Q. Referring to page 33, lines 10-11: Please provide complete references to all the studies cited as well as copies of those that are available to you.

A. Please see the attached seven articles.

1. Eugene F. Fama and Kenneth R. French, “The Equity Premium”, *Journal of Finance,* Vol. LVII, No. 2, April 2002.

2. Elroy Dimson, Paul Marsh and Mike Staunton, “Global Evidence on the Equity Risk Premium”, forthcoming in *Journal of Applied Corporate Finance,* September 2002.

3. John R. Graham, Campbell R. Harvey, “Expectations of equity risk premia, volatility and asymmetry from a corporate finance perspective”, Working Paper, Duke University, November 30, 2001.

4. Robert D. Arnott and Peter L. Bernstein, “What Risk Premium is ‘Normal’”?, *Financial Analysts Journal,* March/April 2002.

5. Jeremy J. Siegel, “The Shrinking Equity Premium”, *The Journal of Portfolio Management*, Fall 1999.

6. James Claus and Jacob Thomas, “Equity Premia as Low as Three Percent? Evidence from Analysts’ Earnings Forecasts for Domestic and International Stock Markets”, *The Journal of Finance*, Vol. LVI, No. 5, October 2001.

7. Roger G. Ibbotson and Peng Chen, "Long-Run Stock Returns: Participating in the Real Economy", *Financial Analysts Journal*, January/February 2003.

THE JOURNAL OF FINANCE • VOL. LVII, NO. 2 • APRIL 2002

The Equity Premium

EUGENE F. FAMA and KENNETH R. FRENCH*

ABSTRACT

We estimate the equity premium using dividend and earnings growth rates to measure the expected rate of capital gain. Our estimates for 1951 to 2000, 2.55 percent and 4.32 percent, are much lower than the equity premium produced by the average stock return, 7.43 percent. Our evidence suggests that the high average return for 1951 to 2000 is due to a decline in discount rates that produces a large unexpected capital gain. Our main conclusion is that the average stock return of the last half-century is a lot higher than expected.

THE EQUITY PREMIUM—the difference between the expected return on the market portfolio of common stocks and the risk-free interest rate—is important in portfolio allocation decisions, estimates of the cost of capital, the debate about the advantages of investing Social Security funds in stocks, and many other applications. The average return on a broad portfolio of stocks is typically used to estimate the expected market return. The average real return for 1872 to 2000 on the S&P index (a common proxy for the market portfolio, also used here) is 8.81 percent per year. The average real return on six-month commercial paper (a proxy for the risk-free interest rate) is 3.24 percent. This large spread (5.57 percent) between the average stock return and the interest rate is the source of the so-called equity premium puzzle: Stock returns seem too high given the observed volatility of consumption (Mehra and Prescott (1985)).

We use fundamentals (dividends and earnings) to estimate the expected stock return. Along with other evidence, the expected return estimates from fundamentals help us judge whether the realized average return is high or low relative to the expected value.

The logic of our approach is straightforward. The average stock return is the average dividend yield plus the average rate of capital gain:

$$A(R_t) = A(D_t/P_{t-1}) + A(GP_t), \tag{1}$$

* Fama is from the University of Chicago and French is from Dartmouth College. The comments of John Campbell, John Cochrane, Kent Daniel, John Heaton, Jay Ritter, Andrei Shleifer, Rex Sinquefield, Tuomo Vuolteenaho, Paul Zarowin, and seminar participants at Boston College, Dartmouth College, the NBER, Purdue University, the University of Chicago, and Washington University have been helpful. Richard Green (the editor) and the two referees get special thanks.

where D_t is the dividend for year t, P_{t-1} is the price at the end of year $t - 1$, $GP_t = (P_t - P_{t-1})/P_{t-1}$ is the rate of capital gain, and $A(\ )$ indicates an average value. (Throughout the paper, we refer to D_t/P_{t-1} as the dividend yield and D_t/P_t is the dividend–price ratio. Similarly, Y_t/P_{t-1}, the ratio of earnings for year t to price at the end of year $t - 1$, is the earnings yield and Y_t/P_t is the earnings–price ratio.)

Suppose the dividend–price ratio, D_t/P_t, is stationary (mean reverting). Stationarity implies that if the sample period is long, the compound rate of dividend growth approaches the compound rate of capital gain. Thus, an alternative estimate of the expected stock return is

$$A(RD_t) = A(D_t/P_{t-1}) + A(GD_t), \tag{2}$$

where $GD_t = (D_t - D_{t-1})/D_{t-1}$ is the growth rate of dividends. We call (2) the dividend growth model.

The logic that leads to (2) applies to any variable that is cointegrated with the stock price. For example, the dividend–price ratio may be nonstationary because firms move away from dividends toward share repurchases as a way of returning earnings to stockholders. But if the earnings–price ratio, Y_t/P_t, is stationary, the average growth rate of earnings, $A(GY_t) = A((Y_t - Y_{t-1})/Y_{t-1})$, is an alternative estimate of the expected rate of capital gain. And $A(GY_t)$ can be combined with the average dividend yield to produce another estimate of the expected stock return:

$$A(RY_t) = A(D_t/P_{t-1}) + A(GY_t). \tag{3}$$

We call (3) the earnings growth model.[1]

We should be clear about the expected return concept targeted by (1), (2), and (3). D_t/P_t and Y_t/P_t vary through time because of variation in the conditional (point-in-time) expected stock return and the conditional expected growth rates of dividends and earnings (see, e.g., Campbell and Shiller (1989)). But if the stock return and the growth rates are stationary (they have constant unconditional means), D_t/P_t and Y_t/P_t are stationary. Then, like the average return (1), the dividend and earnings growth models (2) and (3) provide estimates of the unconditional expected stock return. In short, the focus of the paper is estimates of the unconditional expected stock return.

The estimate of the expected real equity premium for 1872 to 2000 from the dividend growth model (2) is 3.54 percent per year. The estimate from the average stock return, 5.57 percent, is almost 60 percent higher. The difference between the two is largely due to the last 50 years. The equity premium for 1872 to 1950 from the dividend growth model, 4.17 percent per year, is close to the estimate from the average return, 4.40 percent. In con-

[1] Motivated by the model in Lettau and Ludvigson (2001), one can argue that if the ratio of consumption to stock market wealth is stationary, the average growth rate of consumption is another estimate of the expected rate of capital gain. We leave this path to future work.

trast, the equity premium for 1951 to 2000 produced by the average return, 7.43 percent per year, is almost three times the estimate, 2.55 percent, from (2). The estimate of the expected real equity premium for 1951 to 2000 from the earnings growth model (3), 4.32 percent per year, is larger than the estimate from the dividend growth model (2). But the earnings growth estimate is still less than 60 percent of the estimate from the average return.

Three types of evidence suggest that the lower equity premium estimates for 1951 to 2000 from fundamentals are closer to the expected premium. (a) The estimates from fundamentals are more precise. For example, the standard error of the estimate from the dividend growth model is less than half the standard error of the estimate from the average return. (b) The Sharpe ratio for the equity premium from the average stock return for 1951 to 2000 is just about double that for 1872 to 1950. In contrast, the equity premium from the dividend growth model has a similar Sharpe ratio for 1872 to 1950 and 1951 to 2000. (c) Most important, valuation theory specifies relations among the book-to-market ratio, the return on investment, and the cost of equity capital (the expected stock return). The estimates of the expected stock return for 1951 to 2000 from the dividend and earnings growth models line up with other fundamentals in the way valuation theory predicts. But the book-to-market ratio and the return on investment suggest that the expected return estimate from the average stock return is too high.

Our motivation for the dividend growth model (2) is simpler and more general, but (2) can be viewed as the expected stock return estimate of the Gordon (1962) model. Our work is thus in the spirit of a growing literature that uses valuation models to estimate expected returns (e.g., Blanchard (1993), Claus and Thomas (2001), and Gebhardt, Lee, and Swaminathan (2001)). Claus and Thomas and Gebhardt, Lee, and Swaminathan use forecasts by security analysts to estimate expected cash flows. Their analyst forecasts cover short periods (1985 to 1998 and 1979 to 1995). We use realized dividends and earnings from 1872 to 2000. This 129-year period provides a long perspective, which is important for judging the competing expected return estimates from fundamentals and realized stock returns. Moreover, though the issue is controversial (Keane and Runkle (1998)), Claus and Thomas find that analyst forecasts are biased; they tend to be substantially above observed growth rates. The average growth rates of dividends and earnings we use are unbiased estimates of expected growth rates.

Like us, Blanchard (1993) uses dividend growth rates to estimate the expected rate of capital gain, which he combines with an expected dividend yield to estimate the expected stock return. But his focus is different and his approach is more complicated than ours. He is interested in the path of the conditional expected stock return. His conditional expected return is the sum of the fitted values from time-series regressions of the realized dividend yield and a weighted average of 20 years of future dividend growth rates on four predetermined variables (the dividend yield, the real rate of capital gain, and the levels of interest rates and inflation). He focuses on describing the path of the conditional expected return in terms of his four explanatory variables.

In contrast, our prime interest is the unconditional expected return, which we estimate more simply as the sum of the average dividend yield and the average growth rate of dividends or earnings. This approach is valid if the dividend–price and earnings–price ratios are stationary. And we argue below that it continues to produce estimates of the average expected stock return when the price ratios are subject to reasonable forms of nonstationarity. Given its simplicity and generality, our approach is an attractive addition to the research toolbox for estimating the expected stock return.

