted portfolio; then the portfolio's mean daily return is compounded to obtain an estimate of the expected return over a longer interval. With a buy-and-hold method, individual stock returns are first obtained for the longer interval by linking together the daily individual returns; then an equally-weighted portfolio's mean return is computed by averaging the longer-term (individual) returns.

Defining a 'longer interval' as one year, the arithmetic method produces an average annual return difference of 14.9 percent between AMEX and NYSE stocks[1] over the 19 complete calendar years, 1963–1981 inclusive. The buy-and-hold method gives an annual return difference of only 7.45 percent. Assuming that annual returns are statistically independent, the arithmetic

*Comments and suggestions by Gordon Alexander, Kenneth French, Stephen Ross and the referee, Allan Kleidon, are gratefully acknowledged.

[1]The effect of smallness can be measured by the difference in returns of stock listed on the American Exchange (AMEX) and the New York Exchange (NYSE) because AMEX issues are, on average, much smaller than NYSE issues. Most of the results presented here are based on the AMEX–NYSE differential because it is convenient and easy to use. Some confirmatory results based directly on measured size will also be presented.

method's return differential had an associated t-statistic of 3.07 while the buy-and-hold method yielded a t-statistic of 1.53.

Speculation on possible causes of the small firm premium has occupied the attention of many finance theorists over the past few years; but perhaps this attention has been premature. If the estimated small firm premium can be cut in half simply by compounding individual returns before averaging them, some consideration should be given to whether the magnitude of the true premium is really all that large. The various explanations for the premium offered so far would become more plausible if the premium is actually smaller than has been previously reported.

This paper investigates why the mean return computational method can be such a significant choice in some empirical research. The reason seems to be that individual asset returns are not as well-behaved as we might like. Individual assets do not trade continuously and there are significant trading costs. In some empirical studies, the effect of these factors might be safely ignored; but when the object of investigation is related to trading volume (and thus to trading frequency and trading costs), there can be measurement problems. Firm size is related to trading volume and it is used as an example throughout the paper. Other variables related to size and to trading, such as dividend yield, price/earnings ratio, and beta, could also present similar empirical difficulties. Section 2 gives a brief theoretical discussion of mean return computational methods and section 3 presents details of the empirical results for small firm premia.

2. Compounding and the bias in mean return calculation

2.1. Formulae for computing mean returns

To elucidate the differences in mean return computation and explain why they might produce different results, consider a sample of N securities, each having returns observed for T periods. Let R_{it} be the value relative (1 + return), of security i in period t. Suppose also that investment results are reviewed every τ periods. For example, if data were available daily but returns were to be reviewed every month, we would have $\tau \cong 21$ since there are usually about 21 trading days per month.

Two alternative methods of computing the mean equally-weighted return over the review period can be written algebraically as

$$\bar{R}_{\mathrm{AR}} = \left[\frac{1}{N \cdot \tau} \sum_i \sum_t R_{it}\right]^{\tau}, \tag{1}$$

$$\bar{R}_{\mathrm{BH}} = \frac{1}{N} \sum_i \left[\prod_t R_{it}\right], \tag{2}$$

where the subscripts 'AR' and 'BH' denote 'arithmetic' and 'buy-and-hold', respectively. These labels are intended to portray the sense of the computation method. The first method (1) is simply an arithmetic mean raised to the τth power while the second method gives the actual investment results an investor would achieve from buying equal dollar amounts of N securities and holding the shares for τ periods.

There is also a third possible definition of mean return,

$$\bar{R}_{\mathrm{RB}}=\prod_{t}\left[\frac{1}{N}\sum_{i}R_{it}\right], \tag{3}$$

where the subscript 'RB' stands for 'rebalanced'. This would be the actual investment return (ignoring transactions costs) on a portfolio which begins with equal investments in the N securities and *maintains* equal investments by rebalancing at the end of each period, $t=1,\ldots,\tau$.

To compare results over different review periods, we must choose some typical and familiar calendar interval, say a year, and express the results as percentage returns over that common calendar interval. In the tables below, annualization is accomplished and reported for 'linked' returns; the review period returns within each calendar year are simply multiplied together (or linked) in order to obtain an annual return.[2] Linked annualization includes *every* daily observation in some review period during the year. This assures that in any comparison of the results across review periods, the observed differences are due to review period alone and cannot be ascribed to slightly different sample observations.

The next two subsections investigate some properties of these sample mean returns. Subsection 2.2 derives their expected values under the assumption of temporally independent individual asset returns. Subsection 2.3 then examines the effect of intertemporal dependence.

[2]The exact formulae for linked returns can be written as follows. Let $\bar{R}_m(y,\tau)$ denote the mean annualized linked return for year y ($y=1,\ldots,Y$) using a review period whose length is τ trading days and using method ($m=\mathrm{BH,AR,RB}$), to compute the review period returns. Then,

$$\bar{R}_{\mathrm{BH}}(y,\tau)=\prod_{j=(y-1)k_\tau+1}^{y\cdot k_\tau}\left[\frac{1}{N}\sum_{i}\prod_{t=(j-1)\tau+1}^{j\cdot\tau}(R_{it})\right],$$

$$\bar{R}_{\mathrm{AR}}(y,\tau)=\prod_{j=(y-1)k_\tau+1}^{y\cdot k_\tau}\left[\frac{1}{N\cdot\tau}\sum_{i}\sum_{t=(j-1)\tau+1}^{j\cdot\tau}R_{it}\right]^{\tau},$$

$$\bar{R}_{\mathrm{RB}}(y,\tau)=\prod_{j=(y-1)k_\tau+1}^{y\cdot k_\tau}\left\{\prod_{t=(j-1)\tau+1}^{j\cdot\tau}\left[\frac{1}{N}\sum_{i}R_{it}\right]\right\},$$

where $k_\tau=T/(Y\cdot\tau)$ is the number of review periods per year and T is the total number of trading days in the entire sample. When returns are reviewed in natural calendar intervals such as months, the review period cannot be a fixed number of trading days and thus τ in the formulae above varies slightly with the actual number of trading days.

2.2. Sample mean return biases with temporal independence

Following Blume (1974), assume that each individual asset return is drawn from a stationary distribution with temporally independent disturbances; that is,

$$\tilde{R}_{it} = \mu_i + \tilde{\varepsilon}_{it}, \quad \forall i, \tag{4}$$

with $E(\tilde{R}_{it}) = \mu_i$, a constant for all t, and where the unexpected return, $\tilde{\varepsilon}_{it}$, satisfies $\mathrm{cov}(\tilde{\varepsilon}_{i,t}, \tilde{\varepsilon}_{i,t-j}) = 0$ for $j \neq 0$.

The expected value of the arithmetic mean (1) can be expressed as

$$E(\bar{R}_{AR}) = E\left[\left(\frac{1}{N}\sum_i \mu_i + \tilde{h}\right)^{\tau}\right], \tag{5}$$

where

$$\tilde{h} \equiv \frac{1}{N \cdot \tau}\sum_i \sum_t \tilde{\varepsilon}_{it}$$

is the average disturbance on the equally-weighted portfolio over the sample review period τ.

The expected value of the buy-and-hold mean (2) is

$$E(\bar{R}_{BH}) = \frac{1}{N}\sum_i \left[E\prod_t (\mu_i + \tilde{\varepsilon}_{it})\right] = \frac{1}{N}\sum_i (\mu_i^{\tau}). \tag{6}$$

This follows since the expectation can be taken inside the product with independent returns and since $E(\tilde{\varepsilon}) = 0$, by definition.

The rebalancing method (3) produces a mean return whose expectation is

$$E(\bar{R}_{RB}) = \prod_t \left[\frac{1}{N}\sum_i \mu_i\right] = \left(\frac{1}{N}\sum_i \mu_i\right)^{\tau}, \tag{7}$$

where, again, the expectation can be taken inside the product because of time independence.

Expressions (5), (6) and (7) imply that the three different mean return definitions do not produce the same results. By Jensen's inequality,

$$E(\bar{R}_{AR}) \geqq E(\bar{R}_{RB}),$$ [3]

[3]Jensen's inequality for a random variable $\tilde{x}$ and a convex function $f(x)$ is $E[f(\tilde{x})] \geqq f[E(x)]$. Let $\tilde{x} \equiv (1/N)\sum_i \mu_i + \tilde{h}$; then $f(\tilde{x}) = \tilde{x}^{\tau}$ is convex since $\tau > 1$. $E(\bar{R}_{AR}) > E(\bar{R}_{RB})$ follows immediately from (5) and (7) since $E(\tilde{h}) = 0$.

with strict inequality if $\mathrm{var}(\tilde{h})>0$, and

$$\mathrm{E}(\bar{R}_{\mathrm{BH}}) \geqq \mathrm{E}(\bar{R}_{\mathrm{RB}}),^4$$

with strict inequality if $N>1$ and at least two assets have different returns. Since we generally have some randomness $[\mathrm{var}(\tilde{h})>0]$, and many securities, $(N>1)$, the rebalanced method generally should produce lower mean returns than either the arithmetic or the buy-and-hold method, provided that returns are temporally independent.

The relation between the buy-and-hold and arithmetic means is more complex; and, indeed, neither is invariably smaller than the other. The larger the cross-sectional dispersion of individual expected returns, the larger $\mathrm{E}(\bar{R}_{\mathrm{BH}})$ relative to $\mathrm{E}(\bar{R}_{\mathrm{AR}})$. But there is an offsetting influence: the larger the intertemporal dispersion of unexpected returns $(\tilde{h})$, the larger $\mathrm{E}(\bar{R}_{\mathrm{AR}})$ relative to $\mathrm{E}(\bar{R}_{\mathrm{BH}})$.[5] Their relation in a given sample depends, therefore, on the characteristics of the underlying individual returns.

2.3. Time series dependence and its effect on estimated expected returns

The effect of serial dependence is seen most easily by examining expected mean returns when the review period is doubled, say from daily to bi-daily or from bi-weekly to monthly. Assume first that returns are collected for the shorter review period and then let $\tau=2$ (a doubling of the period). Over the doubled review period, the three mean returns are

$$\bar{R}_{\mathrm{AR}}=\left[\frac{1}{N}\sum_i\left(\mu_i+\frac{\varepsilon_{i1}+\varepsilon_{i2}}{2}\right)\right]^2, \tag{8}$$

[4]Define $f(\mu_i)=\mu_i^\tau$, a convex function for $\tau>1$. With $1/N$ used as a (pseudo) probability, $\mathrm{E}(\bar{R}_{\mathrm{BH}})\geqq\mathrm{E}(\bar{R}_{\mathrm{RB}})$ follows immediately from (6) and (7). (Cf. footnote 3.) Strict inequality holds if at least two μ_i's are different. [This result was noted by Cheng and Deets in (1971).]

The inequality above grows with the cross-sectional dispersion in μ_i, ceteris paribus. To prove this, expand μ_i^τ in a Taylor series about $\bar{\mu}\equiv(1/N)\sum_i\mu_i$; the second-order term is a positive function of the cross-sectional variance in μ_i. If μ_i were cross-sectionally normally distributed, the variance *alone* would determine the size of the inequality.

[5]This can be confirmed by using a Taylor series expansion of $\mathrm{E}(\bar{R}_{\mathrm{AR}})$. Define $\bar{\mu}=(1/N)\sum_i\mu_i$; then

$$\mathrm{E}(\bar{R}_{\mathrm{AR}})=\bar{\mu}^\tau\,\mathrm{E}\left[1+\frac{\tilde{h}^2}{2}(\tau)(\tau-1)\bar{\mu}^{-2}+\frac{\tilde{h}^3}{3!}(\tau)(\tau-1)(\tau-2)\bar{\mu}^{-3}+\ldots+\tilde{h}^\tau\bar{\mu}^{-\tau}\right].$$

Jensen's inequality (see footnote 4 above), implies that $\mathrm{E}(\bar{R}_{\mathrm{BH}})>\bar{\mu}^\tau$ with the inequality being larger the larger the cross-sectional variance in μ_i. But the term in brackets just above shows that $\mathrm{E}(\bar{R}_{\mathrm{AR}})$ increases with the higher moments of $\tilde{h}$ (since $\bar{\mu}$ is strictly positive). For example, the second term in brackets involves the variance of $\tilde{h}$. Conceivably, this term could more than offset the cross-sectional variance in μ_i. If the unexpected arithmetic portfolio return h happens to be normally-distributed, the expression above simplifies to $\mathrm{E}(\bar{R}_{\mathrm{AR}})=\bar{\mu}^\tau[1+k\cdot\mathrm{var}(\tilde{h})]$ with the constant $k>0$. In this case, there is a simple and direct tradeoff between the cross-sectional variance in expected return, μ_i, and the variance of the unexpected portfolio return, $\tilde{h}$.

$$\bar{R}_{\mathrm{BH}}=\frac{1}{N}\sum_i\left[(\mu_i+\varepsilon_{i1})(\mu_i+\varepsilon_{i2})\right], \tag{9}$$

$$\bar{R}_{\mathrm{RB}}=\left[\frac{1}{N}\sum_i(\mu_i+\varepsilon_{i1})\right]\left[\frac{1}{N}\sum_i(\mu_i+\varepsilon_{i2})\right], \tag{10}$$

where $R_{it}=\mu_i+\varepsilon_{it}$ is the observed return on individual stock i $(i=1,\ldots,N)$ in period t, and μ_i is i's single-period (i.e., shorter review period) expected return.

For notational convenience, define the cross-sectional averages

$$\bar{\mu}=\frac{1}{N}\sum_i\mu_i \quad \text{and} \quad \bar{\varepsilon}_t=\frac{1}{N}\sum_i\varepsilon_{it}.$$

Then the three mean returns have expected values,

$$\mathrm{E}(\bar{R}_{\mathrm{AR}})=\bar{\mu}^2+\tfrac{1}{2}(\sigma^2_{\bar{\varepsilon}}+\sigma_{\bar{\varepsilon}_1,\bar{\varepsilon}_2}), \tag{11}$$

$$\mathrm{E}(\bar{R}_{\mathrm{BH}})=\frac{1}{N}\sum_i\mu_i^2+\frac{1}{N}\sum_i\sigma_{\varepsilon_{i1},\varepsilon_{i2}}, \tag{12}$$

$$\mathrm{E}(\bar{R}_{\mathrm{RB}})=\bar{\mu}^2+\sigma_{\bar{\varepsilon}_1,\bar{\varepsilon}_2}, \tag{13}$$

where σ^2_x is the variance of x and $\sigma_{x,y}$ is the covariance of x and y.

Even with serial dependence, the expected arithmetic mean still exceeds the expected rebalanced mean in all circumstances since,

$$\mathrm{E}(\bar{R}_{\mathrm{AR}}-\bar{R}_{\mathrm{RB}})=\tfrac{1}{2}(\sigma^2_{\bar{\varepsilon}}-\sigma_{\bar{\varepsilon}_1,\bar{\varepsilon}_2})>0. \tag{14}$$

Comparing the buy-and-hold means and the rebalanced means, we have

$$\mathrm{E}(\bar{R}_{\mathrm{BH}}-\bar{R}_{\mathrm{RB}})=\sigma^2_{\mu_i}+\left(\frac{1}{N}\sum_i\sigma_{\varepsilon_{i1},\varepsilon_{i2}}-\sigma_{\bar{\varepsilon}_1,\bar{\varepsilon}_2}\right).$$

With no serial dependence in the ε's, the term in parentheses is zero and the BH mean would exceed the RB mean by the cross-sectional variance in expected individual returns.

However, with negative serial dependence in unexpected individual returns (ε_{i1} and ε_{i2}) or positive dependence in portfolio returns ($\bar{\varepsilon}_1$ and $\bar{\varepsilon}_2$), the rebalanced mean would become larger; enough such dependence could conceivably render it larger that the buy-and-hold mean. Since the expected arithmetic mean exceeds the expected rebalanced mean, it too could be larger than the BH mean with enough serial dependence of the right type.

There is some reason to anticipate just this type of serial dependence because of the intertemporal characteristics of individual returns. Scholes and Williams (1977, pp. 313–314) explain that because of non-synchronous trading individual assets display first-order *negative* serial dependence while diversified portfolios display *positive* dependence. A difference in the sign of serial dependence between individual assets and portfolios is relevant here because buy-and-hold (BH) means are mainly affected by individual asset serial dependence [see (12)], while the arithmetic (AR) and rebalanced (RB) means are affected by portfolio serial dependence [see (11) and (13)]. The Scholes/Williams explanation implies that BH means would tend to fall as review period lengthens while the AR and RB means would tend to rise.

There is also negative serial dependence induced in very short-term returns because of the institutional arrangement of trading. Neiderhoffer and Osborne (1966) pointed out that negative serial dependence should be anticipated when a market maker is involved in most transactions (because successive transactions are conducted at either the bid or the asked price).[6]

First-order negative serial dependence in individual returns has the effect of widening the disparity between the buy-and-hold mean and the arithmetic and rebalanced means as the review period lengthens. This follows from the fact that a doubling of the review period introduces serial covariance terms in *addition* to those already present. However, the marginal effect of lengthening the review period should probably diminish as the review period becomes longer; the effect on measured mean return should be greater when changing from, say, a daily to a weekly review period than from a monthly to an annual period. The exact impact of serial dependence can, of course, only be determined empirically and we now turn to an examination of the data.

3. The empirical small firm premium

3.1. Results

In the previous section, we found that the computational formula for sample mean returns can affect the estimated expected return. The buy-and-hold (BH) mean (2) gives an unbiased estimate of the holding period return on a realistic portfolio. The rebalanced (RB) mean (3), gives an unbiased estimate of return for its strategy but it is not realistic if the period is short since rebalancing is so costly. Except under a fortuitous combination of circumstances, the arithmetic (AR) mean (3) gives a biased estimate of *both* the rebalanced and the buy-and-hold investment returns.

[6]A paper by Blume and Stambaugh (1983), which came to my attention after the first version of this paper was written, investigates this explanation for serial dependence in detail. They find empirical results very similar to those reported here. See also Cohen et al. (1979).

Although the arithmetic and rebalanced methods of calculating the mean return probably do not portray realistic investment experience, the small-firm premium is calculated as the *difference* between the two mean returns and one might hope that the improper portrayal in these methods would cancel. Unfortunately, this is not likely for several reasons. The intertemporal variance in the portfolio disturbance, $\tilde{h}$, and the cross-sectional variance in individual security expected returns, μ_i, will not be the same in samples of large and small firms. The disturbance, $\tilde{h}$, will almost certainly have a larger variance for portfolios of small firms while the cross-sectional variances of μ_i within large- and small-firm portfolios could conceivably differ in either direction. Furthermore, serial dependence has an effect which is stronger for stocks with lower trading volumes and thus with less synchronous trading and with larger bid/ask spreads.

Empirical evidence is reported in table 1. Small Firm Premia (AMEX–NYSE) are given for the 19 complete calendar years, 1963–1981, according to the method of computation and the 'review' period. As explained earlier, the 'review' period refers to the rebalancing interval for buy-and-hold returns. For example, with a monthly review period, an equal allocation is made to stocks listed on the first day of the month and the original positions are held until the end of the month. This is repeated for each calendar month of the sample. The daily rebalancing method uses the same available returns, but it re-initializes equal positions every day during the month. The arithmetic method simply averages the same available returns during the month.

In order to compare results across the different review periods, returns are annualized by linking together the review period returns obtained during the calendar year.[7] Thus, there are 19 annual observations (one for each calendar year, 1963–81), regardless of the review period.[8] Means and t-statistics are calculated from the 19 annual returns differences between exchanges; t-

[7]See footnote 2 for exact computational formulae.

[8]Daily and bi-daily returns are over trading day intervals, while weekly and longer returns use actual calendar intervals. In the weekly case, the first week of the year ends on the same day of the week as the last trading day of the previous year, say Thursday for a given year. Then weekly returns are computed from Thursday to Thursday during that year. If the year does not terminate on a Thursday trading day, the last 'weekly' return of the year is over the remaining fraction of a calendar week. This method of year-end padding was used to ensure that every daily return during a year was included, regardless of the review period. Only the bi-daily, weekly, and bi-weekly returns are subject to such padding because the other intervals are evenly divisible into years.

Weekly returns are not always for five trading day intervals. During 1968, the exchanges were closed on Wednesdays for part of the year so that a week was composed of only four trading days. Holidays are also a problem for weekly returns; if the calendar week ended on a holiday, the return was computed through the next trading day. Then the subsequent week's return covered four trading days. Bi-weekly returns were treated identically to weekly returns with respect to year-end padding, holidays, and exchange closings.

Table 1
The small firm premium as measured by the difference in returns between American Exchange and New York Exchange listed stocks, 1963–1981 (basic data are daily, January 2, 1963 — December 31, 1981).

Review period[a] (number of review periods in sample)	Return computation method[b]		
	Buy-and-hold (BH)	Arithmetic (AR)	Daily rebalancing (RB)
	AMEX–NYSE mean return differential (% per annum)[c]		
Daily (4767)	14.9 (3.16) [7.76]	14.9 (3.16) [7.76]	14.9 (3.16) [7.76]
Bi-daily (2389)	12.3 (2.64) [5.58]	14.9 (3.16) [7.06]	14.8 (3.15) [7.01]
Weekly (992)	9.81 (2.16) [3.35]	14.8 (3.15) [5.64]	14.7 (3.14) [5.62]
Bi-weekly (498)	8.27 (1.84) [2.46]	14.9 (3.14) [5.09]	14.7 (3.13) [5.07]
Monthly (228)	7.06 (1.58) [1.82]	14.9 (3.14) [4.40]	14.7 (3.11) [4.38]
Quarterly (76)	6.42 (1.43) [1.67]	15.0 (3.15) [3.88]	14.8 (3.12) [3.85]
Annual (19)	7.45 (1.53) [1.53]	15.1 (3.10) [3.10]	14.9 (3.07) [3.07]

[a]For the daily and bi-daily cases, one- and two-trading-day intervals were used respectively. For all other cases, actual calendar intervals were used. (In the weekly and bi-weekly cases, a residual interval was necessary to fill out each calendar year). All returns were compounded to an annual basis by linking successive observations within each year (see footnote 2 of the text).

[b]The computation method follows expressions (1), (2) and (3) of the text. For interested readers, the author will gladly supply a mimeographed sheet containing details on the treatment of delisting and listing securities. The main feature of the treatment of new listings and delistings was to assure that all three mean return methods employed exactly the same sample observations.

[c]t-statistics based on the 19 annual (linked) observations are in parentheses; t-statistics based on the review period returns as independent observations are given in brackets. To understand the difference in the two reported t-statistics, consider the example of the daily review period of which there are 4767 in the sample. The t-statistic in brackets is calculated from these 4767 (daily) observations (mean *daily* return divided by standard error of mean daily return). The t-statistic in brackets is calculated from 19 annual observations; each annual observation having been calculated by linking together approximately 250 (4767/19) daily observations observed during that year. In calculating the review-period-based t-statistics for the weekly and bi-weekly cases, ten days were omitted; these ten days were the reminders of partial weeks at year end. It turned out that in 10 years of the 19, the year was exactly 52 weeks plus one trading day long. An earlier version of the paper, available on request, details the effect of omitting these single-day partial weeks. N.B. This is an issue only for the bracketed t-statistics. The linked annual returns include *every* sample day.

statistics are also given based on review period returns taken as independent observations.[9]

The results most like actual investment experience are those in the first column, buy-and-hold returns. Most actual portfolios pursue a buy-and-hold strategy within a given review period with only minor modifications induced by new information about particular individual issues. The results are frequently expressed on an annual percentage basis by comparing wealth levels at the ends of successive years, i.e., after linking sub-year results.

The review period seems to have little effect on the AR and RB means. The annual average difference in returns between AMEX and NYSE issues is about fifteen percent. But for the BH means, the review period has a large impact. Monthly and longer review periods give an AMEX–NYSE return differential of only around seven percent (and the t-statistic does not indicate an overwhelming probability that the differential is even positive). The drop in the BH mean with lengthening review period is statistically significant and so is the difference between the BH and the other means.[10]

[9]Note that the t-statistics in these tables are based on the assumption that the annual returns (t-statistics in parentheses) and review period returns (t-statistics in brackets) are temporally independent. The results indicate that the AR and RB returns are, in fact, close to independent while there is negative serial dependence in the BH returns. This implies that the t-statistics for the BH means are actually *understated.*

[10]A statistical test of the significance of the review period can be conducted by considering each year's mean difference, AMEX–NYSE, as an independent observation. Let $D_{m,y,\tau}$ be the difference for year y, review period τ, and the method m (m=BH, AR, RB). Then the time series mean of $D_{m,y,\tau}-D_{m,y,\tau'}$ ($\tau \neq \tau'$) can be tested for significance under the presumption that the years constitute independent observations. t-statistics for the AR and RB means, for all combinations of τ and τ', never indicated significance. Of the 42 combinations (21 for each mean AR and RB) none exceeded 2.0, five exceeded 1.5, and 28 were less than 1.0. In contrast, the t-statistics for the BH mean comparisons across review periods are given below:

Review period τ'	Review period τ Daily	Bi-daily	Weekly	Bi-weekly	Monthly	Quarterly
Bi-daily	6.21					
Weekly	6.75	6.82				
Bi-weekly	7.67	8.37	10.8			
Monthly	8.11	8.89	11.3	9.82		
Quarterly	8.10	7.68	8.65	6.49	3.27	
Annual	5.08	4.42	2.81	1.04	−0.532	−1.67

All BH means are significantly different across-review periods except the annual mean versus the bi-weekly, monthly and quarterly means. Note that these table entries are not statistically independent of one another (they were all calculated from the same underlying data).

A similar procedure can be employed to test the statistical significance of mean computational method. The difference $D_{m,y,\tau}-D_{m',y,\tau}$ ($m \neq m'$) forms another time series across years. Based on 19 annual observations, t-statistics for the significance of this difference from zero are as follows:

Given that the BH results in table 1 are most likely to portray actual investment experience, we now turn to the interesting econometric question: What explains the observed pattern of means? To aid in answering this question, the mean returns for each exchange are presented separately in table 2. Notice that the pattern is not predicted by the expected values of the mean returns derived in section 2.2 under the assumption of temporally independent returns. With serial independence, the BH expected mean should be greater than the RB expected mean. The empirical results in table 2 show, however, that serial dependence must be present since $\bar{R}_{BH}$ falls below $\bar{R}_{RB}$ as the review period lengthens.

The arithmetic (AR) mean is larger than the rebalanced (RB) mean as was expected with or without serial dependence. However, these two means are very close and this suggests that serial dependence in *portfolio* returns is not much of an influence [Cf. eq. (14)]. Indeed, the strikingly different behavior of the BH means from the other two means indicates that negative serial dependence in individual securities is the dominant influence on the results.

In order to be certain that the AMEX–NYSE comparison measures the small firm effect properly, table 3 is presented. It contains results for the annual review period and for portfolios classified directly by size. Firm size was calculated as market capitalization (market price times number shares), at the end of each year, 1962–1980. Firms were assigned to fractiles based on market capitalization and their returns were calculated for the following year according to three mean return methods, BH, AR, and RB.

Not surprisingly, the results are consistent with the AMEX corresponding to lower size quintiles and the NYSE to higher quintiles. The overall implication is identical: viz., the estimated small firm premium is much smaller and less significant when mean returns are computed with the buy-

Review period τ	$m=$ AR, $m'=$ BH	$m=$ RB, $m'=$ BH	$m=$ AR, $m'=$ RB
	t-statistic for difference		
Bi-daily	6.82	6.30	1.47
Weekly	7.33	6.80	1.59
Bi-weekly	8.14	7.59	1.74
Monthly	8.44	7.90	2.17
Quarterly	8.21	7.69	2.72
Annual	5.85	5.48	3.16

No statistic was computed in the daily case because all three means are identical by construction in that case. Notice that the BH means are significantly smaller than the other two means for all review periods.

Although the difference between the AR and RB small firm premium is very small (cf. table 1), the AR mean premium is always larger and is significantly larger for monthly, quarterly and annual review periods. This is predicted by eq. (14); the AR mean grows with review period relative to the RB mean.