Moreover, our focus is comparing alternative estimates of the unconditional expected stock return over the long 1872 to 2000 period, and explaining why the expected return estimates for 1951 to 2000 from fundamentals are much lower than the average return. Our evidence suggests that much of the high return for 1951 to 2000 is unexpected capital gain, the result of a decline in discount rates.

Specifically, the dividend–price and earnings–price ratios fall from 1950 to 2000; the cumulative percent capital gain for the period is more than three times the percent growth in dividends or earnings. All valuation models agree that the two price ratios are driven by expectations about future returns (discount rates) and expectations about dividend and earnings growth. Confirming Campbell (1991), Cochrane (1994), and Campbell and Shiller (1998), we find that dividend and earnings growth rates for 1950 to 2000 are largely unpredictable. Like Campbell and Shiller (1998), we thus infer that the decline in the price ratios is mostly due to a decline in expected returns. Some of this decline is probably expected, the result of reversion of a high 1950 conditional expected return to the unconditional mean. But most of the decline in the price ratios seems to be due to the unexpected decline of expected returns to ending values far below the mean.

The paper proceeds as follows. The main task, addressed in Sections I and II, is to compare and evaluate the estimates of the unconditional annual expected stock return provided by the average stock return and the dividend and earnings growth models. Section III then considers the issues that arise if the goal is to estimate the long-term expected growth of wealth, rather than the unconditional expected annual (simple) return. Section IV concludes.

I. The Unconditional Annual Expected Stock Return

Table I shows estimates of the annual expected real equity premium for 1872 to 2000. The market portfolio is the S&P 500 and its antecedents. The deflator is the Producer Price Index until 1925 (from Shiller (1989)) and the Consumer Price Index thereafter (from Ibbotson Associates). The risk-free interest rate is the annual real return on six-month commercial paper, rolled over at midyear. The risk-free rate and S&P earnings data are from Shiller, updated by Vuolteenaho (2000) and us. Beginning in 1925, we construct S&P book equity data from the book equity data in Davis, Fama, and French (2000), expanded to include all NYSE firms. The data on dividends, prices, and returns for 1872 to 1925 are from Shiller. Shiller's annual data on the

Table I

Real Equity Premium and Related Statistics for the S&P Portfolio

The inflation rate for year t is $Inf_t = L_t/L_{t-1} - 1$, where L_t is the price level at the end of year t. The real return for year t on six-month (three-month for the year 2000) commercial paper (rolled over at midyear) is F_t. The nominal values of book equity and price for the S&P index at the end of year t are b_t and p_t. Nominal S&P dividends and earnings for year t are d_t and y_t. Real rates of growth of dividends, earnings, and the stock price are $GD_t = (d_t/d_{t-1})*(L_{t-1}/L_t) - 1$, $GY_t = (y_t/y_{t-1})*(L_{t-1}/L_t) - 1$, and $GP_t = (p_t/p_{t-1})*(L_{t-1}/L_t) - 1$. The real dividend yield is $D_t/P_{t-1} = (d_t/p_{t-1})*(L_{t-1}/L_t)$. The real income return on investment is $Y_t/B_{t-1} = (1 + y_t/b_{t-1})*(L_{t-1}/L_t) - 1$. The dividend growth estimate of the real S&P return for t is $RD_t = D_t/P_{t-1} + GD_t$, the earnings growth estimate is $RY_t = D_t/P_{t-1} + GY_t$, and R_t is the realized real S&P return. The dividend and earnings growth estimates of the real equity premium for year t are $RXD_t = RD_t - F_t$ and $RXY_t = RY_t - F_t$, and $RX_t = R_t - F_t$ is the real equity premium from the realized real return. The Sharpe ratio for $RD_t - F_t$ (the mean of $RD_t - F_t$ divided by the standard deviation of R_t) is SD, SY is the Sharpe ratio for $RY_t - F_t$ (the mean of $RY_t - F_t$ divided by the standard deviation of R_t), and SR is the Sharpe ratio for $R_t - F_t$ (the mean of $R_t - F_t$ divided by the standard deviation of R_t). Except for the Sharpe ratios, all variables are expressed as percents, that is, they are multiplied by 100.

	Inf_t	F_t	D_t/P_{t-1}	GD_t	GY_t	GP_t	RD_t	RY_t	R_t	RXD_t	RXY_t	RX_t	SD	SY	SR
Means of annual values of variables															
1872–2000	2.16	3.24	4.70	2.08	NA	4.11	6.78	NA	8.81	3.54	NA	5.57	0.20	NA	0.31
1872–1950	0.99	3.90	5.34	2.74	NA	2.96	8.07	NA	8.30	4.17	NA	4.40	0.22	NA	0.23
1951–2000	4.00	2.19	3.70	1.05	2.82	5.92	4.74	6.51	9.62	2.55	4.32	7.43	0.15	0.25	0.44
Standard deviations of annual values of variables															
1872–2000	7.51	8.48	1.39	12.37	NA	17.83	12.56	NA	18.03	13.00	NA	18.51			
1872–1950	9.11	10.63	1.12	15.28	NA	18.48	15.41	NA	18.72	16.02	NA	19.57			
1951–2000	3.11	2.46	1.17	5.09	13.79	16.77	5.21	13.51	17.03	5.62	14.02	16.73			
Means of annual continuously compounded returns and growth rates															
1872–2000	1.86	2.87		1.34	NA	2.48			7.00						
1872–1950	0.59	3.33		1.60	NA	1.22			6.41						
1951–2000	3.88	2.14		0.92	1.89	4.46			7.94						

	b_t/p_t	RD_t	RY_t	R_t	Y_t/B_{t-1}
Means of annual values of variables					
1951–2000	0.66	4.74	6.51	9.62	7.60

level of the S&P (used to compute returns and other variables involving price) are averages of daily January values. The S&P dividend, price, and return data for 1926 to 2000 are from Ibbotson Associates, and the returns for 1926 to 2000 are true annual returns.

Without showing the details, we can report that the CRSP value-weight portfolio of NYSE, AMEX, and Nasdaq stocks produces average returns and dividend growth estimates of the expected return close to the S&P estimates for periods after 1925 when both indices are available. What one takes to be the risk-free rate has a bigger effect. For example, substituting the one-month Treasury bill rate for the six-month commercial paper rate causes estimates of the annual equity premium for 1951 to 2000 to rise by about one percent. But for our main task—comparing equity premium estimates from (1), (2), and (3)—differences in the risk-free rate are an additive constant that does not affect inferences.

One can estimate expected returns in real or nominal terms. Since portfolio theory says the goal of investment is consumption, real returns seem more relevant, and only results for real returns are shown. Because of suspicions about the quality of the price deflator during the early years of 1872 to 2000, we have replicated the results for nominal returns. They support all the inferences from real returns.

The dividend and earnings growth models (2) and (3) assume that the market dividend–price and earnings–price ratios are stationary. The first three annual autocorrelations of D_t/P_t for 1872 to 2000 are 0.73, 0.51, and 0.47. For the 1951 to 2000 period that occupies much of our attention, the autocorrelations are 0.83, 0.72, and 0.69. The autocorrelations are large, but their decay is roughly like that of a stationary first-order autoregression (AR1). This is in line with formal evidence (Fama and French (1988), Cochrane (1994), and Lamont (1998)) that the market dividend–price ratio is highly autocorrelated but slowly mean-reverting. S&P earnings data for the early years of 1872 to 2000 are of dubious quality (Shiller (1989)), so we estimate expected returns with the earnings growth model (3) only for 1951 to 2000. The first three autocorrelations of Y_t/P_t for 1951 to 2000, 0.80, 0.70, and 0.61, are again roughly like those of a stationary AR1.

We emphasize, however, that our tests are robust to reasonable nonstationarity of D_t/P_t and Y_t/P_t. It is not reasonable that the expected stock return and the expected growth rates of dividends and earnings that drive D_t/P_t and Y_t/P_t are nonstationary processes that can wander off to infinity. But nonstationarity of D_t/P_t and Y_t/P_t due to structural shifts in productivity or preferences that permanently change the expected return or the expected growth rates is reasonable. Such regime shifts are not a problem for the expected return estimates from (2) and (3), as long as D_t/P_t and Y_t/P_t mean-revert within regimes. If the regime shift is limited to expected dividend and earnings growth rates, the permanent change in expected growth rates is offset by a permanent change in the expected dividend yield, and (2) and (3) continue to estimate the (stationary) expected stock return. (An Appendix, available on request, provides an example.) If there is a permanent shift in the expected stock return, it is nonstationary, but like the average return in (1), the dividend and earnings growth models in (2) and (3) estimate the average expected return during the sample period.

Indeed, an advantage of the expected return estimates from fundamentals is that they are likely to be less sensitive than the average return to long-lived shocks to dividend and earnings growth rates or the expected stock return. For example, a permanent shift in the expected return affects the average dividend yield, which is common to the three expected return estimates, but it produces a shock to the capital gain term in the average return in (1) that is not shared by the estimates in (2) and (3). In short, the estimates of the expected stock return from fundamentals are likely to be more precise than the average stock return.