Table 2
Mean returns on NYSE and AMEX listed securities, 1963–1981.[a]

Review period	Buy-and-hold (BH)		Arithmetic (AR)		Daily rebalancing (RB)	
	NYSE	AMEX	NYSE	AMEX	NYSE	AMEX
	Mean returns (% per Annum)					
Daily	17.24	32.09	17.24	32.09	17.24	32.09
	(2.94)	(3.29)	(2.94)	(3.29)	(2.94)	(3.29)
	[5.09]	[7.72]	[5.09]	[7.72]	[5.09]	[7.72]
Bi-daily	16.93	29.23	17.53	32.42	17.24	32.09
	(2.89)	(3.03)	(2.98)	(3.31)	(2.94)	(3.29)
	[4.59]	[6.25]	[4.76]	[6.96]	[4.68]	[6.88]
Weekly	16.38	26.19	17.79	32.61	17.26	31.99
	(2.80)	(2.78)	(3.02)	(3.34)	(2.94)	(3.28)
	[4.47]	[5.32]	[4.81]	[6.44]	[4.68]	[6.32]
Bi-weekly	15.86	24.14	17.95	32.83	17.29	32.08
	(2.72)	(2.58)	(3.05)	(3.36)	(2.95)	(3.28)
	[4.29]	[4.66]	[4.71]	[5.85]	[4.58]	[5.74]
Monthly	15.34	22.39	18.07	32.96	17.34	32.08
	(2.65)	(2.42)	(3.07)	(3.36)	(2.95)	(3.28)
	[3.11]	[3.08]	[3.67]	[4.54]	[3.51]	[4.41]
Quarterly	15.01	21.42	18.17	33.17	17.38	32.19
	(2.63)	(2.33)	(3.09)	(3.38)	(2.96)	(3.29)
	[2.73]	[2.62]	[3.22]	[3.84]	[3.09]	[3.73]
Annual	15.18	22.63	17.96	33.07	17.16	32.03
	(2.69)	(2.39)	(3.11)	(3.36)	(2.98)	(3.27)
	[2.69]	[2.39]	[3.11]	[3.36]	[2.98]	[3.27]

[a]See footnotes to table 1.

and-hold method than when means are computed with the AR and RB methods.

3.2. *Implications for previous research and for the 'risk-adjusted' small firm premium*

The implications of these findings for previously-published estimates of the small firm premium are: if the basic data were very short-term and arithmetic or rebalanced means were used, the estimated premium overstates the reward investors can expect from a buy-and-hold position in small firms. Papers by Reinganum (1981a, b, 1982) and Roll (1981) used daily data and arithmetic mean returns. Reinganum's (1982) paper gives monthly and quarterly returns but these were computed with the daily rebalancing method since the author states that '... these holding period returns are created by compounding the daily *portfolio* returns' (p. 34, emphasis added).

Table 3
Mean returns and small firm premia for portfolios classified by size[a] at year-end, 1963–1981, annual review period.

Size quintile	Return computation method[b]		
	Buy-and-hold (BH)	Arithmetic (AR)	Daily rebalancing (RB)
	Mean return (% per annum)[c]		
Smallest	27.9 (2.42)	46.0 (3.68)	44.9 (3.61)
2	21.1 (2.51)	27.6 (3.15)	26.6 (3.04)
3	17.1 (2.41)	20.7 (2.86)	19.7 (2.73)
4	14.6 (2.53)	16.9 (2.89)	16.1 (2.75)
Largest	10.8 (2.50)	12.2 (2.85)	11.5 (2.68)
	Small firm premium, smallest–largest quintile (% per annum)		
	17.1 (1.88)	33.9 (3.47)	33.4 (3.46)
	Small firm premium, smallest–largest decile (% per annum)		
	22.8 (2.07)	49.1 (3.84)	48.3 (3.83)

[a]Firms are included in the kth size fractile if the closing price times the number of outstanding shares is ranked in that fractile among all listed AMEX and NYSE firms.

[b]The computation method follows expressions (1), (2) and (3) of the text. An unpublished appendix (available from the author) contains details on the treatment of listing and delisting.

[c]t-statistics based on 19 annual observations are in parentheses.

Papers with monthly returns are apparently much less subject to mean return estimation problems. Tables 1 and 2 show that there is little additional discrepancy between the BH and other means in going from monthly to annual data. The well-known paper by Banz (1981) used monthly data as did earlier papers on the closely-related stock price effect [Blume and Husic (1973), Bachrach and Galai (1979)]. Thus, it seems unlikely that the results presented in those papers will be much affected by the problem investigated here. In a more recent paper, Reinganum, (1983) used the buy-and-hold method and found results close to those reported above. Reinganum did not, however, contrast the buy-and-hold with other mean returns.

It is important to ascertain whether the *risk-adjusted* small firm premium is attributable solely to econometric problems. Is underestimation of risk for small firms [Roll (1981), Reinganum (1982)], combined with overestimation of expected returns, sufficient to induce the observed risk-adjusted premium; or is the premium really evidence of a misspecified capital asset pricing model (CAPM), perhaps because of omitted factors in the single index CAPM?

This is tantamount to asking whether the implicit CAPM market risk premium $\hat{p}$ ($\hat{p} \equiv \hat{\mathrm{E}}(R_{\mathrm{small}} - R_{\mathrm{large}})/(\hat{\beta}_{\mathrm{small}} - \hat{\beta}_{\mathrm{large}})$), is in a reasonable range. $\hat{p}$ was computed by Reinganum (1983) as 37.5 percent per annum using (a) buy-and-hold means on the smallest and largest deciles of NYSE and AMEX stocks, (b) Dimson's (1979) aggregated coefficient betas, (c) the value-weighted C.R.S.P. index and (d) daily data for 1963–1980. The return on the value-weighted index during this period was only about 9.5 percent, so $\hat{p}$ is grossly too large, thereby indicating a substantial risk-adjusted small firm premium.

The main problem with such a test was described some time ago [Roll (1977)]. Even if we make the dubious assumption that the value-weighted C.R.S.P. index is ex-ante mean/variance efficient, there is no necessity in the generalized Black (1972) C.A.P.M. that $\mathrm{E}(\hat{p}) = \mathrm{E}(R_M - R_F)$. Instead, the model requires that $\mathrm{E}(\hat{p}) = \mathrm{E}(R_M - R_Z)$ where Z is M's 'zero-beta' portfolio. Depending upon M's position on the efficient frontier, $\mathrm{E}(R_Z)$ can be negative and large.

To illustrate the difference in inferences that can be obtained with a different index, I recomputed $\hat{p}$ using (a) buy-and-hold annual means on the smallest and largest deciles of NYSE and AMEX stocks, (b) simple OLS beta coefficients estimated from annual returns,[11] (c) the *equally*-weighted C.R.S.P. index, and (d) annual data for 1963–1981.

The beta estimates (t-statistics) were $\beta_{\mathrm{small}} = 1.78$ (5.59), $\beta_{\mathrm{large}} = 0.598$ (8.60). Using the estimated premium $\mathrm{E}(R_{\mathrm{small}} - R_{\mathrm{large}}) = 22.8\%$ from table 3, we have $\hat{p} = 19.3$ percent. The actual ex post return on this market index was 15.3 percent, so $\hat{p}$ is still somewhat too high (thus indicating a risk-adjusted small-firm premium). Nevertheless, the discrepancy between a $\hat{p}$ of 19.3 and a market return of 15.3 is much less aberrant than the difference Reinganum (1983) reports between $\hat{p} = 37.5$ and $\bar{R}_M = 9.5$ percent.

It still seems that investigation of the observed small firm premium in the context of a more general asset pricing model would be a worthwhile endeavor; but estimation problems in expected returns and in simple risk parameters can explain much of the apparent anomaly.

[11]Instead of the Dimson aggregated coefficient betas, I used betas from annual data because of the now well-documented annual seasonal [Keim (1983), Roll (1983)], which has the potential to induce biases into any betas, including the Dimson type, when they are computed from non-yearly data.

5. Conclusion

Computing mean returns in order to estimate investment experience is not as easy as it sounds. Common stock data have serial dependence which, though seemingly slight, substantially affects the estimates obtained under alternate mean return computational methods. Investment experience is best portrayed by buy-and-hold portfolio returns but scholars often use arithmetic or rebalanced portfolio returns because they are easier to compute.

Perhaps this makes little difference for some studies; but if serial dependence differs systematically with the item being investigated, the computational method can be quite material.

For the small firm premium, as measured by the difference in mean returns of American Exchange and New York Exchange listed stocks, the buy-and-hold mean return difference is only about $7\frac{1}{2}$ percent per annum (for 1963–81) while the rebalanced and arithmetic methods produce annual return differences *with the same stocks and time periods* of over 14 percent. The annual difference in returns between the smallest and largest size quintiles (deciles) is about 34 (49.1) percent using the rebalanced and arithmetic methods and about 17 (22.8) percent using the buy-and-hold method.

The annual small-firm premium is only marginally significant at usual significance levels if mean returns are measured with the buy-and-hold method.

References

Bachrach, Benjamin and Dan Galai, 1979, The risk return relationship and stock prices, Journal of Financial and Quantitative Analysis 14, June, 421–441.

Banz, Rolf W., 1981, The relationship between return and market value of common stocks, Journal of Financial Economics 9, March, 3–18.

Black, Fischer, 1982, Capital market equilibrium with restricted borrowing, Journal of Business 45, July, 444–454.

Blume, Marshall, 1974, Unbiased estimators of long-run expected rates of return, Journal of the American Statistical Association 69, Sept., 634–638.

Blume, Marshall and Frank Husic, 1973, Price, beta, and exchange listing, Journal of Finance 28, May, 283–299.

Blume, Marshall and Robert F. Stambaugh, 1983, Biases in computed returns: An application to the size effect, Journal of Financial Economics, this issue.

Cheng, Pao L. and M. King Deets, 1971, Statistical biases and security rates of return, Journal of Financial and Quantitative Analysis 6, June, 977–994.

Cohen, Kalman T., Steven F. Maier, Robert A. Schwartz and David K. Whitcomb, 1979, On the existence of serial correlation in an efficient securities market, TIMS Studies in the Management Sciences 11, 151–168.

Dimson, Elroy, 1979, Risk measurement when shares are subject to infrequent trading, Journal of Financial Economics 7, June, 197–226.

Keim, Donald, 1983, Size related anomalies and stock return seasonality: Further empirical evidence, Journal of Financial Economics 12, June, 13–32.

Niederhoffer, Victor and M.F.M. Osborne, 1966, Market making and reversal on the stock exchange, Journal of the American Statistical Association 61, Dec., 898–916.

Reinganum, Marc R., 1981, The arbitrage pricing theory: Some empirical results, Journal of Finance 36, May, 313–321.
Reinganum, Marc R., 1981, Misspecification of capital asset pricing: Empirical anomalies based on earnings yields and market values, Journal of Financial Economics 9, March, 19–46.
Reinganum, Marc R., 1982, A direct test of Roll's conjecture on the firm size effect, Journal of Finance 37, March, 27–36.
Reinganum, Marc R., 1983, Portfolio strategies based on market capitalization, Journal of Portfolio Management 9, Winter, 29–36.
Roll, Richard, 1977, A critique of the asset pricing theory's tests, Journal of Financial Economics 4, March, 129–176.
Roll, Richard, 1981, A possible explanation of the small firm effect, Journal of Finance 36, Sept., 879–888.
Roll, Richard, 1983, Vas ist Das? The turn of the year effect and the return premia of small firms, Journal of Portfolio Management 9, Winter, 18–28.
Scholes, Myron and Joseph Williams, 1977, Estimating betas from nonsynchronous data, Journal of Financial Economics 5, Dec., 309–327.

The Equity Risk Premium: An Annotated Bibliography

Zhiyi Song, CFA
Greenwich Alternative Investments
Hong Kong

The equity risk premium is broadly defined as the difference between the expected total return on an equity index and the return on a riskless asset. The magnitude of the equity risk premium, arguably the most important variable in financial economics, affects the asset allocation decisions of individual and institutional investors, and the premium is a critical factor in estimating companies' costs of capital. This literature review explores research by academics and practitioners on this topic during the past three decades.

The equity risk premium (or, simply, equity premium) is broadly defined as the difference between the expected total return on an equity index and the return on a riskless asset. (Which index and which riskless asset need to be defined precisely before numerically estimating this premium.) The equity premium is considered the most important variable in financial economics. The magnitude of the equity premium strongly affects the asset allocation decisions of individual investors and institutional investors, including pensions, endowment funds, foundations, and insurance companies, and is a critical factor in estimating companies' costs of capital.

History of Research on the Equity Risk Premium

The topic of the equity risk premium (ERP) has attracted attention from academics and practitioners. There are three major themes in the intellectual history of the equity premium. The first theme builds on Gordon and Shapiro's suggestion that a dividend discount model (DDM) be used to estimate the required return on capital for a corporate project, and, by extension, the expected return on an equity (if the equity is fairly priced).[1] Specifically, the DDM says that expected total equity return equals the dividend yield plus the expected dividend growth rate; the equity premium is this sum minus the riskless rate. The DDM was widely used by practitioners to estimate the equity premium until Ibbotson and Sinquefield (1976) introduced a different approach based on historical returns. An early work by Diermeier, Ibbotson, and Siegel (1984) attempted to bolster the use of the DDM for long-range forecasting, but it was not widely used; the recent, and quite remarkable, revival of the DDM as an estimator of the equity premium dates back only to the late 1990s.

The second theme arose from Ibbotson and Sinquefield's 1976 article, which decomposed historical returns on an equity index into a part attributable to the riskless rate and a part attributable to the equity premium. The arithmetic mean of the equity premium part is assumed to be stationary—that is, the same in the future as in the past. Thus, if equities had beaten riskless Treasury bills by an arithmetic mean margin of 7 percent a year over the historical measurement period, which was usually 1926 through the then-current time, then equities were forecast to beat bills by the same amount in the future. This approach dominated practitioners' estimates of the equity premium starting in the late 1970s, but its influence has faded recently, under attack from both the DDM and the "puzzle" literature that began with Mehra and Prescott (1985).

Mehra and Prescott's 1985 article, "The Equity Premium: A Puzzle," began a third theme. The puzzle they described is that the historical equity risk premium during the period of 1889–1978 (or any other similarly long period, such as 1926 to the present) was too high, by at least an order of magnitude, to be explained by standard

[1]Myron J. Gordon and Eli Shapiro, "Capital Equipment Analysis: The Required Rate of Profit," *Management Science*, vol. 3, no. 1 (October 1956):102–110.

"general equilibrium" or "macroeconomic" asset-pricing models. Using these models, such a high premium can only be explained by a very high coefficient of risk aversion, one in the range of 30 to 40. Risk aversion parameters observed in other aspects of financial behavior are around 1. So, Mehra and Prescott argued, either the model used to describe investors' behavior is flawed or equity investors have received a higher return than they expected.

We call the asset-pricing models referenced by Mehra and Prescott (1985) "macroeconomic" because they originated in that specialty, but more importantly to distinguish them from asset-pricing models commonly used in investment finance—such as the capital asset pricing model, the three-factor Fama–French model, and arbitrage pricing theory—that are silent on the absolute size of the risk premium (in fact, requiring it as an input) and that distinguish instead among the expected *relative* returns on specific securities or portfolios.

The rest of this introductory essay focuses on attempts to resolve the equity premium "puzzle" identified by Mehra and Prescott (1985). Their "puzzle" has stimulated a remarkable response in the academic literature. Most practitioners today, however, use estimates of the equity premium that emerge from the DDM—the earliest method. Moreover, practitioner debates tend to focus on which DDM estimate to use and the extent to which the estimate should be influenced by historical returns, not the question of whether either the DDM or the historical approach can be reconciled with that of Mehra and Prescott. Reflecting practitioners' concerns, this annotated bibliography covers all three major themes in the literature.

Reconciling the "Puzzle"

Research on the question of why the realized equity premium was so large can be grouped into two broad categories: (1) studies alleging bias in the historical data and (2) studies suggesting improvements in the macroeconomic model. A third category, studies that set forth methods for estimating for the equity risk premium *independent* of the macroeconomic model, is also addressed in this review.

Biases in Historical Data. Potential biases in the historical data vary from survivorship bias and variations in transaction and tax costs to the choice of short-term bills versus long-term bonds as the riskless asset.

- *Survivorship bias*. Brown, Goetzmann, and Ross (1995) argued that the historical equity premium calculated using U.S. data is likely to overstate the true (expected) premium because the U.S. stock market turned out to be the most successful in world history. However, Dimson, Marsh, and Staunton (2006) examined stock and bond returns using data from 1900 to 2005 for 17 countries and concluded that the high historical equity premium obtained for the United States is comparable with that of other countries.
- *Transaction costs, regulations, and taxes*. McGrattan and Prescott (2001) suggested that the higher historical equity premium is mainly because of a large run-up in the equity price caused by the sharp decline in the tax rate on dividends. In their 2003 article, they claimed that the equity premium is less than 1 percent after accounting for taxes, regulations, and costs.
- *Short-term bills vs. long-term bonds as the riskless asset*. McGrattan and Prescott (2003) argued that short-term bills provide considerable liquidity services and are a negligible part of individuals' long-term debt holdings. As a result, long-term bonds should be used as the riskless asset in equity premium calculations. Siegel (2005) argued that the riskless asset that is relevant to most investors (that is, to long-term investors) is "an annuity that provides a constant real return over a long period of time" (p. 63). And the return on long-term inflation-indexed government bonds is the closest widely available proxy for such an annuity.
- *Unanticipated repricing of equities*. Bernstein (1997) suggested that because equities started the sample period (which begins in 1926) at a price-to-earnings ratio (P/E) of about 10, and ended the period at a P/E of about 20, the actual return on equities was higher than investors expected or required. Thus, the historical return overstates the future expected return. This finding was bolstered by Fama and French (2002), who used the DDM to show that investors expected an equity risk premium of about 3 percent, on average, from 1926 to the present.
- *Unanticipated poor historical bond returns*. Historical bond returns may have been biased downward because of unexpected double-digit inflation in the 1970s and 1980s (Arnott and Bernstein 2002; Siegel 2005). However, subsequent disinflation and declines in bond yields have caused the bond yield to end the historical study period only a little above where it started, thus mostly negating the validity of this objection.

Improvements in the Theoretical Model. The second broad category of research on the equity risk premium is a large body of literature exploring a variety of improvements in the original Mehra and Prescott (1985) model.

- *Rare events*. Rietz (1988) suggested that the ERP puzzle can be solved by incorporating a very small probability of a very large drop in consumption. If such a probability exists, the predicted equity premium is large (to compensate investors for the small risk of a very bad outcome). In the same year, Mehra and Prescott countered that Rietz's model requires a 1 in 100 chance of a 25 percent decline in consumption to reconcile the equity premium with a risk aversion parameter of 10, which is the approximate degree of risk aversion that would be required to predict an equity premium equal to that which was realized.[2] However, they argued, the largest aggregate consumption decline in the last 100 years was only 8.8 percent. Campbell, Lo, and MacKinlay pointed out in 1997 that "the difficulty with Rietz's argument is that it requires not only an economic catastrophe, but one which affects stock market investors more seriously than investors in short-term debt instruments" (p. 311).[3] Recently, Barro (2006) extended Rietz's model and argued that it does provide a plausible resolution of the equity premium "puzzle."
- *Recursive utility function*. One critique of the power utility function used by Mehra and Prescott (1985) is the tight link between risk aversion and intertemporal substitution. Hall argued that this link is inappropriate because the intertemporal substitution concerns the willingness of an investor to move consumption between different time periods whereas the risk aversion parameter concerns the willingness of an investor to move consumption between states of the world.[4] However, Weil (1989) showed that the ERP puzzle cannot be solved by simply separating risk aversion from intertemporal substitution. More recently, Bansal and Yaron (2004) argued that risks related to varying growth prospects and fluctuating economic uncertainty, combined with separation between the intertemporal substitution and risk aversion, can help to resolve the ERP puzzle.
- *Habit formation*. Constantinides (1990) introduced habit persistence in an effort to explain the ERP puzzle. His model assumes that an investor's utility is affected by both current and past consumption and that a small fall in consumption can generate a large drop in consumption net of the subsistence level. This preference makes investors extremely averse to consumption risk even when risk aversion is small. Constantinides showed that the historical equity premium can be explained if past consumption generates a subsistence level of consumption that is about 80 percent of the normal consumption rate.

Abel defined a similar preference, called "catching up with the Joneses," where one's utility depends not on one's absolute level of consumption, but on how one is doing relative to others.[5]

- *Borrowing constraints and life-cycle issues*. Constantinides, Donaldson, and Mehra (2002) introduced life-cycle and borrowing constraints. They argued that as the correlation of equities with personal income changes over the life of the investor, so too does the attractiveness of equities to that investor. The young, who should borrow to smooth consumption and to invest in equities, cannot do so. Therefore, equities are priced almost exclusively by middle-aged investors, who find equities to be unattractive. Thus, equities are underpriced and bonds are overpriced, producing a higher equity risk premium than predicted by Mehra and Prescott (1985).
- *Limited market participation*. Mankiw and Zeldes (1991) examined whether the consumption of stockholders differs from that of nonstockholders and whether this difference helps explain the historical equity risk premium. They showed that aggregate consumption of stockholders is more highly correlated with the stock market and is more volatile than the consumption of nonstockholders. A risk aversion parameter of 6 can explain the size of the equity premium based on consumption of stockholders alone. Although this value is still too large to be plausible, it is much less than the magnitude of 30 to 40 derived by Mehra and Prescott (1985) using the aggregate consumption data of both stockholders and nonstockholders.

---

[2] Rajnish Mehra and Edward C. Prescott, "The Equity Premium: A Solution?" *Journal of Monetary Economics*, vol. 22, no. 1 (July 1988):133–136.
[3] John Y. Campbell, Andrew W. Lo, and A. Craig MacKinlay, *The Econometrics of Financial Markets* (Princeton, NJ: Princeton University Press, 1997).
[4] Robert E. Hall, "Intertemporal Substitution in Consumption," *Journal of Political Economy*, vol. 96, no. 2 (December 1988):212–273.
[5] Andrew B. Abel, "Asset Prices under Habit Formation and Catching Up with the Joneses," *American Economic Review Papers and Proceedings*, vol. 80, no. 2 (May 1990):38–42.

■ *Incomplete markets*. Heaton and Lucas introduced uninsurable, idiosyncratic income risk into standard and dynamic general equilibrium models and showed that it can increase the risk premium.[6] Brav, Constantinides, and Geczy (2002) showed that the equity premium can be "explained with a stochastic discount factor calculated as the weighted average of the individual households' marginal rate of substitution with low and economically plausible values of the rate of risk aversion coefficient." This explanation relies on incomplete markets in that all risks would be insurable if markets were "complete."

■ *Behavioral approach*. Starting with prospect theory as proposed by Kahneman and Tversky,[7] a large swath of behavioral finance literature argues that the combination of "myopic" loss aversion and narrow framing can help to resolve the ERP puzzle, including works by Benartzi and Thaler (1995), Barberis, Huang, and Santos (2001), and Barberis and Huang (2006).

Summary

The various (and quite different, almost unrelated) approaches to estimating the equity risk premium is best summarized by Ibbotson and Chen, who categorized the estimation methods into four groups:[8]

1. *Historical method*. The historical equity risk premium, or difference in realized returns between stocks and bonds (or stocks and cash), is projected forward into the future. See Ibbotson and Sinquefield (1976), which is updated annually by Ibbotson Associates (now Morningstar), and Dimson, Marsh, and Staunton (2002).
2. *Supply-side models*. This approach uses fundamental information, such as earnings, dividends, or overall economic productivity, to estimate the equity risk premium. See Diermeier, Ibbotson, and Siegel (1984); Siegel (1999); Shiller (2000); Fama and French (1999); Arnott and Ryan (2001); Campbell, Diamond, and Shoven (2001); Arnott and Bernstein (2002); and Grinold and Kroner (2002).
3. *Demand-side models*. This approach uses a general equilibrium or macroeconomic model to calculate the expected equity return by considering the payoff demanded by investors for bearing the risk of equity investments. Mehra and Prescott (1985) is the best known example of this approach, and the "puzzle debate" is an attempt to reconcile the results of this approach with the much higher ERP estimates given by the other approaches.
4. *Surveys*. An estimate of the equity risk premium is obtained by surveying financial professionals or academics (e.g., Welch 2000). Such results presumably incorporate information from the other three methods.

In closing, the equity risk premium has been the topic of intense and often contentious research over at least the last three decades. As Siegel (2005) said, although there are good reasons why the future equity risk premium should be lower than it has been historically, a projected equity premium of 2 percent to 3 percent (over long-term bonds) will still give ample reward for investors willing to bear the risk of equities.

[6]John Heaton and Deborah Lucas, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy*, vol. 104, no. 3 (June 1996):443–487.

[7]Daniel Kahneman and Amos Tversky, "Prospect Theory: An Analysis of Decisions under Risk," *Econometrica*, vol. 47, no. 2 (March 1979):263–292.

[8]Roger Ibbotson and Peng Chen, "The Supply of Stock Market Returns," Ibbotson Associates, 2001.

Bibliography

Aït-Sahalia, Yacine, Jonathan A. Parker, and Yogo Motohiro. 2004. "Luxury Goods and the Equity Premium." *Journal of Finance*, vol. 59, no. 6 (December):2959–3004.

This article proposes a partial solution to the ERP puzzle by distinguishing between the consumption of basic goods and that of luxury goods. The authors argue that the aggregate consumption does not measure the marginal risk of investing in the stock market. Using several novel datasets on luxury goods consumption, such as sales of imported luxury automobiles, this study shows that the covariance of luxury goods and excess returns implies a risk aversion parameter of 7, significantly lower than that implied by aggregate consumption data.

Ang, Andrew, and Angela Maddaloni. 2005. "Do Demographic Changes Affect Risk Premiums? Evidence from International Data." *Journal of Business*, vol. 78, no. 1 (January):341–379.

This article examines empirically the relation between the equity risk premium and demographics using a long-term data sample (1900–2001) from the United States, Japan, the United Kingdom, Germany, and France as well as a shorter-term data sample (1970–2000) for 15 countries. By pooling international data, the authors show that a negative relation exists between the expected equity risk premium and the percentage of adults over 65 years old. The international results from this study support Abel's prediction that the equity risk premium is likely to decrease as the Baby Boom generation enters retirement.[9]

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.

Contrary to the predictions of Ibbotson and Chen (2003) and others who apply Modigliani and Miller (M&M) dividend invariance intertemporally, earnings growth has been fastest when dividend payout is highest, not lowest, because of diminishing marginal productivity of capital. Thus, investors should not look to today's low payout ratios as a sign of stronger-than-historical earnings growth in the future.

Arnott, Robert D., and Peter L. Bernstein. 2002. "What Risk Premium Is 'Normal'?" *Financial Analysts Journal*, vol. 58, no. 2 (March/April):64–85.

The expected equity return equals the dividend yield, plus dividend growth, plus the expected change in valuation, if any. As of year-end 1925, investors expected about 5.1 percent (about 1.4 percent more than the bond yield). The subsequent positive surprise was because of four historical accidents: (1) bonds had unanticipated losses; (2) valuations quadrupled, as measured by the price-to-dividend ratio (P/D); (3) the market survived; and (4) accelerated growth in real dividends and earnings occurred because of regulatory reform. These observations are used to construct a framework for estimating the equity risk premium at each point in time, including the present. The "normal" equity risk premium, or historical average of what investors were actually expecting, is 2.4 percent, and the current equity risk premium is around zero.

[9]Andrew B. Abel, "Will Bequests Attenuate the Predicted Meltdown in Stock Prices when Baby Boomers Retire?" *Review of Economics and Statistics*, vol. 83, no. 2 (November 2001):589–595; "The Effects of a Baby Boom on Stock Prices and Capital Accumulation in the Presence of Social Security," *Econometrica*, vol. 71, no. 2 (March 2003):551–578.

Arnott, Robert, 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.

Applying the dividend discount model to then-current (January 2000) valuations produces an equity risk premium of –0.9 percent, consisting of a real equity expected return of 3.2 percent minus a real Treasury Inflation-Protected Securities (TIPS) yield of 4.1 percent. A similar analysis of the equity risk premium at the end of 1925 shows that it was 2.7 percent. Pension funds, especially (because of their liability characteristics), should invest more in bonds given these estimates.

Avramov, Doron, and Tarun Chordia. 2006. "Predicting Stock Returns." *Journal of Financial Economics*, vol. 82, no. 2 (November):387–415. [added April 2008; abstract by Luis Garcia-Feijoo, CFA]

The authors construct optimal portfolios that allow for company-level equity expected returns, variances, and covariances to vary conditional on a set of macroeconomic variables. Predictability-based investments outperform static and dynamic investments in the market, the Fama–French plus momentum factors, and strategies that invest in stocks with similar size, book-to-market, and prior return characteristics. Returns on individual stocks are predictable out-of-sample because of alpha variation, not because of equity premium predictability.