A. *The Equity Premium*

For much of the period from 1872 to 2000—up to about 1950—the dividend growth model and the average stock return produce similar estimates of the expected return. Thereafter, the two estimates diverge. To illustrate, Table I shows results for 1872 to 1950 (79 years) and 1951 to 2000 (50 years). The year 1950 is a big year, with a high real stock return (23.40 percent), and high dividend and earnings growth estimates of the return (29.96 percent and 24.00 percent). But because the three estimates of the 1950 return are similarly high, the ordering of expected return estimates, and the inferences we draw from them, are unaffected by whether 1950 is allocated to the earlier or the later period. Indeed, pushing the 1950 break-year backward or forward several years does not affect our inferences.

For the earlier 1872 to 1950 period, there is not much reason to favor the dividend growth estimate of the expected stock return over the average return. Precision is not an issue; the standard errors of the two estimates are similar (1.74 percent and 2.12 percent), the result of similar standard deviations of the annual dividend growth rate and the rate of capital gain, 15.28 percent and 18.48 percent. Moreover, the dividend growth model and the average return provide similar estimates of the expected annual real return for 1872 to 1950, 8.07 percent and 8.30 percent. Given similar estimates of the expected return, the two approaches produce similar real equity premiums for 1872 to 1950, 4.17 percent (dividend growth model) and 4.40 percent (stock returns).

The competition between the dividend growth model and the average stock return is more interesting for 1951 to 2000. The dividend growth estimate of the 1951 to 2000 expected return, 4.74 percent, is less than half the average return, 9.62 percent. The dividend growth estimate of the equity premium, 2.55 percent, is 34 percent of the estimate from returns, 7.43 percent. The 1951 to 2000 estimates of the expected stock return and the equity premium from the earnings growth model, 6.51 percent and 4.32 percent, are higher than for the dividend growth model. But they are well below the estimates from the average return, 9.62 percent and 7.43 percent.

B. Evaluating the Expected Return Estimates for 1951 to 2000

We judge that the estimates of the expected stock return for 1951 to 2000 from fundamentals are closer to the true expected value, for three reasons.

(a) The expected return estimates from the dividend and earnings growth models are more precise than the average return. The standard error of the dividend growth estimate of the expected return for 1951 to 2000 is 0.74 percent, versus 2.43 percent for the average stock return. Since earnings growth is more volatile than dividend growth, the standard error of the expected return from the earnings growth model, 1.93 percent, is higher than the estimate from the dividend growth model, but it is smaller than the 2.43 percent standard error of the average stock return. Claus and Thomas (2001) also argue that expected return estimates from fundamentals are more precise than average returns, but they provide no direct evidence.

(b) Table I shows Sharpe ratios for the three equity premium estimates. Only the average premium in the numerator of the Sharpe ratio differs for the three estimates. The denominator for all three is the standard deviation of the annual stock return. The Sharpe ratio for the dividend growth estimate of the equity premium for 1872 to 1950, 0.22, is close to that produced by the average stock return, 0.23. More interesting, the Sharpe ratio for the equity premium for 1951 to 2000 from the dividend growth model, 0.15, is lower than but similar to that for 1872 to 1950. The Sharpe ratio for the 1951 to 2000 equity premium from the earnings growth model, 0.25, is somewhat higher than the dividend growth estimate, 0.15, but it is similar to the estimates for 1872 to 1950 from the dividend growth model, 0.22, and the average return, 0.23.

In asset pricing theory, the Sharpe ratio is related to aggregate risk aversion. The Sharpe ratios for the 1872 to 1950 and 1951 to 2000 equity premiums from the dividend growth model and the earnings growth model suggest that aggregate risk aversion is roughly similar in the two periods. In contrast, though return volatility falls a bit, the equity premium estimate from the average stock return increases from 4.40 percent for 1872 to 1950 to 7.43 percent for 1951 to 2000, and its Sharpe ratio about doubles, from 0.23 to 0.44. It seems implausible that risk aversion increases so much from the earlier to the later period.

(c) Most important, the behavior of other fundamentals favors the dividend and earnings growth models. The average ratio of the book value of equity to the market value of equity for 1951 to 2000 is 0.66, the book-to-market ratio B_t/P_t is never greater than 1.12, and it is greater than 1.0 for only 6 years of the 50-year period. Since, on average, the market value of equity is substantially higher than its book value, it seems safe to conclude that, on average, the expected return on investment exceeds the cost of capital.

Suppose investment at time $t - 1$ generates a stream of equity earnings for t, $t + 1, \ldots, t + N$ with a constant expected value. The average income return on book equity, $A(Y_t/B_{t-1})$, is then an estimate of the expected return on equity's share of assets. It is an unbiased estimate when N is infinite and it is upward biased when N is finite. In either case, if the expected return on investment exceeds the cost of capital, we should find that (except for sampling error) the average income return on book equity is greater than estimates of the cost of equity capital (the expected stock return):

$$A(Y_t/B_{t-1}) > E(R). \tag{4}$$

Table I shows that (4) is confirmed when we use the dividend and earnings growth models to estimate the expected real stock return for 1951 to 2000. The estimates of $E(R)$, 4.74 percent (dividend growth model) and 6.51 percent (earnings growth model), are below 7.60 percent, the average real income return on book equity, $A(Y_t/B_{t-1})$. In contrast, the average real stock return for 1951 to 2000, 9.62 percent, exceeds the average income return by more than 2 percent. An expected stock return that exceeds the expected income return on book equity implies that the typical corporate investment has a negative net present value. This is difficult to reconcile with an average book-to-market ratio substantially less than one.

To what extent are our results new? Using analyst forecasts of expected cash flows and a more complicated valuation model, Claus and Thomas (2001) produce estimates of the expected stock return for 1985 to 1998 far below the average return. Like us, they argue that the estimates from fundamentals are closer to the true expected return. We buttress this conclusion with new results on three fronts. (a) The long-term perspective provided by the evidence that, for much of the 1872 to 2000 period, average returns and fundamentals produce similar estimates of the expected return. (b) Direct evidence that the expected return estimates for 1951 to 2000 from fundamentals are more precise. (c) Sharpe ratios and evidence on how the alternative expected return estimates line up with the income return on investment. These new results provide support for the expected return estimates from fundamentals, and for the more specific inference that the average stock return for 1951 to 2000 is above the expected return.

II. Unexpected Capital Gains

Valuation theory suggests three potential explanations for why the 1951 to 2000 average stock return is larger than the expected return. (a) Dividend and earnings growth for 1951 to 2000 is unexpectedly high. (b) The expected (post-2000) growth rates of dividends and earnings are unexpectedly high. (c) The expected stock return (the equity discount rate) is unexpectedly low at the end of the sample period.

A. Is Dividend Growth for 1951 to 2000 Unexpectedly High?

If the prosperity of the United States over the last 50 years was not fully anticipated, dividend and earnings growth for 1951 to 2000 exceed 1950 expectations. Such unexpected in-sample growth produces unexpected capital gains. But it does not explain why the average return for 1951 to 2000 (the average dividend yield plus the average rate of capital gain) is so much higher than the expected return estimates from fundamentals (the average dividend yield plus the average growth rate of dividends or earnings). To see the point, note that unexpected in-sample dividend and earnings growth do not affect either the 1950 or the 2000 dividend–price and earnings–price ratios. (The 2000 ratios depend on post-2000 expected returns and growth rates.) Suppose D_t/P_t and E_t/P_t were the same in 1950 and 2000. Then the total percent growth in dividends and earnings during the period would be the same as the percent growth in the stock price. And (1), (2), and (3) would provide similar estimates of the expected stock return.

It is worth dwelling on this point. There is probably survivor bias in the U.S. average stock return for 1872 to 1950, as well as for 1951 to 2000. During the 1872 to 2000 period, it was not a foregone conclusion that the U.S. equity market would survive several financial panics, the Great Depression, two world wars, and the cold war. The average return for a market that survives many potentially cataclysmic challenges is likely to be higher than the expected return (Brown, Goetzmann, and Ross (1995)). But if the positive bias shows up only as higher than expected dividend and earnings growth during the sample period, there is similar survivor bias in the expected return estimates from fundamentals—a problem we do not solve. Our more limited goal is to explain why the average stock return for 1951 to 2000 is so high relative to the expected return estimates from the dividend and earnings growth models.

Since unexpected growth for 1951 to 2000 has a similar effect on the three expected return estimates, the task of explaining why the estimates are so different falls to the end-of-sample values of future expected returns and expected dividend and earnings growth. We approach the problem by first looking for evidence that expected dividend or earnings growth is high at the end of the sample period. We find none. We then argue that the large spread of capital gains over dividend and earnings growth for 1951 to 2000, or equivalently, the low end-of-sample dividend–price and earnings–price ratios, are due to an unexpected decline in expected stock returns to unusually low end-of-sample values.