Bansal, Ravi, and Amir Yaron. 2004. "Risk for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *Journal of Finance*, vol. 59, no. 4 (August):1481–1509.

This article presents a model that can explain the equity risk premium. Dividend and, thus, consumption growth are assumed to consist of two components: a small persistent expected growth rate component and a time-varying economic uncertainty component. The authors show that the historical equity risk premium can be quantitatively justified by the model using a risk aversion parameter of 7.5 to 10.

Barberis, Nicholas, and Ming Huang. 2006. "The Loss Aversion/Narrow Framing Approach to the Equity Premium Puzzle." In *Handbook of Investments: Equity Risk Premium*. Edited by Rajnish Mehra. Amsterdam: North Holland.

The authors review the behavioral approach to understanding the ERP puzzle. The key elements of this approach are loss aversion and narrow framing, two well-known features of decision making under risk in experimental settings. By incorporating these features into traditional utility functions, Barberis and Huang show that a large equity premium and a low and stable risk-free rate can be generated simultaneously, even when consumption growth is smooth and only weakly correlated with the stock market.

Barberis, Nicholas, Ming Huang, and Tano Santos. 2001. "Prospect Theory and Asset Prices." *Quarterly Journal of Economics*, vol. 116, no. 1 (February):1–53.

This paper proposes a new approach for pricing assets by incorporating two psychological ideas into the traditional consumption-based model. Investors are assumed to be more sensitive to losses than to gains, and their risk aversion changes over time depending on their prior investment outcomes. The authors show that this framework can help explain the high historical equity risk premium.

Barro, Robert. 2006. "Rare Disasters and Asset Markets in the Twentieth Century." *Quarterly Journal of Economics*, vol. 121, no. 3 (August):823–866.

This paper extends the analysis of Rietz (1988) and argues that it does provide a plausible resolution of the ERP puzzle. The author suggests that the rare-disasters framework (i.e., the allowance for low-probability disasters proposed by Rietz) can explain the ERP puzzle while "maintaining the tractable framework of a representative agent, time-additive and iso-elastic preferences, and complete markets" (p. 823). These technical terms refer to assumptions that are embedded in Mehra and Prescott (1985) and that are considered standard in general equilibrium or macroeconomic models.

Benartzi, Shlomo, and Richard H. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics*, vol. 110, no. 1 (February):73–92.

This article proposes an explanation for the equity premium based on two concepts from the psychology of decision making. The first concept is called "loss aversion," meaning that investors are more sensitive to losses than to gains. The second concept is called "mental accounting," which points out that investors mentally separate their portfolios into subportfolios for which they have quite different utility functions or risk aversion parameters. For example, investors may have one set of portfolios that they never evaluate and another set that they evaluate every day. Benartzi and Thaler show that the size of the historical equity premium can be explained if investors evaluate their portfolio at least annually.

Bernstein, Peter L. 1997. "What Rate of Return Can You Reasonably Expect... or What Can the Long Run Tell Us about the Short Run?" *Financial Analysts Journal*, vol. 53, no. 2 (March/April):20–28.

By studying historical intervals when stock valuation (P/D or P/E) was the same at the end of the interval as at the beginning, one can avoid incorporating unexpected valuation changes into long-term rate of return studies. The analysis gives an equity risk premium of 3 percent, although the more interesting finding is that equity returns are mean-reverting whereas bond returns have no mean to which to regress. Thus, in the very long run and in real terms, stocks are safer than bonds.

Blanchard, Olivier J., Robert Shiller, and Jeremy J. Siegel. 1993. "Movements in the Equity Premium." *Brookings Papers on Economic Activity*, no. 2:75–138.

The authors show that the expected equity premium has gone steadily down since the 1950s from an unusually high level in the late 1930s and 1940s. Blanchard et al. show the positive relation between inflation and the equity premium, and they conclude that the equity premium is expected to stay at its current level of 2–3 percent if inflation remains low. Implications of this forecast for the macroeconomy are explored.

Brav, Alon, George M. Constantinides, and Christopher C. Geczy. 2002. "Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence." *Journal of Political Economy*, vol. 110, no. 4 (August):793–824.

This paper shows that the equity risk premium can be explained with a stochastic discount factor (SDF) calculated as the weighted average of the individual households' marginal rate of substitution. Important components of the SDF are cross-section variance and skewness of the households' consumption growth rates.

Brown, Stephen J., William N. Goetzmann, and Stephen A. Ross. 1995. "Survival." *Journal of Finance*, vol. 50, no. 3 (July):853–873.

This paper suggests that survival could induce a substantial spurious equity premium and at least partially explain the equity premium puzzle documented by Mehra and Prescott (1985). (That is, to explain it away, because the returns used to frame the "puzzle" were neither expected nor were they achieved by many investors.)

Campbell, John Y., Peter A. Diamond, and John B. Shoven. 2001. "Estimating the Real Rate of Return on Stocks over the Long Term." Social Security Advisory Board. (www.ssab.gov/Publications/Financing/estimated%20rate%20of%20return.pdf)

This collection of papers presented to the Social Security Advisory Board explores expected equity rates of return for the purpose of assessing proposals to invest Social Security assets in the stock market.

Under certain stringent conditions, the earnings-to-price ratio (E/P) is an unbiased estimator of the expected equity return. Noting that earnings are highly cyclical, Campbell, in "Forecasting U.S. Equity Returns in the 21st Century," produces a more stable numerator for E/P by taking the 10-year trailing average of real earnings, E* (after Graham and Dodd;[10] see also Campbell and Shiller 1998, Shiller 2000, and Asness[11]). From this perspective, current data suggest that the structural equity risk premium is now close to zero or that prices will fall, causing the equity risk premium to rise to a positive number. A little of each is the most likely outcome. Departing from the steady-state assumptions used to equate E/P with the expected equity return and using a macroeconomic growth forecast and sensible assumptions about the division, by investors, of corporate risk between equities and bonds, a real interest rate of 3–3.5 percent is forecast, along with an equity risk premium of 1.5–2.5 percent geometric (3–4 percent arithmetic).

In "What Stock Market Returns to Expect for the Future?" Diamond explores the implications of an assumed 7 percent real rate of return on equities. Stocks cannot earn a real total return of 7 percent or else they will have a market capitalization of 39.5 times U.S. GDP by the year 2075 (assuming a 2 percent dividend-plus-share-buyback yield). In contrast, the current capitalization/GDP ratio is 1.5. Changing the GDP growth rate within realistic bounds does not change the answer much. To justify a real total return of 7 percent, stocks must fall by 53 percent in real terms over the next 10 years (assuming a 2 percent dividend yield). Increasing the dividend payout does reduce the projected capitalization/GDP ratio materially, but in no case does it reduce the ratio below 7.86 in 2075.

In "What Are Reasonable Long-Run Rates of Return to Expect on Equities?" Shoven examines what is likely to happen to rates of return over the next 75 years. Dividends are irrelevant, because of tax policy; what counts is total cash flow to the investor. In a steady state, the expected return on equities (per share) equals the dividend yield, plus the share buyback yield, plus the growth rate of macroeconomic aggregates. This analysis produces an expected real total return on equities of 6.125 percent (say, 6–6.5 percent). Because of high (3 percent) real rates as projected—not the very high, current TIPS yield—the equity risk premium is only 3–3.5 percent, but these projections require one to reduce the 7 percent real equity return projection used by the Social Security Advisory Board only a little. At a P/E of 15, the real equity return projection would be a little better than 7 percent.

Campbell, John Y., and Robert J. Shiller. 1998. "Valuation Ratios and the Long-Run Stock Market Outlook." *Journal of Portfolio Management*, vol. 28, no. 2 (Winter):11–26. (Updated in Cowles Foundation Discussion Paper #1295, Yale University, March 2001.)

The dividend-to-price ratio (D/P) can forecast either changes in dividend, which is what efficient market theory suggests, or changes in price, or both. Empirically, it forecasts only changes in price. At the current D/P, the forecast is extraordinarily bearish: The stock market will lose about two-thirds of its real value. The forecast becomes less drastically bearish (although still quite bearish) when one uses (dividend + share buybacks), earnings, the 10-year moving average of earnings in constant dollars, or other variables in the denominator. Real stock returns close to zero over the next 10 years are forecast. A number of statistical weaknesses in the analysis are acknowledged: The historical observations are not independent, and the analysis depends on valuation ratios regressing to their historical means, whereas the actual means are not known and could conceivably lie outside the historical range.

The 2001 update reaches the same conclusion and an even more bearish forecast.

[10] Benjamin Graham and David Dodd, *Security Analysis* (New York: McGraw-Hill, 1934).

[11] Clifford S. Asness, "Stocks versus Bonds: Explaining the Equity Risk Premium," *Financial Analysts Journal*, vol. 56, no. 2 (March/ April 2000):96–113.

Carhart, Mark M., and Kurt Winkelmann. 2003. "The Equity Risk Premium." In *Modern Investment Management*. Edited by William N. Goetzmann and Roger G. Ibbotson. Hoboken, NJ: John Wiley & Sons:44–54.

Historical perspective and an equilibrium estimate of the equity risk premium are discussed. The authors estimate that the U.S. corporate bond yield above Treasury bonds is 2.25 percent, and the expected U.S. corporate bond risk premium is thus 1.5 percent after subtracting an expected default loss of 0.75 percent. This amount (1.5 percent) is considered to be the lower bound of the current equity risk premium. Because equity volatility is two or three times higher than that of corporate bonds, the authors "cautiously" suggest an equity risk premium of 3 percent or higher.

Claus, James, and Jacob Thomas. 2001. "Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets." *Journal of Finance*, vol. 56, no. 5 (October):1629–1666.

The Ibbotson or historical-extrapolation method gives ERP estimates that are much too high, relative to both purely utility-based estimates (Mehra and Prescott 1985) and estimates based on valuation (for example, Campbell and Shiller 1998). Estimates of the equity risk premium were calculated for each year since 1985 by subtracting the 10-year risk-free rate from the discount rate that equates U.S. stock market valuations with forecasted future flows, and results suggest that the equity risk premium is probably no more than 3 percent. International evidence from Canada, France, Germany, Japan, and the United Kingdom also support this claim. Known upward biases in analysts' earnings forecasts are corrected in making the estimates. Possible reasons why the historical method might have overstated the expected equity risk premium in recent years are discussed.

Cochrane, John H. 1997. "Where Is the Market Going? Uncertain Facts and Novel Theories." *Economic Perspectives*, Federal Reserve Bank of Chicago, vol. 21, no. 6 (November/December):3–37.

This paper summarizes the statistical evidence on average stock return and surveys economic theories that try to explain it. Standard models can only justify a low equity risk premium, whereas new models that can explain the 8 percent historical equity premium drastically modify the description of stock market risk. The author concludes that low forecast stock returns do not imply that the investor should change his portfolio unless he is different from the average investor in risk exposure, attitude, or information.

Constantinides, George M. 1990. "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy*, vol. 98, no. 3 (June):519–543.

Constantinides introduces habit persistence in an effort to explain the ERP puzzle. This model assumes that an investor's utility is affected by both current and past consumption and that a small drop in consumption can generate a large drop in consumption net of the subsistence level. The author shows that the historical equity premium can be explained if past consumption generates a subsistence level of consumption that is about 80 percent of the normal consumption rate.

———. 2002. "Rational Asset Prices." *Journal of Finance*, vol. 57, no. 4 (August):1567–1591.

This article examines the extent to which historical asset returns can be explained by relaxing the assumptions of the traditional asset pricing model. Constantinides reviews statistical evidence on historical equity returns and premiums and discusses the limitations of existing theories. The author suggests that it is promising to try to explain the equity risk premium by integrating the notions of incomplete market, life-cycle issues, borrowing constraints, and limited stock participation (i.e., stockholdings are concentrated in the hands of the wealthiest few), along with investors' deviation from rationality.

Constantinides, George M., John B. Donaldson, and Rajnish Mehra. 2002. "Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle." *Quarterly Journal of Economics*, vol. 117, no. 1 (February):269–296.

As the correlation of equities with personal income changes over the life of the investor, so does the attractiveness of equities to that investor. The young, who should borrow to smooth consumption and to invest in equities, can't do so. Therefore, equities are priced almost exclusively by middle-aged investors, who find equities to be unattractive. (Middle-aged investors have a shorter time horizon and also prefer bonds because they smooth consumption in retirement, as wages do when one is working.) The result is a decreased demand for equities and an increased demand for bonds relative to what it would be in a perfectly competitive market. Thus, equities are (on average, over time) underpriced and bonds are overpriced, producing a higher equity risk premium than predicted by Mehra and Prescott (1985).

Cornell, Bradford. 1999. *The Equity Risk Premium*. New York: Wiley.

The literature on the equity risk premium is extensively reviewed and somewhat popularized in this book. The conclusion is that the equity risk premium will be lower in the future than it was in the past. A premium of 3.5–5.5 percent over Treasury bonds and 5–7 percent over Treasury bills is projected.

Dichev, Ilia D. 2007. "What Are Stock Investors' Actual Historical Returns? Evidence from Dollar-Weighted Returns." *American Economic Review*, vol. 97, no. 1 (March):386–401. [added April 2008, abstract by Bruce D. Phelps, CFA]

For the NYSE and Amex, the author finds that dollar-weighted returns are 1.9 percent per year lower on average than value-weighted (or buy-and-hold) returns. For the NASDAQ, dollar-weighted returns are 5.3 percent lower. Similar results hold internationally. Because actual investor returns are lower than published returns, empirical measurements of the equity risk premium and companies' cost of equity are potentially overstated.

Diermeier, Jeffrey J., Roger G. Ibbotson, and Laurence B. Siegel. 1984. "The Supply of Capital Market Returns." *Financial Analysts Journal*, vol. 40, no. 2 (March/April):74–80.

Stock total returns must equal dividend yields plus the growth rate of dividends, which cannot, in the long run, exceed the growth rate of the economy. If infinite-run expected dividend growth exceeded infinite-run expected economic growth, then dividends would crowd out all other economic claims. Net new issues, representing new capital (transferred from the labor market) that is needed so the corporate sector can grow, may cause the dividend growth rate to be slower than the GDP growth rate. Thus, the equity risk premium equals the dividend yield (minus new issues net of share buybacks), plus the GDP growth rate, minus the riskless rate.

As far as we know, this is the first direct application of the dividend discount model of John Burr Williams (writing in the 1930s) and Myron Gordon and Eli Shapiro (in the 1950s) to the question of the equity risk premium for the whole equity market as opposed to an individual company. The "supply side" thread thus begins with this work.

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

This book provides a comprehensive examination of returns on stocks, bonds, bills, inflation, and currencies for 16 countries over the period from 1900 to 2000. This evidence suggests that the high historical equity premium obtained for the United States is comparable with that of other countries. The point estimate of the historical equity premium for the United States and the United Kingdom is about 1.5 percent lower than reported in previous studies, and the authors attribute the difference to index construction bias (for the United Kingdom) and a longer time frame (for the United States). The prospective risk premium that investors can expect going forward is also discussed. The estimated geometric mean premium for the United States is 4.1 percent, 2.4 percent for the United Kingdom, and 3.0 percent for the 16-country world index. Implications for individual investors, investment institutions, and companies are carefully explored.

———. 2003. "Global Evidence on the Equity Risk Premium." *Journal of Applied Corporate Finance*, vol. 15, no. 4 (Summer):27–38.

This article examines the historical equity risk premium for 16 countries using data from 1900 to 2002. The geometric mean annualized equity risk premium for the United States was 5.3 percent, and the average risk premium across the 16 countries was 4.5 percent. The forward-looking risk premium for the world's major markets is likely to be around 3 percent on a geometric mean basis and about 5 percent on an arithmetic mean basis.

———. 2006. "The Worldwide Equity Premium: A Smaller Puzzle." Working paper.

This paper is an updated version of Dimson, Marsh, and Staunton (2003). Using 1900–2005 data for 17 countries, the authors show that the annualized equity premium for the rest of the world was 4.2 percent, not too much below the U.S. equity premium of 5.5 percent over the same period.

The historical equity premium is decomposed into dividend growth, multiple expansion, the dividend yield, and changes in the real exchange rate. Assuming zero change in the real exchange rate and no multiple expansion, and a dividend yield 0.5–1 percent lower than the historical mean (4.49 percent), the authors forecast a geometric equity premium on the world index around 3–3.5 percent and 4.5–5 percent on an arithmetic mean basis.

Elton, Edwin J. 1999. "Presidential Address: Expected Return, Realized Return and Asset Pricing Tests." *Journal of Finance*, vol. 54, no. 4 (August):1199–1220.

At one time, researchers felt they had to (weakly) defend the assumption that expected returns were equal to realized returns. Now, they just make the assumption without defending it. This practice embeds the assumption that information surprises cancel to zero; evidence, however, shows they do not. The implications of this critique are applied to asset-pricing tests, not to the equity risk premium.

Fama, Eugene F., and Kenneth R. French. 1999. "The Corporate Cost of Capital and the Return on Corporate Investment." *Journal of Finance*, vol. 54, no. 6 (December):1939–1967.

The authors use Compustat data to estimate the internal rate of return (IRR) of the capitalization-weighted corporate sector from 1950 to 1996. This IRR, 10.72 percent, is assumed to have been the nominal weighted average cost of capital (WACC). By observing the capital structure and assuming a corporate debt yield 150 bps above Treasuries, and making the usual tax adjustment to the cost of debt, a nominal expected equity total return of 12.8 percent is derived, which produces an equity risk premium of 6.5 percent. The cash flow from the "sale" of securities in 1996 is a large proportion of the total cash flow studied, so the sensitivity of the result to the 1996 valuation is analyzed. Because the period studied is long, the result is not particularly sensitive to the exit price.

———. 2002. "The Equity Premium." *Journal of Finance*, vol. 57, no. 2 (April):637–659.

This paper compares alternative estimates of the unconditional expected stock return between 1872 and 2000, and provides explanation to the low expected return estimates derived from fundamentals such as dividends and earnings for the 1951–2000 period. The authors conclude that the decline in discount rates largely causes the unexplained capital gain of the last half-century.

Faugère, Christophe, and Julian Van Erlach. 2006. "The Equity Premium: Consistent with GDP Growth and Portfolio." *Financial Review*, vol. 41, no. 4 (November):547–564. [added April 2008; abstract by Stephen Phillip Huffman, CFA]

Two macroeconomic equity premium models are derived and tested for consistency with historical data. The first model illustrates that the long-term equity premium is directly related to per capita growth in GDP. The second model, based on a portfolio insurance strategy of buying put options, illustrates that debtholders are paying stockholders an insurance premium, which is essentially the equity premium.

Fisher, Lawrence, and James H. Lorie. 1964. "Rates of Return on Investments in Common Stocks." *Journal of Business*, vol. 37, no. 1 (January):1–21.

This paper presents the first comprehensive data on rates of return on investments in common stocks listed on New York Stock Exchange over the period from 1926 to 1960. The authors show that the annually compounded stock return was 9 percent with reinvestment of dividend for tax-exempt institutions during this period.

Geweke, John. 2001. "A Note on Some Limitations of CRRA Utility." *Economic Letters*, vol. 71, no. 3 (June): 341–345.

This paper points out that the equity premium calculated from the standard growth model in Mehra and Prescott (1985) is quite sensitive to small changes in distribution assumptions. As such, it is questionable to use this kind of growth model to interpret observed economic behavior.

Goyal, Amit, and Ivo Welch. 2006. "A Comprehensive Look at the Empirical Performance of Equity Premium Prediction." Working paper.

This paper examines a wide range of variables that have been proposed by economists to predict the equity premium. The authors find that the prediction models have failed both in sample and out of sample using data from 1975 to 2004 and that out-of-sample predictions of the models are unexpectedly poor. They conclude that "the models would not have helped an investor with access only to the information available at the time to time the market" (p. 1).

Grinold, Richard, and Kenneth Kroner. 2002. "The Equity Risk Premium." *Investment Insights*, Barclays Global Investors, vol. 5, no. 3 (July):1–24.

The authors examine the four components of the expected equity risk premium separately (income return, expected real earnings growth, expected inflation, and expected repricing) and suggest a current risk premium of about 2.5 percent. The authors argue that neither the "rational exuberance" view (5.5 percent equity risk premium) and "risk premium is dead" (zero or negative premium) view can be justified without making extreme and/or irrational assumptions.

The authors also forcefully attack the "puzzle" literature by arguing that literature on the equity risk premium puzzle is too academic and is dependent on unrealistic asset-pricing models.

Ibbotson, Roger G., and Peng Chen. 2003. "Long-Run Stock Returns: Participating in the Real Economy." *Financial Analysts Journal*, vol. 59, no. 1 (January/February):88–98.

If one simply uses the dividend discount model to forecast stock returns, the forecast violates M&M dividend invariance because the current dividend yield is much lower than the average dividend yield over the period from which historical earnings growth rates were taken. Applying M&M intertemporally, lower dividend payouts should result in higher earnings growth rates. The solution is to add, to the straight dividend discount model estimate, an additional-growth term of 2.28 percent *as well as* using a current-dividend number of 2.05 percent, which is what the dividend yield would have been in 2000 if the dividend payout ratio had equaled the historical average of 59.2 percent. The equity risk premium thus estimated is about 4 percent (geometric) or 6 percent (arithmetic), about 1.25 percent lower than the straight historical estimate.

Ibbotson, Roger G., and Rex A. Sinquefield, 1976. "Stocks, Bonds, Bills and Inflation: Year-by-Year Historical Returns (1926–74)." *Journal of Business*, vol. 49, no. 1 (January):11–47. (Updated in *Stocks, Bonds, Bills and Inflation: 2006 Yearbook*; Chicago: Morningstar, 2006.)

Total equity returns consist of a stationary part (the equity risk premium) and a nonstationary part (the interest rate component, which consists of a real interest rate plus compensation for expected inflation). The estimator of the future arithmetic mean equity risk premium is the past arithmetic

mean premium, which is currently about 7 percent. To this is added the current interest rate, 4.8 percent (on 20-year Treasury bonds). The sum of these, about 12 percent, is the arithmetic mean expected total return on equities. This method is justified by the assertion that in the long run, investors should and do conform their expectations to what is actually realizable. As a result, the historical equity risk premium reflects equilibrium at all times and forms the proper estimator of the future equity risk premium. (Note that the 2006 update discusses other methods rather than supporting a doctrinaire "future equals past" interpretation of historical data.)

Jagannathan, Ravi, Ellen R. McGrattan, and Anna Scherbina. 2000. "The Declining U.S. Equity Premium." *Quarterly Review*, Federal Reserve Bank of Minneapolis, vol. 24, no. 4 (Fall):3–19.

The IRR equating expected future dividends from a stock portfolio with the current price is the expected total return on equities; subtracting the bond yield, one arrives at the equity risk premium. This number is estimated at historical points in time and is shown to have declined over the sample period (1926–1999). The expected total return on equities is about the same in the 1990s as it was in the 1960s, but the equity risk premium is smaller because bond yields have increased. The equity risk premium in 1999 is –0.27 percent for the S&P 500, –0.05 percent for the "CRSP portfolio," and 2.71 percent for the "Board of Governors stock portfolio" (a broad-cap portfolio with many small stocks that pay high dividend yields). The analysis is shown to be reasonably robust when tested for sensitivity to the dividend yield being too low because of share repurchases and the bond yield being too high. If dividend growth is assumed equal to GNP growth, instead of being 1.53 percentage points lower as it was historically, then the equity risk premium based on the S&P 500 rises to 1.26 percent.

Jorion, Philippe, and William N. Goetzmann. 1999. "Global Stock Markets in the Twentieth Century." *Journal of Finance*, vol. 54, no. 3 (June):953–980.

The U.S. equity market experience in the 20th century is an unrepresentative sample of what can and does happen. The high equity risk premium observed globally is mostly a result of high equity returns in the United States (with a 4.3 percent real capital appreciation return), which had a large initial weight in the GDP-weighted world index. All other surviving countries had lower returns (with a median real capital appreciation return of 0.8 percent), and there were many nonsurviving countries. Although the large capitalization of the United States was in a sense the market's forecast of continued success, investors did not know in advance that they would be in the highest-returning country or even in a surviving one. Nonsurvival or survival with poor returns should be factored in when reconstructing the history of investor expectations (and should conceivably be factored into current expectations too). This finding contrasts with that of Dimson, Marsh, and Staunton (2002, 2003, 2006).

Kocherlakota, Narayana R. 1996. "The Equity Premium: It Is Still a Puzzle." *Journal of Economic Literature*, vol. 34, no. 1 (March):42–71.

After reviewing the literature on modifications of investor risk preference and on market friction, the author suggests that the ERP puzzle is still unsolved. Kocherlakota concludes that the equity risk premium puzzle should be solved by discovering the fundamental features of goods and asset markets rather than patching existing models.

Kritzman, Mark P. 2001. "The Equity Risk Premium Puzzle: Is It Misspecification of Risk?" *Economics and Portfolio Strategy* (15 March), Peter L. Bernstein, Inc.

Investors do not know when they are going to need their money back (for consumption), so the terminal-wealth criterion used by Mehra and Prescott (1985) to frame the ERP puzzle greatly understates the risk of equities (but not of bonds). In addition, some investors face risk from "breaching a threshold" that is not captured by classical utility theory. Thus, a much higher equity risk premium is justified by utility theory than is proposed by Mehra and Prescott.

Longstaff, Francis A., and Monika Piazzesi. 2004. "Corporate Earnings and the Equity Premium." *Journal of Financial Economics*, vol. 74, no. 3 (December):401–421.

Most studies assume that aggregate dividends equal aggregate consumption. This article argues that separating corporate cash flow from aggregate consumption is critical because "corporate cash flows have historically been far more volatile and sensitive to economic shocks than has aggregate consumption" (p. 402). The authors show that the equity premium consists of three components, identified by allowing aggregate dividends and consumption to follow distinct dynamic processes. The first component is called the consumer-risk premium, which is the Mehra and Prescott (1985) equity risk premium proportional to the variance of consumption growth. The second component is the event-risk premium, which compensates for downward jumps. And the third component is the corporate-risk premium, which is proportional to the covariance between the consumption growth rate and the "corporate fraction" (defined as the ratio of aggregate dividends to consumption). Using a risk aversion parameter of 5, the three components are 0.36 percent, 0.51 percent, and 1.39 percent, summing to a total equity premium of 2.26 percent. The authors admit that their model does not solve the ERP puzzle completely and suggest that the ultimate resolution may lie in the integration of their model with other elements, such as habit formation or investor heterogeneity in incomplete markets.

Lundblad, Christian. 2007. "The Risk Return Tradeoff in the Long Run: 1836–2003." *Journal of Financial Economics*, vol. 85, no. 1 (July):123–150. [added April 2008; abstract by Yazann S. Romahi, CFA]

Although the risk–return trade-off is fundamental to finance, the empirical literature has offered mixed results. The author extends the sample considerably and analyzes nearly two centuries of both U.S. and U.K. market returns and finds a positive and statistically significant risk–return trade-off in line with the postulated theory.

Mankiw, N. Gregory. 1986. "The Equity Premium and the Concentration of Aggregate Shocks." *Journal of Financial Economics*, vol. 17, no. 1 (September):211–219.

This article shows that one cannot judge the appropriateness of the equity premium from aggregate data alone, as Mehra and Prescott (1985) did. In an economy where aggregate shocks are not dispersed equally throughout the population, the equity premium depends on the concentrations of these aggregate shocks in particular investors and can be made arbitrarily large by making the shock more and more concentrated.

Mankiw, N. Gregory, and Stephen P. Zeldes. 1991. "The Consumption of Stockholders and Non-Stockholders." *Journal of Financial Economics*, vol. 29, no. 1 (March):97–112.

This article examines whether the consumption of stockholders differs from that of nonstockholders and whether this difference helps to explain the historical equity risk premium. It shows that aggregate consumption of stockholders is more highly correlated with the stock market and is more volatile than the consumption of nonstockholders. A risk aversion parameter of 6 (relative to the magnitude of 30–40 in Mehra and Prescott 1985) can explain the size of the equity premium based on consumption of stockholders alone.

McGrattan, Ellen R., and Edward C. Prescott. 2000. "Is the Stock Market Overvalued?" *Quarterly Review*, Federal Reserve Bank of Minneapolis (Fall):20–40.

Standard macroeconomic growth theory (Cobb–Douglas, etc.) is used to value the corporate sector in the United States. The current capitalization-to-GDP ratio of 1.8 is justified, so the market is not overvalued. "[T]heory . . . predicts that the real returns on debt and equity should both be near 4 percent" (p. 26). Thus, the predicted equity risk premium is small.