B. Are Post-2000 Expected Dividend and Earnings Growth Rates Unusually High?

The behavior of dividends and earnings provides little evidence that rationally assessed (i.e., true) long-term expected growth is high at the end of the sample period. If anything, the growth rate of real dividends declines during the 1951 to 2000 period (Table II). The average growth rate for the first two decades, 1.60 percent, is higher than the average growth rates for the last three, 0.68 percent. The regressions in Table III are more formal evidence on the best forecast of post-2000 real dividend growth rates. Re-

Table II

Means of Simple Real Equity Premium and Related Statistics for the S&P Portfolio for 10-year Periods

The inflation rate for year t is $Inf_t = L_t/L_{t-1} - 1$, where L_t is the price level at the end of year t. The real return for year t on six-month (three-month for the year 2000) commercial paper (rolled over at midyear) is F_t. The nominal price of the S&P index at the end of year t is p_t. Nominal S&P dividends and earnings for year t are d_t and y_t. Real rates of growth of dividends, earnings, and the stock price are $GD_t = (d_t/d_{t-1})*(L_{t-1}/L_t) - 1$, $GY_t = (y_t/y_{t-1})*(L_{t-1}/L_t) - 1$, and $GP_t = (p_t/p_{t-1})*(L_{t-1}/L_t) - 1$. The real dividend yield is $D_t/P_{t-1} = (d_t/p_{t-1})*(L_{t-1}/L_t)$. The dividend growth estimate of the real S&P return for t is $RD_t = D_t/P_{t-1} + GD_t$, the earnings growth estimate is $RY_t = D_t/P_{t-1} + GY_t$, and R_t is the realized real S&P return. The dividend and earnings growth estimates of the real equity premium for year t are $RXD_t = RD_t - F_t$ and $RXY_t = RY_t - F_t$, and $RX_t = R_t - F_t$ is the real equity premium from the realized real return. All variables are expressed as percents, that is, they are multiplied by 100.

	Inf_t	F_t	D_t/P_{t-1}	GD_t	GY_t	GP_t	RD_t	RY_t	R_t	RXD_t	RXY_t	RX_t
1872–1880	−2.77	9.86	6.29	4.62		7.13	10.91		13.42	1.06		3.56
1881–1890	−1.72	7.23	5.04	0.69		0.04	5.73		5.08	−1.51		−2.15
1891–1900	0.18	5.08	4.40	4.49		4.75	8.89		9.15	3.81		4.08
1901–1910	1.95	3.18	4.45	3.25		2.33	7.70		6.78	4.52		3.60
1911–1920	6.82	0.82	5.70	−3.43		−6.52	2.27		−0.83	1.45		−1.64
1921–1930	−1.70	7.41	5.72	9.07		11.83	14.78		17.54	7.37		10.13
1931–1940	−1.23	2.80	5.31	0.36		2.21	5.67		7.52	2.87		4.72
1941–1950	6.04	−4.57	5.90	3.02		2.33	8.91		8.22	13.48		12.79
1951–1960	1.79	1.05	4.68	1.22		10.64	5.90		15.32	4.85		14.27
1961–1970	2.94	2.27	3.21	1.98		2.69	5.19		5.90	2.92		3.63
1971–1980	8.11	−0.30	4.04	−0.86		−1.92	3.18		2.12	3.48		2.42
1981–1990	4.51	5.32	4.19	2.32		5.40	6.51		9.59	1.19		4.28
1991–2000	2.68	2.61	2.36	0.58		12.80	2.94		15.16	0.32		12.54

Table III

Regressions to Forecast Real Dividend and Earnings Growth Rates, GD_t and GY_t

The price level at the end of year t is L_t. The nominal values of book equity and price for the S&P index at the end of year t are b_t and p_t. Nominal S&P dividends and earnings for year t are d_t and y_t. The real dividend and earnings growth rates for year t are $GD_t = (d_t/d_{t-1})*(L_{t-1}/L_t) - 1$ and $GY_t = (y_t/y_{t-1})*(L_{t-1}/L_t) - 1$, and R_t is the realized real return on the S&P portfolio for year t. The regression intercept is *Int*, and t-Stat is the regression coefficient (*Coef*) divided by its standard error. The regression R^2 is adjusted for degrees of freedom. Except for the dividend payout ratio, d_t/y_t, all variables are expressed as percents, that is, they are multiplied by 100.

Panel A: One Year: The Regressions Forecast Real Dividend Growth, GD_t, with Variables Known at $t-1$

	Int	d_{t-1}/y_{t-1}	d_{t-1}/p_{t-1}	GD_{t-1}	GD_{t-2}	GD_{t-3}	R_{t-1}	R_{t-2}	R_{t-3}	R^2
1875–1950, N = 76 years										
Coef	29.56	−23.12	−2.63	−0.12	−0.07	−0.03	0.22	0.13	0.09	0.38
t-Stat	3.22	−3.17	−1.77	−1.08	−0.64	−0.29	2.24	1.37	1.01	
1951–2000, N = 50 years										
Coef	−2.16	2.97	0.11	−0.07	−0.20	−0.06	0.11	0.07	0.01	0.01
t-Stat	−0.40	0.33	0.16	−0.45	−1.57	−0.45	2.17	1.33	0.22	

Panel B: Two Years: The Regressions Forecast Real Dividend Growth, GD_t, with Variables Known at $t-2$

	Int	d_{t-2}/y_{t-2}	d_{t-2}/p_{t-2}	GD_{t-2}	GD_{t-3}	R_{t-2}	R_{t-3}	R^2
1875–1950, N = 76 years								
Coef	6.61	−11.60	0.31	−0.26	0.05	0.24	0.11	0.07
t-Stat	0.64	−1.28	0.18	−2.02	0.39	2.03	1.00	
1951–2000, N = 50 years								
Coef	−4.11	7.62	0.32	−0.14	−0.03	0.05	−0.01	−0.05
t-Stat	−0.73	0.81	0.46	−1.13	−0.28	0.99	−0.16	

Panel C: One Year: The Regressions Forecast Real Earnings Growth, GY_t, with Variables Known at $t-1$

	Int	Y_{t-1}/B_{t-2}	d_{t-1}/y_{t-1}	y_{t-1}/p_{t-1}	GY_{t-1}	GY_{t-2}	GY_{t-3}	R_{t-1}	R_{t-2}	R_{t-3}	R^2
1951–2000, N = 50 years											
Coef	5.48	0.11	13.06	−1.36	0.21	−0.13	−0.31	0.28	−0.25	0.03	0.40
t-Stat	0.33	0.11	0.52	−1.91	1.17	−0.89	−2.64	2.39	−2.18	0.26	

Panel D: Two Years: The Regressions Forecast Real Earnings Growth, GY_t, with Variables Known at $t-2$

	Int	Y_{t-2}/B_{t-3}	d_{t-2}/y_{t-2}	y_{t-2}/p_{t-2}	GY_{t-2}	GY_{t-3}	R_{t-2}	R_{t-3}	R^2
1951–2000, N = 50 years									
Coef	−7.60	0.46	2.05	−0.74	−0.16	−0.39	−0.31	−0.12	0.23
t-Stat	−0.43	1.66	0.76	−1.02	−0.92	−2.54	−2.59	−0.97	

gressions are shown for forecasts one year ahead (the explanatory variables for year t dividend growth are known at the end of year $t - 1$) and two years ahead (the explanatory variables are known at the end of year $t - 2$).

The regression for 1875 to 1950 suggests strong forecast power one year ahead. The slopes on the lagged payout ratio, the dividend–price ratio, and the stock return are close to or more than two standard errors from zero, and the regression captures 38 percent of the variance of dividend growth. Even in the 1875 to 1950 period, however, power to forecast dividend growth does not extend much beyond a year. When dividend growth for year t is explained with variables known at the end of year $t - 2$, the regression R^2 falls from 0.38 to 0.07. Without showing the details, we can report that extending the forecast horizon from two to three years causes all hint of forecast power to disappear. Thus, for 1875 to 1950, the best forecast of dividend growth more than a year or two ahead is the historical average growth rate.

We are interested in post-2000 expected dividend growth, and even the short-term forecast power of the dividend regressions for 1872 to 1950 evaporates in the 1951 to 2000 period. The lagged stock return has some information ($t = 2.17$) about dividend growth one year ahead. But the 1951 to 2000 regression picks up only one percent of the variance of dividend growth. And forecast power does not improve for longer forecast horizons. Our evidence that dividend growth is essentially unpredictable during the last 50 years confirms the results in Campbell (1991), Cochrane (1991, 1994), and Campbell and Shiller (1998). If dividend growth is unpredictable, the historical average growth rate is the best forecast of future growth.

Long-term expected earnings growth also is not unusually high in 2000. There is no clear trend in real earnings growth during the 1951 to 2000 period. The most recent decade, 1991 to 2000, produces the highest average growth rate, 7.58 percent per year (Table II). But earnings growth is volatile. The standard errors of 10-year average growth rates vary around 5 percent. It is thus not surprising that 1981 to 1990, the decade immediately preceding 1991 to 2000, produces the lowest average real earnings growth rate, 0.37 percent per year.

The regressions in Table III are formal evidence on the predictability of earnings growth during the 1951 to 2000 period. There is some predictability of near-term growth, but it is largely due to transitory variation in earnings that is irrelevant for forecasting long-term earnings. In the 1951 to 2000 regression to forecast earnings growth one year ahead, the slope on the first lag of the stock return is positive (0.28, $t = 2.39$), but the slope on the second lag is negative (-0.25, $t = -2.18$) and about the same magnitude. Thus, the prediction of next year's earnings growth from this year's return is reversed the following year. In the one-year forecast regression for 1951 to 2000, the only variable other than lagged returns with power to forecast earnings growth ($t = -2.64$) is the third lag of earnings growth. But the slope is negative, so it predicts that the strong earnings growth of recent years is soon to be reversed.

In the 1951 to 2000 regression to forecast earnings one year ahead, there is a hint ($t = -1.91$) that the low earnings–price ratio at the end of the period implies higher than average expected growth one year ahead. But the effect peters out quickly; the slope on the lagged earnings–price ratio in the regression to forecast earnings growth two years ahead is -1.02 standard errors from zero. The only variables with forecast power two years ahead are the second lag of the stock return and the third lag of earnings growth. But the slopes on these variables are negative, so again the 2000 prediction is that the strong earnings growth of recent years is soon to be reversed. And again, regressions (not shown) confirm that forecast power for 1951 to 2000 does not extend beyond two years. Thus, beyond two years, the best forecast of earnings growth is the historical average growth rate.

In sum, the behavior of dividends for 1951 to 2000 suggests that future growth is largely unpredictable, so the historical mean growth rate is a near optimal forecast of future growth. Earnings growth for 1951 to 2000 is somewhat predictable one and two years ahead, but the end-of-sample message is that the recent high growth rates are likely to revert quickly to the historical mean. It is also worth noting that the market survivor bias argument of Brown, Goetzmann, and Ross (1995) suggests that past average growth rates are, if anything, upward biased estimates of future growth. In short, we find no evidence to support a forecast of strong future dividend or earnings growth at the end of our sample period.

C. Do Expected Stock Returns Fall during the 1951 to 2000 Period?

The S&P dividend–price ratio, D_t/P_t, falls from 7.18 percent at the end of 1950 to a historically low 1.22 percent at the end of 2000 (Figure 1). The growth in the stock price, P_{2000}/P_{1950}, is thus 5.89 times the growth in dividends, D_{2000}/D_{1950}. The S&P earnings–price ratio, Y_t/P_t, falls from 13.39 percent at the end of 1950 to 3.46 percent at the end of 2000, so the percent capital gain of the last 50 years is 3.87 times the percent growth in earnings. (Interestingly, almost all of the excess capital gain occurs in the last 20 years; Figure 1 shows that the 1979 earnings–price ratio, 13.40 percent, is nearly identical to the 13.39 percent value of 1950.)

All valuation models say that D_t/P_t and E_t/P_t are driven by expected future returns (discount rates) and expectations about future dividend and earnings growth. Our evidence suggests that rational forecasts of long-term dividend and earnings growth rates are not unusually high in 2000. We conclude that the large spread of capital gains for 1951 to 2000 over dividend and earnings growth is largely due to a decline in the expected stock return.

Some of the decline in D_t/P_t and E_t/P_t during 1951 to 2000 is probably anticipated in 1950. The dividend–price ratio for 1950, 7.18 percent, is high (Figure 1). The average for 1872 to 2000 is 4.64 percent. If D_t/P_t is mean-reverting, the expectation in 1950 of the yield in 2000 is close to the unconditional mean, say 4.64 percent. The actual dividend–price ratio for 2000 is

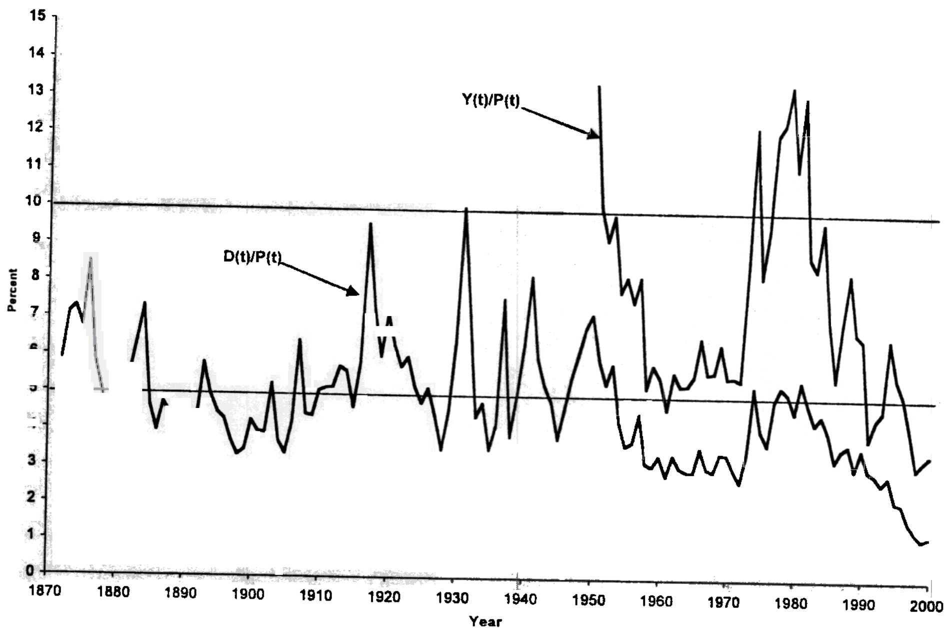


Figure 1. Dividend–price and earnings–price ratios.

1.22 percent. The 2000 stock price is thus 4.64/1.22 = 3.80 times what it would be if the dividend yield for 2000 hit the historical mean. Roughly speaking, this unexpected capital gain adds about 2.67 percent to the compound annual return for 1951 to 2000.

Similarly, part of the large difference between the 1951 to 2000 capital gain and the growth in earnings is probably anticipated in 1950. The 13.39 percent value of Y_t/P_t in 1950 is high relative to the mean for 1951 to 2000, 7.14 percent. If the earnings–price ratio is stationary, the expectation in 1950 of Y_t/P_t for 2000 is close to the unconditional mean, say 7.14 percent. The actual Y_t/P_t for 2000 is 3.46 percent. Thus, the 2000 stock price is 7.14/3.46 = 2.06 times what it would be if the ratio for 2000 hit the 7.14 percent average value for 1951 to 2000. Roughly speaking, this estimate of the unexpected capital gain adds about 1.45 percent to the compound annual return for the 50-year period.

In short, the percent capital gain for 1951 to 2000 is several times the growth of dividends or earnings. The result is historically low dividend–price and earnings–price ratios at the end of the period. Since the ratios are high in 1950, some of their subsequent decline is probably expected, but much of it is unexpected. Given the evidence that rational forecasts of long-term growth rates of dividends and earnings are not high in 2000, we conclude that the unexpected capital gains for 1951 to 2000 are largely due to a decline in the discount rate. In other words, the low end-of-sample price ratios imply low (rationally assessed, or true) expected future returns.

Like us, Campbell (1991), Cochrane (1994), and Campbell and Shiller (1998) find that, for recent periods, dividend and earnings growth are largely unpredictable, so variation in dividend–price and earnings–price ratios is largely due to the expected stock return. The samples in Campbell (1991) and Cochrane (1994) end in 1988 (before the strong subsequent returns that produce sharp declines in the price ratios), and they focus on explaining, in general terms, how variation in D_t/P_t splits between variation in the expected stock return and expected dividend growth. Campbell and Shiller (1998) focus on the low expected future returns implied by the low price ratios of recent years.

In contrast, we are more interested in what the decline in the price ratios says about past returns, specifically, that the average return for 1951 to 2000 is above the expected return. And this inference does not rest solely on the information in price ratios. We buttress it with two types of novel evidence. (a) The perspective from our long sample period that, although the average stock return for 1951 to 2000 is much higher than expected return estimates from fundamentals, the two approaches produce similar estimates for 1872 to 1950. (b) Evidence from Sharpe ratios, the book-to-market ratio, and the income return on investment, which also suggests that the average return for 1951 to 2000 is above the expected value.

III. Estimating the Expected Stock Return: Issues

There are two open questions about our estimates of the expected stock return. (a) In recent years the propensity of firms to pay dividends declines and stock repurchases surge. How do these changes in dividend policy affect our estimates of the expected return? (b) Under rather general conditions, the dividend and earnings growth models (2) and (3) provide estimates of the expected stock return. Are the estimates biased and does the bias depend on the return horizon? This section addresses these issues.

A. Repurchases and the Declining Incidence of Dividend Payers

Share repurchases surge after 1983 (Bagwell and Shoven (1989) and Dunsby (1995)), and, after 1978, the fraction of firms that do not pay dividends steadily increases (Fama and French (2001)). More generally, dividends are a policy variable, and changes in policy can raise problems for estimates of the expected stock return from the dividend growth model. There is no problem in the long-term, as long as dividend policies stabilize and the dividend–price ratio resumes its mean-reversion, though perhaps to a new mean. (An Appendix, available on request, provides an example involving repurchases.) But there can be problems during transition periods. For example, if the fraction of firms that do not pay dividends steadily increases, the market dividend–price ratio is probably nonstationary; it is likely to decline over time, and the dividend growth model is likely to underestimate the expected stock return.

Fortunately, the earnings growth model is not subject to the problems posed by drift in dividend policy. The earnings growth model provides an estimate of the expected stock return when the earnings–price ratio is stationary. And as discussed earlier, the model provides an estimate of the average expected return during the sample period when there are permanent shifts in the expected value of Y_t/P_t, as long as the ratio mean-reverts within regimes.

The earnings growth model is not, however, clearly superior to the dividend growth model. The standard deviation of annual earnings growth rates for 1951 to 2000 (13.79 percent, versus 5.09 percent for dividends) is similar to that of capital gains (16.77 percent), so much of the precision advantage of using fundamentals to estimate the expected stock return is lost. We see next that the dividend growth model has an advantage over the earnings growth model and the average stock return if the goal is to estimate the long-term expected growth of wealth.