———. 2001. "Taxes, Regulations, and Asset Prices." NBER Working Paper #8623.

This paper shows that the large run-up in equity value relative to GDP between 1962 and 2000 is mainly caused by (1) large reductions in individual tax rates, (2) increased opportunities to hold equity in a nontaxed pension plan, and (3) increases in intangible and foreign capital. The authors argue that the high equity risk premium documented by Mehra and Prescott (1985) is not puzzling after these three factors are accounted for. However, in the future, one should expect no further gains from tax policy; the currently expected real return on equities is about 4 percent, down from 8 percent in the early postwar period.

———. 2003. "Average Debt and Equity Returns: Puzzling?" *American Economic Review*, vol. 93, no. 2 (May):392–397.

This article shows that the realized equity premium in the last century was less than 1 percent after accounting for taxes, regulations, and diversification costs. The authors also argue that Treasury bills "provide considerable liquidity services and are a negligible part of individuals' long-term debt holdings" (p. 393). Long-term savings instruments replace short-term government debt in their equity premium calculation.

Mehra, Rajnish. 2003. "The Equity Premium: Why Is It a Puzzle?" *Financial Analysts Journal*, vol. 59, no. 1 (January/February):54–69.

The ERP puzzle literature is easily misunderstood because of its difficulty. Here, the puzzle is stated in language that is accessible to most finance practitioners. First, empirical facts regarding the returns and risks of major asset classes are presented. Then, the theory responsible for the "puzzle" is summarized. Modern asset pricing theory assumes that economic agents pursue and, on average, get fair deals. When one follows this line of reasoning to its conclusion, using the tools of classic growth and real business cycle theory, an equity risk premium of at most 1 percent emerges. An extensive discussion reveals why this is the case and addresses various attempts made by other authors to resolve the puzzle.

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

In this seminal work, Mehra and Prescott first document the "equity premium puzzle" using a consumption-based asset-pricing model in which the quantity of risk is defined as the covariance of excess stock return with consumption growth and the price of risk is the coefficient of relative risk aversion. Because of the low risk resulting from the smooth historical growth of consumption, the 6 percent equity risk premium in the 1889–1978 period can only be explained by a very high coefficient of risk aversion in the magnitude of 30 to 40. Risk aversion parameters observed in other aspects of financial behavior are around 1. Such a risk aversion parameter is consistent with at most a 1 percent equity risk premium, and possibly one as small as 0.25 percent.

Note that Mehra and Prescott assumed that consumption was equal to aggregate dividends. Because consumption is very smooth and dividends are not as smooth, this comparison may be troublesome.

Philips, Thomas K. 1999. "Why Do Valuation Ratios Forecast Long-Run Equity Returns?" *Journal of Portfolio Management*, vol. 25, no. 3 (Spring):39–44.

In this article, the Edwards–Bell–Ohlson equation,

$$P_0 = B_0 + \sum_{i=1}^{\infty} \left\{ \frac{E\left[\left(ROE_i - r\right) B_{i-1}\right]}{\left(1+r\right)} \right\},$$

where *P* is price, *B* is book value, *ROE* is return on book equity, *r* is the expected return on equity, and *i* is the time increment, is first used to derive closed-form expressions for the expected return on equities, stated in terms of both dividends and earnings. Then, the GDP growth rate is introduced as an indicator of earnings growth. Share repurchases are considered to be a part of dividends. This setup leads to the following conclusions: (1) The expected return increases monotonically with book-to-price ratio (B/P), E/P, and D/P; (2) if a corporation's return on equity equals its cost of capital (expected return), then its price-to-book ratio (P/B) should be 1 and its expected return should equal E/P. The analysis suggests that nominal total expected equity returns shrank from almost 14 percent in 1982 to 6.5 percent in 1999 (a larger decline than can be explained by decreases in unanticipated inflation). This decrease in expected return was accompanied by very high concurrent actual returns that were misread by investors as evidence of an *increase* in the expected return. Going forward, investors will not get an increased return.

Rietz, Thomas A. 1988. "The Equity Risk Premium: A Solution." *Journal of Monetary Economics*, vol. 22, no. 1 (July):117–131.

Rietz suggests that the ERP puzzle can be solved by incorporating a very small probability of a very large drop in consumption. In such a scenario, the risk-free rate is much lower than the equity return. In an article published in the same issue, Mehra and Prescott argued that Rietz's model requires a 1 in 100 chance of a 25 percent decline in consumption to reconcile the equity premium with a risk aversion parameter of 10. However, the author says, the largest consumption decline in the last 100 years was only 8.8 percent. Campbell, Lo, and MacKinlay (see Note 3) point out that "the difficulty with Rietz's argument is that it requires not only an economic catastrophe, but one which affects stock market investors more seriously than investors in the short-term debt instruments" (p. 311).

But during the Great Depression, the stock market fell by 86 percent from peak to trough and dividends fell by about half; consumption by stockholders over that period thus probably fell by much more than 8.8 percent. Aggregate consumption at that time included many lower-income people, especially farmers, whose consumption was not directly affected by falling stock prices.

Shiller, Robert J. 2000. *Irrational Exuberance*. Princeton, NJ: Princeton University Press.

This influential book provides a wealth of historical detail on the equity risk premium. Using 10 years of trailing real earnings (see, originally, Graham and Dodd) to estimate normalized P/Es, Shiller concludes that the market is not only overpriced but well outside the range established by previous periods of high stock prices.

Siegel, Jeremy J. 1999. "The Shrinking Equity Premium." *Journal of Portfolio Management*, vol. 26, no. 1 (Fall):10–19.

In contrast to Siegel (2002), analysis of dividend and earnings multiples suggests a real return (not an equity risk premium) of only 3.1–3.7 percent for stocks, lower than the then-current real TIPS yield. Although then-current high prices suggest higher-than-historical earnings growth, investors are likely to realize lower returns than in the past. (Incidentally, past achieved returns are lower than index returns because of transaction costs and lack of diversification.) On the positive side, the Jorion and Goetzmann (1999) finding that world markets returned a real capital gain of only 0.8 percent from 1921 to the present, compared with 4.3 percent in the United States, is misstated because the analysis is of the median portfolio, not the average. The GDP-weighted average is only 0.28 percent short of the U.S. return and is higher than the U.S. return if converted to dollars (although Jorion and Goetzmann point out that the large initial size of the United States causes the annualized world index return to lie within 1 percent of the U.S. return by construction).

———. 2002. *Stocks for the Long Run*. 3rd ed. New York: McGraw-Hill.

Siegel argues for a U.S. equity risk premium of 2–3 percent, about half of the historic equity risk premium. He expects a future real return on equity of about 6 percent, justified by several positive factors. Siegel considers an equity risk premium as low as 1 percent but clearly sees that stocks must yield more than inflation-indexed bond yields (3.5 percent at the time of the book). He turns to earnings yield arguments to answer the question of how much more. A Tobin's *q* greater than 1 in 2001 leads Siegel to see the earnings yield as understated. In addition, the overinvestment in many technology companies led to a drop in the cost of productivity-enhancing investments, which allows companies to buy back shares or raise dividends. In technology, an excess supply of capital, overbuilding, and a subsequent price collapse provide a technological base to benefit the economy and future shareholder returns. Also, the United States is still seen as an entrepreneurial nation to attract a growing flow of investment funds seeking a safe haven, leading to higher equity prices. Furthermore, short-run room for growth in corporate profits is another positive factor for future real return enhancement.

———. 2005. "Perspectives on the Equity Risk Premium." *Financial Analysts Journal*, vol. 61, no. 6 (November/December):61–73.

This article reviews and discusses the ERP literature as follows: (1) a summary of data used in equity premium calculation and their potential biases, (2) a discussion of academic attempts to find models to fit the data, (3) the practical applications of some proposed models, and (4) a discussion of the future equity risk premium.

Siegel, Jeremy J., and Richard H. Thaler. 1997. "Anomalies: The Equity Premium Puzzle." *Journal of Economic Perspectives*, vol. 11, no. 1 (Winter):191–200.

Proposed resolutions of the ERP puzzle fall into two categories: (1) observations that the stock market is riskier, or the equity risk premium is smaller, than generally thought, and (2) different theoretical frameworks that would make the observed risk aversion rational. Neither approach has been "completely successful" in explaining why, if stocks are so rewarding, investors don't hold more of them.

Weil, Philippe. 1989. "The Equity Premium Puzzle and the Risk-Free Rate Puzzle." *Journal of Monetary Economics*, vol. 24, no. 3 (November):401–421.

A critique of the power utility function used by Mehra and Prescott (1985) is the tight link between risk aversion and intertemporal substitution. This article shows that the ERP puzzle cannot be solved by simply separating risk aversion for intertemporal substitution.

Weitzman, Martin L. "Prior-Sensitive Expectations and Asset-Return Puzzles." Forthcoming. *American Economic Review*.

This article presents one unified Bayesian theory that explains the ERP puzzle, risk-free rate puzzle, and excess volatility puzzle. The author shows that Bayesian updating of unknown structural parameters introduces a permanent thick tail to posterior expectation that can account for, and even reverse, major asset-return puzzles.

Welch, Ivo. 2000. "Views of Financial Economists on the Equity Premium and Professional Controversies." *Journal of Business*, vol. 73, no. 4 (October):501–537.

This paper presents the results of a comprehensive survey of 226 financial economists. The main findings are: (1) the average arithmetic 30-year equity premium forecast is about 7 percent; (2) short-term forecasts are lower than the long-term forecast, in the range of 6–7 percent; (3) economists perceive that their consensus is about 0.5–1 percent higher than it actually is.

———. 2001. "The Equity Premium Consensus Forecast Revisited." Working paper, Yale University.

The equity premium forecast in this 2001 survey declined significantly compared with the 1998 survey. The one-year forecast is 3–3.5 percent, and the 30-year forecast stands at 5–5.5 percent.

I would like to thank Laurence Siegel, research director of the Research Foundation of CFA Institute, for his assistance and for providing much of the foundation for this project with his earlier work on the equity risk premium. I am also grateful to the Research Foundation for financial support.

This publication qualifies for 1 CE credit.

UTILITY STOCKS AND THE SIZE EFFECT: AN EMPIRICAL ANALYSIS

Annie Wong*

I. Introduction

The objective of this study is to examine whether the firm size effect exists in the public utility industry. Public utilities are regulated by federal, municipal, and state authorities. Every state has a public service commission with board and varying powers. Often their task is to estimate a fair rate of return to a utility's stockholders in order to determine the rates charged by the utility. The legal principles underlying rate regulation are that "the return to the equity owner should be commensurate with returns on investments in other enterprises having corresponding risks," and that the return to a utility should be sufficient to "attract capital and maintain credit worthiness." However, difficulties arise from the ambiguous interpretation of the legal definition of *fair and reasonable rate of return* to an equity owner.

Some finance researchers have suggested that the Capital Asset Pricing Model (CAPM) should be used in rate regulation because the CAPM beta can serve as a risk measure, thus making risk comparisons possible. This approach is consistent with the spirit of a Supreme Court ruling that equity owners sharing similar level of risk should be compensated by similar rate of return.

The empirical studies of Banz (1981) and Reinganum (1981) showed that small firms tend to earn higher returns than large firms after adjusting for beta. This phenomenon leads to the proposition that firm size is a proxy for omitted risk factors in determining stock returns. Barry and Brown (1984) and Brauer (1986) suggested that the omitted risk factor could be the differential information environment between small and large firms. Their argument is based on the fact that investors often have less publicly available information to assess the future cash flows of small firms than that of large firms. Therefore, an additional risk premium should be included to determine the appropriate rate of return to shareholders of small firms.

The samples used in prior studies are dominated by industrial firms, no one has examined the size effect in public utilities. The objective of this study is to extend the empirical findings of the existing studies by investigating whether the size effect is also present in the utility industry. The findings of this study have important implications for investors, public utility firms, and state regulatory agencies. If the size effect does exist in the utility industry, this would suggest that the size factor should be considered when the CAPM is being used to determine the fair rate of return for public utilities in regulatory proceedings.

II. Information Environment of Public Utilities

In general, utilities differ from industriales in that utilities are heavily regulated and they follow similar accounting procedures. A public utility's financial reporting is mainly regulated by the Securities and Exchange Commission (SEC) and the Federal Energy Regulatory Commission (FERC). Under the Public Utility Holding Company Act of 1935, the SEC is empowered to regulate the holding company systems of electric and gas utilities. The Act requires registration of public utility holding companies with the SEC. Only under strict conditions would the purchase, sale or issuance of securities by these holding companies be permitted. The purpose of the Act is to keep the SEC and investors informed of the financial conditions of these firms. Moreover, the FERC is in charge of the interstate operations of electric and gas companies. It requires utilities to follow the accounting procedures set forth in its Uniform Systems of Accounts. In particular, electric and gas utilities must request their Certified Public Accountants to certify that certain schedules in the financial reports are in conformity with the Commission's accounting requirements. These detailed reports are submitted annually and are open to the public.

*Western Connecticut State University. The author thanks Philip Perry, Robert Hagerman, Eric Press, the anonymous referee, and Clay Singleton for their helpful comments.

The FERC requires public utilities to keep accurate records of revenues, operating costs, depreciation expenses, and investment in plant and equipment. Specific financial accounting standards for these purposes are also issued by the Financial Accounting Standards Board (FASB). Uniformity is required so that utilities are not subject to different accounting regulations in each of the states in which they operate. The ultimate objective is to achieve comparability in financial reporting so that factual matters are not hidden from the public view by accounting flexibility.

Other regulatory reports tend to provide additional financial information about utilities. For example, utilities are required to file the FERC Form No. 1 with the state commission. This form is designed for state commissions to collect financial and operational information about utilities, and serves as a source for statistical reports published by state commissions.

Unlike industriales, a utility's earnings are predetermined to a certain extent. Before allowed earnings requests are approved, a utility's performance is analyzed in depth by the state commission, interest groups, and other witnesses. This process leads to the disclosure of substantial amount of information.

III. Hypothesis and Objective

Due to the Act of 1935, the Uniform Systems of Accounts, the uniform disclosure requirements, and the predetermined earnings, all utilities are reasonably homogeneous with respect to the information available to the public. Barry and Brown (1984) and Brauer (1986) suggested that the difference of risk-adjusted returns between small and large firms is due to their differential information environment. Assuming that the differential information hypothesis is true, then uniformity of information availability among utility firms would suggest that the size effect should not be observed in the public utility industry. The objective of this paper is to provide a test of the size effect in public utilities.