B. The Investment Horizon

The return concept in discrete time asset pricing models is a one-period simple return, and our empirical work focuses on the one-year return. But many, if not most, investors are concerned with long-term returns, that is, terminal wealth over a long holding period. Do the advantages and disadvantages of different expected return estimates depend on the return horizon? This section addresses this question.

B.1. The Expected Annual Simple Return

There is downward bias in the estimates of the expected annual simple return from the dividend and earnings growth models—the result of a variance effect. The expected value of the dividend growth estimate of the expected return, for example, is the expected value of the dividend yield plus the expected value of the annual simple dividend growth rate. The expected annual simple return is the expected value of the dividend yield plus the expected annual simple rate of capital gain. If the dividend–price ratio is stationary, the compound rate of capital gain converges to the compound dividend growth rate as the sample period increases. But because the dividend growth rate is less volatile than the rate of capital gain, the expected simple dividend growth rate is less than the expected simple rate of capital gain.

The standard deviation of the annual simple rate of capital gain for 1951 to 2000 is 3.29 times the standard deviation of the annual dividend growth rate (Table I). The resulting downward bias of the average dividend growth rate as an estimate of the expected annual simple rate of capital gain is roughly 1.28 percent per year (half the difference between the variances of the two growth rates). Corrected for this bias, the dividend growth estimate of the equity premium in the simple returns of 1951 to 2000 rises from 2.55 to 3.83 percent (Table IV), which is still far below the estimate from the average return, 7.43 percent. Since the earnings growth rate and the annual rate of capital gain have similar standard deviations for 1951 to 2000,

Table IV

Estimates of the Real Equity Premium in Simple Annual and Long-term Returns: 1951 to 2000

The inflation rate for year t is $Inf_t = L_t/L_{t-1}$, where L_t is the price level at the end of year t. The real return for year t on six-month (three-month for the year 2000) commercial paper (rolled over at midyear) is F_t. The nominal value of the S&P index at the end of year t is p_t. Nominal S&P dividends and earnings for year t are d_t and y_t. Real rates of growth of dividends, earnings, and the stock price are $GD_t = (d_t/d_{t-1})*(L_{t-1}/L_t) - 1$, $GY_t = (y_t/y_{t-1})*(L_{t-1}/L_t) - 1$, and $GP_t = (p_t/p_{t-1})*(L_{t-1}/L_t) - 1$. The real dividend yield is $D_t/P_{t-1} = (d_t/p_{t-1})*(L_{t-1}/L_t)$. The dividend growth estimate of the real S&P return for t is $RD_t = D_t/P_{t-1} + GD_t$, the earnings growth estimate is $RY_t = D_t/P_{t-1} + GY_t$, and R_t is the realized real S&P return. The dividend and earnings growth estimates of the real equity premium for year t are $RXD_t = RD_t - F_t$ and $RXY_t = RY_t - F_t$, and $RX_t = R_t - F_t$ is the real equity premium from the realized real return. The average values of the equity premium estimates are $A(RXD_t)$, $A(RXY_t)$, and $A(RX_t)$. The first column of the table shows unadjusted estimates of the annual simple equity premium. The second column shows bias-adjusted estimates of the annual premium. The bias adjustment is one-half the difference between the variance of the annual rate of capital gain and the variance of either the dividend growth rate or the earnings growth rate. The third column shows bias-adjusted estimates of the expected equity premium relevant if one is interested in the long-term growth rate of wealth. The bias adjustment is one-half the difference between the variance of the annual dividend growth rate and the variance of either the growth rate of earnings or the rate of capital gain. The equity premiums are expressed as percents.

		Bias-adjusted	
	Unadjusted	Annual	Long-term
$A(RXD_t)$		3.83	
$A(RXY_t)$		4.78	
$A(RX_t)$		7.43	

13.79 percent and 16.77 percent (Table I), the bias of the earnings growth estimate of the expected return is smaller (0.46 percent). Corrected for bias, the estimate of the equity premium for 1951 to 2000 from the earnings growth model rises from 4.32 to 4.78 percent (Table IV), which again is far below the 7.43 percent estimate from the average return.

B.2. Long-term Expected Wealth

The (unadjusted) estimate of the expected annual simple return from the dividend growth model is probably the best choice if we are concerned with the long-term expected wealth generated by the market portfolio. The annual dividend growth rates of 1951 to 2000 are essentially unpredictable. If the dividend growth rate is serially uncorrelated, the expected value of the compounded dividend growth rate is the compounded expected simple growth rate:

$$E\left[\prod_{t=1}^{T}(1 + GD_t)\right] = [1 + E(GD)]^T. \tag{5}$$

And if the dividend–price ratio is stationary, for long horizons the expected compounded dividend growth rate is the expected compounded rate of capital gain:

$$E\left[\prod_{t=1}^{T}(1+GD_t)\right] = E\left[\prod_{t=1}^{T}(1+GP_t)\right]. \tag{6}$$

Thus, when the horizon T is long, compounding the true expected annual simple return from the dividend growth model produces an unbiased estimate of the expected long-term return:

$$[1+E(RD)]^T = E\left[\prod_{t=1}^{T}(1+R_t)\right]. \tag{7}$$

In contrast, if the dividend growth rate is unpredictable and the dividend–price ratio is stationary, part of the higher volatility of annual rates of capital gain is transitory, the result of a mean-reverting expected annual return (Cochrane (1994)). Thus, compounding even the true unconditional expected annual simple return, $E(R)$, yields an upward biased measure of the expected compounded return:

$$[1+E(R)]^T > E\left[\prod_{t=1}^{T}(1+R_t)\right] \tag{8}$$

There is a similar problem in using the average (simple) earnings growth rate to estimate long-term expected wealth. The regressions in Table III suggest that the predictability of earnings growth for 1951 to 2000 is due to transitory variation in earnings. As a result, annual earnings growth is 2.71 times more volatile than dividend growth (Table I). The compound growth rate of earnings for 1951 to 2000, 1.89 percent, is 2.05 times the compound dividend growth rate, 0.92 percent. But because earnings are more volatile, the average simple growth rate of earnings, 2.82 percent, is 2.69 times the average simple growth rate of dividends, 1.05 percent. As a result, the average simple growth rate of earnings produces an upward biased estimate of the compound rate of growth of long-term expected wealth.

We can correct the bias by subtracting half the difference between the variance of earnings growth and the variance of dividend growth (0.82 percent) from the average earnings growth rate. The estimate of the expected rate of capital gain provided by this adjusted average growth rate of earnings is 2.00 percent per year. Using this adjusted average growth rate of earnings, the earnings growth estimate of the expected real stock return for 1951 to 2000 falls from 6.51 to 5.69 percent. The estimate of the equity premium falls from 4.32 to 3.50 percent (Table IV), which is closer to the 2.55 percent obtained when the average dividend growth rate is used to estimate the expected rate of capital gain. Similarly, adjusting for the effects of transitory return volatility causes the estimate of the equity premium from realized stock returns to fall from 7.43 to 6.16 percent, which is still far above the bias-adjusted estimate of the earnings growth model (3.50 percent) and the estimate from the dividend growth model (2.55 percent).

Finally, we only have estimates of the expected growth rates of dividends and earnings and the expected rate of capital gain. Compounding estimates rather than true expected values adds upward bias to measures of expected long-term wealth (Blume (1974)). The bias increases with the imprecision of the estimates. This is another reason to favor the more precise estimate of the expected stock return from the dividend growth model over the earnings growth estimate or the estimate from the average stock return.

IV. Conclusions

There is a burgeoning literature on the equity premium. Our main additions are on two fronts. (a) A long (1872 to 2000) perspective on the competing estimates of the unconditional expected stock return from fundamentals (the dividend and earnings growth models) and the average stock return. (b) Evidence (estimates of precision, Sharpe ratios, and the behavior of the book-to-market ratio and the income return on investment) that allows us to choose between the expected return estimates from the two approaches.

Specifically, the dividend growth model and the realized average return produce similar real equity premium estimates for 1872 to 1950, 4.17 percent and 4.40 percent. For the half-century from 1951 to 2000, however, the equity premium estimates from the dividend and earnings growth models, 2.55 percent and 4.32 percent, are far below the estimate from the average return, 7.43 percent.

We argue that the dividend and earnings growth estimates of the equity premium for 1951 to 2000 are closer to the true expected value. This conclusion is based on three results.

(a) The estimates from fundamentals, especially the estimate from the dividend growth model, are more precise; they have lower standard errors than the estimate from the average return.

(b) The appealing message from the dividend and earnings growth models is that aggregate risk aversion (as measured by the Sharpe ratio for the equity premium) is on average roughly similar for the 1872 to 1949 and 1950 to 1999 periods. In contrast, the Sharpe ratio for the equity premium from the average return just about doubles from the 1872 to 1950 period to the 1951 to 2000 period.

(c) Most important, the average stock return for 1951 to 2000 is much greater than the average income return on book equity. Taken at face value, this says that investment during the period is on average unprofitable (its expected return is less than the cost of capital). In contrast, the lower estimates of the expected stock return from the dividend and earnings growth models are less than the income return on investment, so the message is that investment is on average profitable. This is more consistent with book-to-market ratios that are rather consistently less than one during the period.