IV. Methodology

1. Sample and Data

To test for the size effect, a sample of public utilities and a sample of industriales matched by equity value are formed so that their results can be compared. Companies in both samples are listed on the Center for Research in Security Prices (CRSP) Daily and Monthly Returns files. The utility sample includes 152 electric and gas companies. For each utility in the sample, two industrial firms with similar firm size (one is slightly larger and the other is slightly smaller than the utility) are selected. Thus, the industrial sample includes 304 non-regulated firms.

The size variable is defined as the natural logarithm of market value of equity at the beginning of each year. Both the equally-weighted and value-weighted CRSP indices are employed as proxies for the market returns. Daily, weekly and monthly returns are used. The Fama-MacBeth (1973) procedure is utilized to examine the relation between risk-adjusted returns and firm size.

2. Research Design

All utilities in the sample are ranked according to the equity size at the beginning of the year, and the distribution is broken down into deciles. Decile one contains the stocks with the lowest market values while decile ten contains those with the highest market values. These portfolios are denoted by MV_1, MV_2, ..., and MV_{10}, respectively.

The combinations of the ten portfolios are updated annually. In the year after a portfolio is formed, equally-weighted portfolio returns are computed by combining the returns of the component stocks within the portfolio. The betas for each portfolio at year t, $\hat{\beta}_{pt}$'s, are estimated by regressing the previous five years of portfolio returns on market returns:

$$\tilde{R}_{pt} = \alpha_p + \hat{\beta}_{pt}\tilde{R}_{mt} + \tilde{U}_{pt} \qquad (1)$$

where

R_{pt} = periodic return in year t on portfolio p

R_{mt} = periodic market return in year t

U_{pt} = disturbance term.

Banz (1981) applied both the ordinary and generalized least squares regressions to estimate β; and concluded that the results are essentially identical (p.8). Since adjusting for heteroscedasticity does not necessarily lead to more efficient estimators, the ordinary least squares procedures are used in this study to estimate β in equation (1).

The following cross-sectional regression is then run for the portfolios to estimate γ_{it}, i = 0, 1, and 2:

$$R_{pt} = \gamma_{0t} + \gamma_{1t}\hat{\beta}_{pt} + \gamma_{2t}\hat{S}_{pt} + U_{pt} \qquad (2)$$

where

$\hat{\beta}_{pt}$ = estimated beta for portfolio p at year t, t=1968, ..., 1987

$\hat{S}_{pt}$ = mean of the logarithm of firm size in portfolio p at the beginning of year t

U_{pt} = disturbance term.

Depending on whether daily, weekly or monthly returns are used, a portfolio's average return changes periodically while its beta and size only change once a year. The γ_1 and γ_2 coefficients are estimated over the following four subperiods: 1968-72, 1973-77, 1978-82 and 1983-1987. If portfolio betas can fully account for the differences in returns, one would expect the average coefficient for the beta variable to be positive and for the size variable to be zero. A t-statistic will be used to test the hypothesis. The coefficients of a matched sample are also examined so that the results between industrial and utility firms can be compared.

V. Analysis of Results

1. Equity Value of the Utility Portfolios

The mean equity values of the ten size-based utility portfolios are reported in Table 1. Panels A and B present the average firm size of these portfolios at the beginning and end of the test period, 1968-1987. The first interesting observation from Table 1 is that the difference in magnitude between the smallest and the largest market value utility portfolios is tremendous. In Panel A, the average size of MV_1 is about $31 million while that of MV_{10} is over $1.4 billion. In Panel B, that is twenty years later, they are $62 million and $5.2 billion, respectively. Another interesting finding is that there is a substantial increase in average firm size from MV_9 to MV_{10}. Since these two findings are consistent over the entire test period, the average portfolio market values for interim years are not reported. These results are similar to the empirical evidence provided by Reinganum (1981).

The utility sample in this study contains 152 firms whereas Reinganum's sample contains 535 firms that are mainly industrial companies. Two conclusions may be drawn from the results of the Reinganum study and this one. First, utilities and industriales are similar in the sense that their market values vary over a wide spectrum. Second, the fact that there is a huge jump in firm size from MV_9 to MV_{10} indicates that the distribution of firm size is positively skewed. To correct for the skewness problem, the natural logarithm of the mean equity value of each portfolio is calculated. This variable is then used in later regressions instead of the actual mean equity value.

2. Betas of the Utility and Industrial Samples

The betas based on monthly, weekly and daily returns are reported for the utility and industrial samples. For simplicity, they will be referred to as monthly, weekly, and daily betas. In all cases, five years of returns are used to estimate the systematic risk. The betas estimated over the 1963-67 time period are used to proxy for the betas in 1968, which is the beginning of the test period. By the same token, the betas obtained from the time period 1982-86 are used as proxies for the betas in 1987, which is the end of the test period.

The betas from using the equally-weighted and value-weighted indices are calculated in order to check whether the results are affected by the choice of market index. Since the results are similar, only those obtained from the equally-weighted index are reported and analyzed.

Table 2 reports the monthly, weekly and daily betas of the two samples at the beginning and end of the test period. Panel A shows the various betas of the industrial portfolios. Two conclusions may be drawn. First, in the 1960's, smaller market value portfolios tend to have relatively larger betas. This is consistent with the empirical findings by Banz (1981) and Reinganum (1981). Second, this trend seems to vanish in the 1980's, especially when weekly and daily returns are used.

The betas of the utility portfolios are presented in Panel B. The table shows that none of the utility betas are greater than 0.71. A comparison between Panels A and B reveals that utility portfolios are relatively less risky than industrial portfolios after controlling for firm size. The comparison also reveals that, unlike industrial stocks, betas of the utility portfolios are not related to the market values of equity.

The negative correlation between firm size and beta in the industrial sample may introduce a multicolinearity problem in estimating equation (2). Banz (p.11) had addressed this issue and concluded that the test results are not sensitive to the multicolinearity problem. For the utility sample, this problem does not exist.

3. Tests on the Coefficients of Beta and Size

The beta and firm size are used to estimate γ_1 and γ_2 in equation (2). A t-statistic is used to test if the mean values of the gammas are significantly different from zero. The tests were performed for four 5-year periods which are reported in Table 3. The mean of the gammas and their t-statistic are presented in Panel A for the utilities and in Panel B for the industrial firms.

The empirical results for the utility sample are reported in Panel A of Table 3. When monthly returns are used, 60 regressions were run to obtain 60 pairs of gammas for each of the 5-year periods. When daily returns are used, over 1200 regressions were run for each period to obtain the gammas. The results are similar: in all of the time periods tested, none of the average coefficients for beta and size are significantly different from zero. When weekly returns are used, 260 pairs of gammas were obtained. The average coefficients for beta are not significant in any test period, and the average coefficients for size are not significant in three of the test periods. For the test period of 1978-82, the average coefficient for size is significantly negative at a 5% level.

The test results for the industrial sample are reported in Panel B of Table 3. When monthly returns are used, the average coefficient estimates for size and beta are significant and have the expected sign only in the 1983-87 test period. When weekly returns are used, only the size variable is significantly negative in the 1978-82 period. When daily returns are used, the coefficient estimates for betas and size are not significant at any conventional level.

According to the CAPM, beta is the sole determinant of stock returns. It is expected that the coefficient for beta is significantly positive. However, the empirical findings reported in this study and in Fama and French (1992) only provide weak support for beta in explaining stock returns. The empirical findings in this study also suggest that the size effect varies over time. It is not unusual to document the firm size effect at certain time periods but not at others. Banz (1981) found that the size effect is not stable over time with substantial differences in the magnitude of the coefficient of the size factor (p.9, Table 1). Brown, Kleidon and Marsh (1983) not only have shown that size effect is not constant over time but also have reported a reversal of the size anomaly for certain years.

The research design of this study allows us to keep the sample, test period, and methodology the same with the holding-period being the only variable. The size effect is documented for the industrial sample in one of the four test periods when monthly returns are used and in another when weekly returns are used. When daily returns are used, no size effect is observed. For the utility sample, the size effect is significant in only one test period when weekly returns are used. When monthly and daily returns are used, no size effect is found. Therefore, this study concludes that the size effect is not only time-period specific but also holding-period specific.

VI. Concluding Remarks

The fact that the two samples show different, though weak, results indicates that utility and industrial stocks do not share the same characteristics. First, given firm size, utility stocks are consistently less risky than industrial stocks. Second, industrial betas tend to decrease with firm size but utility betas do not. These findings may be attributed to the fact that all public utilities operate in an environment with regional monopolistic power and regulated financial structure. As a result, the business and financial risks are very similar among the utilities regardless of their sizes. Therefore, utility betas would not necessarily be expected to be related to firm size.

The objective of this study is to examine if the size effect exists in the utility industry. After controlling for equity values, there is some weak evidence that firm size is a missing factor from the CAPM for the industrial but not for the utility stocks. This implies that although the size phenomenon has been strongly documented for the industriales, the findings suggest that there is no need to adjust for the firm size in utility rate regulations.

References

Banz, R.W. "The Relationship Between Return and Market Value of Common Stocks," *Journal of Financial Economics*, (March 1981): 3-18.

Barry, C.B. and S.J. Brown. "Differential Information and the Small Firm Effect," *Journal of Financial Economics*, (1984): 283-294,

Brauer, G.A. "Using Jump-Diffusion Return Models to Measure Differential Information by Firm

Size," *Journal of Financial and Quantitative Analysis*, (December 1986): 447-458.

Brown, P., A.W. Kleidon, and T.A. Marsh. "New Evidence on the Nature and Size Related Anomalies in Stock Prices," *Journal of Financial Economics*, (1983): 33-56.

Fama, E.F. and K.R. French. "The Cross-Section of Expected Stock Returns," *Journal of Finance*, (June 1992): 427-465.

Fama, E.F. and J.D. MacBeth. "Risk Return and Equilibrium: Empirical Tests," *Journal of Political Economy*, (May/June 1973): 607-636.

Reinganum, M.R. "Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings' Yields and Market Values," *Journal of Financial Economics*, (March 1981): 19-46.

Table 1

Average Equity Size of the Utility Portfolios at the Beginning and End of the Test Period
(Dollar figures in millions)

	A: Beginning (1968)	B: End (1987)
MV_1	$31	$62
MV_2	$77	$177
MV_3	$113	$334
MV_4	$161	$475
MV_5	$220	$715
MV_6	$334	$957
MV_7	$437	$1,279
MV_8	$505	$1,805
MV_9	$791	$2,665
MV_{10}	$1,447	$5,399

Table 2

Betas of the Two Samples at the Beginning and End of the Test Period

	Monthly Betas		Weekly Betas		Daily Betas	
	1963-67	1982-86	1963-67	1982-86	1963-67	1982-86
Panel A: Industrial Firms						
MV_1	0.89	1.00	1.15	0.95	1.11	0.92
MV_2	0.94	0.87	1.07	1.01	1.14	1.01
MV_3	0.88	0.82	1.12	0.86	1.14	1.04
MV_4	0.69	0.74	1.00	0.83	1.03	0.86
MV_5	0.73	0.80	1.05	0.96	1.13	1.01
MV_6	0.66	0.82	1.03	1.01	1.05	1.04
MV_7	0.64	0.81	0.97	1.04	0.98	1.09
MV_8	0.62	0.75	0.97	1.11	1.00	1.20
MV_9	0.52	0.78	0.84	1.06	0.94	1.16
MV_{10}	0.43	0.65	0.78	1.01	0.86	1.22
Panel B: Public Utilities						
MV_1	0.30	0.37	0.31	0.43	0.30	0.40
MV_2	0.28	0.38	0.37	0.47	0.36	0.44
MV_3	0.22	0.42	0.33	0.42	0.31	0.49
MV_4	0.27	0.35	0.36	0.52	0.34	0.54
MV_5	0.25	0.45	0.37	0.61	0.35	0.62
MV_6	0.25	0.41	0.39	0.54	0.40	0.65
MV_7	0.20	0.35	0.34	0.54	0.37	0.63
MV_8	0.17	0.38	0.34	0.65	0.33	0.68
MV_9	0.19	0.34	0.35	0.60	0.34	0.71
MV_{10}	0.18	0.29	0.38	0.59	0.39	0.71

Table 3

Tests on the Mean Coefficients of Beta (γ_1) and Size (γ_2)

$$R_{pt} = \gamma_{ot} + \gamma_{1t}\hat{\beta}_{pt} + \gamma_{2t}\hat{S}_{pt} + U_{pt}$$

Returns Used:		Monthly (t-value)	Weekly (t-value)	Daily (t-value)
Panel A: Utility Sample				
1968-72	γ_1	-0.46% (-0.26)	-0.32% (-0.42)	-0.02% (-0.18)
	γ_2	-0.07% (-0.78)	-0.01% (-0.51)	-0.00% (-0.46)
1973-77	γ_1	-0.28% (-0.13)	0.14% (0.14)	-0.03% (-0.21)
	γ_2	-0.11% (-0.70)	-0.03% (-0.67)	-0.00% (-0.53)
1978-82	γ_1	0.55% (0.36)	0.54% (1.00)	0.05% (0.43)
	γ_2	-0.10% (-0.75)	-0.05% (-1.71)*	-0.01% (-1.60)
1983-87	γ_1	1.74% (1.28)	-0.24% (-0.51)	-0.02% (-0.18)
	γ_2	-0.16% (-1.54)	-0.03% (-0.86)	-0.01% (-0.63)
Panel B: Industrial Sample				
1968-72	γ_1	-0.36% (-0.27)	-0.28% (-0.55)	-0.02% (-0.32)
	γ_2	0.07% (0.43)	-0.01% (-0.19)	0.00% (0.51)
1973-77	γ_1	1.34% (0.64)	-0.23% (-0.31)	0.14% (1.45)
	γ_2	-0.01% (-0.06)	-0.04% (-0.85)	-0.00% (-0.64)
1978-82	γ_1	-0.84% (-0.28)	-0.56% (-0.91)	-0.09% (-0.81)
	γ_2	-0.29% (-0.75)	-0.01% (-1.72)*	-0.00% (-1.33)
1983-87	γ_1	2.51% (1.83)*	0.34% (0.64)	0.11% (1.40)
	γ_2	-0.25% (-1.90)*	-0.01% (-0.43)	0.00% (0.14)

* Significant at the 5% level based on a one-tailed test.