If the average stock return for 1951 to 2000 exceeds the expected return, stocks experience unexpected capital gains. What is the source of the gains? Growth rates of dividends and earnings are largely unpredictable, so there is no basis for extrapolating unusually high long-term future growth. This leaves a decline in the expected stock return as the prime source of the unexpected capital gain. In other words, the high return for 1951 to 2000 seems to be the result of low expected future returns.

Many papers suggest that the decline in the expected stock return is in part permanent, the result of (a) wider equity market participation by individuals and institutions, and (b) lower costs of obtaining diversified equity portfolios from mutual funds (Diamond (1999), Heaton and Lucas (1999), and Siegel (1999)). But there is also evidence that the expected stock return is slowly mean reverting (Fama and French (1989) and Cochrane (1994)). Moreover, there are two schools of thought on how to explain the variation in expected returns. Some attribute it to rational variation in response to macroeconomic factors (Fama and French (1989), Blanchard (1993), and Cochrane (1994)), while others judge that irrational swings in investor sentiment are the prime moving force (e.g., Shiller (1989)). Whatever the story for variation in the expected return, and whether it is temporary or partly permanent, the message from the low end-of-sample dividend–price and earnings–price ratios is that we face a period of low (true) expected returns.

Our main concern, however, is the unconditional expected stock return, not the end-of-sample conditional expected value. Here there are some nuances. If we are interested in the unconditional expected annual simple return, the estimates for 1951 to 2000 from fundamentals are downward biased. The bias is rather large when the average growth rate of dividends is used to estimate the expected rate of capital gain, but it is small for the average growth rate of earnings. On the other hand, if we are interested in the long-term expected growth of wealth, the dividend growth model is probably best, and the average stock return and the earnings growth estimate of the expected return are upward biased. But our bottom line inference does not depend on whether one is interested in the expected annual simple return or long-term expected wealth. In either case, the bias-adjusted expected return estimates for 1951 to 2000 from fundamentals are a lot (more than 2.6 percent per year) lower than bias-adjusted estimates from realized returns. (See Table IV.) Based on this and other evidence, our main message is that the unconditional expected equity premium of the last 50 years is probably far below the realized premium.

REFERENCES

Bagwell, Laurie S., and John B. Shoven, 1989, Cash distributions to shareholders, *Journal of Economic Perspectives* 3, 129–149.

Blanchard, Olivier J., 1993, Movements in the equity premium, *Brooking Papers on Economic Activity* 2, 75–138.

Blume, Marshall, 1974, Unbiased estimators of long-run expected rates of return, *Journal of the American Statistical Association* 69, 634–638.
Brown, Stephen J., William N. Goetzmann, and Stephen Ross, 1995, Survival, *Journal of Finance* 50, 853–873.
Campbell, John Y., 1991, A variance decomposition for stock returns, *Economic Journal* 101, 157–179.
Campbell, John Y., and Robert J. Shiller, 1989, The dividend–price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.
Campbell, John Y., and Robert J. Shiller, 1998, Valuation ratios and the long-run stock market outlook, *Journal of Portfolio Management* 24, 11–26.
Claus, James, and Jacob Thomas, 2001, Equity premia as low as three percent? Evidence from analysts' earnings forecasts for domestic and international stock markets, *Journal of Finance* 56, 1629-1666.
Cochrane, John, 1991, Volatility tests and efficient markets: A review essay, *Journal of Monetary Economics* 27, 463–487.
Cochrane, John H., 1994, Permanent and transitory components of GNP and stock prices, *Quarterly Journal of Economics* 109, 241–265.
Davis, James L., Eugene F. Fama, and Kenneth R. French, 2000, Characteristics, covariances, and average returns, *Journal of Finance* 55, 389–406.
Diamond, Peter A., 1999, What stock market returns to expect for the future? Center of Retirement Research at Boston College, September, Number 2.
Dunsby, Adam, 1995, *Share Repurchases, Dividends, and Corporate Distribution Policy* (Ph.D. thesis, The Wharton School, University of Pennsylvania).
Fama, Eugene F., and Kenneth R. French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–25.
Fama, Eugene F., and Kenneth R. French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.
Fama, Eugene F., and Kenneth R. French, 2001, Disappearing dividends: Changing firm characteristics or lower propensity to pay, *Journal of Financial Economics* 60, 3–43.
Gebhardt, William R., Charles M. C. Lee, and Bhaskaram Swaminathan, 2001, Toward an implied cost of capital, *Journal of Accounting Research* 39, June, 135–176.
Gordon, Myron, 1962, *The Investment Financing and Valuation of the Corporation* (Irwin, Homewood, IL).
Heaton, John, and Deborah Lucas, 1999, Stock prices and fundamentals, in Ben Bernanke and Julio Rotemberg, eds.: *Macroeconomics Annual 1999* (National Bureau of Economic Research, MIT Press, Cambridge, MA).
Keane, Michael P., and David E. Runkle, 1998, Are financial analysts' forecasts of corporate profits rational? *Journal of Political Economy* 106, 768–805.
Lamont, Owen, 1998, Earnings and expected returns, *Journal of Finance* 53, 1563–1587.
Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849.
Mehra, Rajnish, and Edward Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145–161.
Shiller, Robert, 1989, *Market Volatility* (MIT Press, Cambridge, MA).
Siegel, Jeremy J., 1999, The shrinking equity premium, *Journal of Portfolio Management* 26, 10–17.
Vuolteenaho, Tuomo, 2000, Understanding the aggregate book-to-market ratio, Manuscript, University of Chicago.

September 2002

Forthcoming in *Journal of Applied Corporate Finance*

GLOBAL EVIDENCE ON THE EQUITY RISK PREMIUM

Elroy Dimson, Paul Marsh and Mike Staunton*

London Business School

One of the most important contemporary issues in corporate finance is the magnitude of the equity risk premium. The risk premium is the incremental return that shareholders require from holding risky equities rather than risk-free securities. The risk premium drives future equity returns and is the key determinant of the cost of capital.

Today, investors have more cause than ever to ask what returns they can expect from equities, and what the future risk-reward tradeoff is likely to be. Companies also need to answer this question in order to understand what returns their shareholders require from projects of differing risk. Regulators, too, need to know the cost of capital in order to set 'fair' rates of return for regulated industries.

This paper sheds light on this important issue by addressing two key questions: What has the size of the equity risk premium been historically? And what can we expect for the future? To answer these questions, we need to look at long periods of capital market history, and extend our horizons beyond just the United States. In this paper, we therefore present evidence for sixteen different countries over the 102-year period from 1900–2001.

The need for a long-run perspective

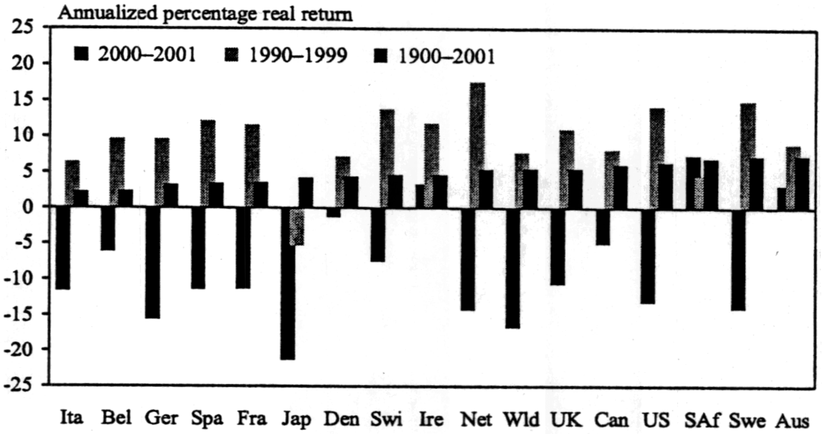
The need for a long-run perspective, and the dangers of focusing just on recent stock market history, are easily demonstrated. Over the last decade of the twentieth century, US equity investors more than trebled their initial stake. In real terms, they achieved a total return (capital gain plus reinvested dividends) of 14.2 percent per annum. During the last five years of the1990s, US equities achieved high returns in every year, varying from a low of twenty-one percent in 1996 to a high of thirty-six percent in 1995. Many investors became convinced that high corporate growth rates could be extrapolated into the indefinite future. With steady growth rates, equity risk appeared lower. Simultaneously, there appeared to be a decline in the premium sought by investors to compensate for exposure to equity market risk. This drove stock prices onward and upward. Surveys suggested that, in consequence, many investors expected long-run stock market returns to continue at double-digit percentage rates of return.

Then the technology bubble burst. Growth projections had been unrealistic. High growth expectations were seen to be associated with high risk. Investors demanded a larger reward for equity market risk exposure. Stock prices fell in 2000 and then again in 2001, with no respite yet in 2002. With markets having fallen, investors started to project lower returns for the future.

---

* This paper draws on, extends, and updates the research that underpinned our recent book, "*Triumph of the Optimists: 101 Years of Global Investment Returns*" (New Jersey: Princeton University Press, 2002). We are very grateful to ABN Amro for their extensive support and to our many international data contributors—too numerous to mention here, but all of whom are listed and cited in "Triumph". We are also grateful for the many helpful comments received from participants at numerous academic and practitioner seminars held around the world.

FIGURE 1
SHORT-TERM AND LONG-RUN REAL RETURNS ON EQUITIES AROUND THE WORLD*



* The country names listed in abbreviated form along the horizontal axis are (from left to right) Italy, Belgium, Germany, Spain, France, Japan, Denmark, Switzerland, Ireland, The Netherlands, the world (the weighted average of the sixteen individual countries), The United Kingdom, Canada, The United States, South Africa, Sweden, and Australia.

Yet it is dangerous to overreact to recent stock market performance. It would be wrong for investors to conclude that just because equities have delivered a low return since New Year 2000 that there has been either a substantial fall, or indeed rise, in the long-term expected equity premium.

Figure 1 shows how US equity returns compared with those in fifteen other countries and the world index. The black bars show annualized equity returns over 2000–01. In most countries, equities suffered negative returns, underperforming bonds everywhere except Ireland, and falling short of bill returns everywhere except Australia, Ireland, and South Africa. Estimating the expected risk premium from the performance of equities relative to bills or bonds over this period would clearly be nonsense. Investors cannot have required or expected a negative return for assuming risk. Instead, this was simply a very disappointing period for equities.

But while the opening years of the twenty-first century (fortunately) do not provide a basis for generalising about future returns, looking back at the previous decade only confuses the picture. Indeed, it would be equally misleading to estimate future risk premia from data for 1990–99. The light blue bars in Figure 1 show that over this period, equity returns (except in Japan and South Africa) were high. The 1990s was a golden age for stocks, and golden ages, by definition, recur infrequently.

To understand the risk premium—which is the principal objective of this paper—we need to examine periods that are much longer than one or two years, or even a decade. This is because stock markets are volatile, with much variation in year-to-year returns. In order to make inferences we thus need long time series that incorporate the bad times as well as the good. The dark blue bars in Figure 1 provide an insight into the perspective that longer periods of history can bring. These show real equity returns over the 102-year period from 1900–2001. Clearly, these 102-year returns are much less favourable than the returns during the 1990s, but equally, they contrast sharply with the disappointing returns over 2000–01.

Investors' judgements should thus be informed by the full extent of financial market history, and by looking not just at the United States, but at other countries as well.

Limitations of prior estimates of the risk premium

To be fair, financial economists do tend to measure the equity premium over quite long periods. Standard practice, however, draws heavily on the United States, with most textbooks citing only the US experience. By far the most widely cited US source prior to the end of the technology bubble was Ibbotson Associates[1], whose equity premium history starts in 1926. They estimated an annualized return on equities of 11.3 percent, and a risk-free return of 3.8 percent. This implied a geometric premium relative to bills of 7.3 percent (i.e., 1.113/1.038 = 1.073). References to other countries are few and far between, but a few textbooks also cite UK evidence. Before the publication of the research that underpins this paper, the most widely cited sources for the United Kingdom were the studies published by Barclays Capital and CSFB[2], which both started in 1919, and who published equity and risk-free returns of 12.2 and 5.5 percent, implying an annualized risk premium relative to bills of 6.4 percent.

In citing these estimates, financial economists are generally making the implicit assumption that provided the data are of sufficient quality, then the historical risk premium, measured over many decades, will provide an unbiased estimate of the future premium. Yet the twentieth century proved to be a period of remarkable growth in the US economy, and it seems probable that the outcome exceeded the expectations held in 1926 by US investors. Similar arguments apply to the United Kingdom, and the likely expectations of UK investors in 1919, but additionally, the UK evidence turned out to be based on a retrospectively constructed index whose composition, up to 1955, was tainted by survivor bias and narrow coverage.

In recent years, both practitioners and researchers have grown increasingly uneasy about these widely cited estimates, largely because they seem high. Apart from biases in index construction, the finger of suspicion has pointed mainly at success and survivorship bias. One influential study by Jorion and Goetzmann[3], for example, asserted, "the high equity premium obtained for US (and, by implication, UK) equities appears to be the exception rather than the rule" (parenthesis added). Recently, Zvi Bodie[4] argued that high US and UK premia are likely to be anomalous, and underlined the need for comparative international evidence. He pointed out that long-run studies are always of US or UK premia: "There were 36 active stock markets in 1900, so why do we only look at two? I can tell you—because many of the others don't have a 100-year history, for a variety of reasons." This paper helps fill this gap in our knowledge by providing a 102-year back-history of risk premia for sixteen of these markets.

NEW EVIDENCE

The new evidence on long-run risk premia presented in this paper is derived from a unique new database of long-run international returns. This comprises annual returns on stocks, bonds, bills, inflation, and currencies for sixteen countries from 1900–2001. The countries include the two main North American markets, namely, the United States and Canada, the United Kingdom, seven markets from what is now the Euro currency area, three other

1. See Ibbotson Associates, 2000, *Stocks, Bonds, Bills and Inflation Yearbook*, Chicago, Ibbotson Associates

2. Barclays Capital, 1999, *Equity-Gilt Study*, London: Barclays Capital; and Credit Suisse First Boston, 1999, *The CSFB Equity-Gilt Study*, London: Credit Suisse First Boston.

3. Jorion, P. and W. Goetzmann, "Global Stock Markets in the Twentieth Century", *Journal of Finance*, Vol. 54, 1999, pp. 953-80.

4. Bodie, Z, "Longer time horizon 'does not reduce risk'" *Financial Times*, 26 January 2002.

European markets, two Asia-Pacific markets, and one African market. Together, these countries made up 95 percent of the free float market capitalization of all world equities at start-2002, and we estimate that they comprised over 90 percent by value at the start of our period in 1900.

To compile this database, we assembled the best quality indices and returns data available for each national market from previous studies and other sources[5]. Where possible, we used data from peer-reviewed academic papers, although some studies were previously unpublished. To span the full period from 1900 onward, we typically linked more than one index series. For our own home market, the UK, we constructed our own indices, since hitherto there was no satisfactory record of long run returns. For the period since 1955, we used the London Business School Share Price Database to construct an index covering the entire UK equity market[6]. From 1900–55, we constructed an index of the performance of the largest 100 companies by a process of painstaking financial archaeology, collecting data from archives in the City of London. We also used archive data to construct indices for several other countries (e.g., Canada, Ireland, South Africa) for periods for which no data was previously available.

Unlike most previous long-term studies of global markets, all our investment returns include reinvested gross income as well as capital gains. Many early equity indices measure just capital gains, ignoring dividends, thereby introducing serious downward bias. Similarly, many early bond indices record just yields, ignoring price movements. Our database is thus more comprehensive and accurate than previous research, spans a longer period, and the common start-date of 1900 aids international comparisons. We can now set the US risk premia data alongside comparable 102-year risk premia series for fifteen other countries, and make international comparisons that help set the US experience in perspective.

Table 1 shows the historical equity risk premia for the sixteen countries over the 102-year period 1900–2001. We also display equity premia for the world, based on our world equity index. The latter comprises a sixteen-country, common-currency (here taken as US dollars) equity index in which each country is weighted by its start-year market capitalization or (in earlier years) its GDP[7]. The left-hand half of Table 1 shows equity premia measured relative to the return on treasury bills (or the nearest equivalent short-term instrument); the right-hand half shows premia calculated relative to the return on long-term government bonds. Since the world index is computed here from the perspective of a US (dollar) investor, the world equity risk premium relative to bills is calculated relative to the US risk free (i.e., treasury bill) rate. The world equity premium relative to bonds is calculated relative to a GDP-weighted, sixteen-country, common-currency (here taken as US dollars) world bond index.

In each half of the table we show three measures. These are, first, the geometric mean risk premium, namely, the annualized premium over the entire 102 years; second, the arithmetic mean of the 102 one-year premia; and third, the standard deviation of the 102 one-year premia. While the United States and the United Kingdom have indeed performed well, compared to other markets there is no indication that they are hugely out of line.

---

5. Details of our data sources for all sixteen countries together with full citations are provided in Dimson, E, P R Marsh, and M Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press, 2002.

6. Dimson, E and P R Marsh, "UK Financial Market Returns 1955-2000", *Journal of Business*, Vol. 74, pp. 1–31.

7. We use market capitalization weights from 1968 onward and GDP (gross domestic product) weights before then due to the lack of reliable comprehensive data on country capitalizations prior to that date.

TABLE 1
EQUITY RISK PREMIA AROUND THE WORLD 1900–2001

Country	Relative to bills: Geometric mean	Relative to bills: Arithmetic mean	Relative to bills: SD	Relative to bonds: Geometric mean	Relative to bonds: Arithmetic mean	Relative to bonds: SD
Australia	7.0	8.5	17.2	6.3	7.9	18.8
Belgium	2.7	5.0	23.5	2.8	4.7	20.7
Canada	4.4	5.7	16.7	4.2	5.7	17.9
Denmark	1.6	3.2	19.4	1.8	3.1	16.9
France	7.1	9.5	23.9	4.6	6.7	21.7
Germany*	4.6	10.0	35.3	6.3	9.6	28.5
Ireland	3.4	5.3	20.5	3.1	4.5	17.3
Italy	6.6	10.6	32.5	4.6	8.0	30.1
Japan	6.4	9.6	27.9	5.9	10.0	33.2
The Netherlands	4.8	6.8	22.3	4.4	6.4	21.5
South Africa	6.1	8.2	22.4	5.4	7.1	19.6
Spain	3.1	5.2	21.4	2.2	4.1	20.2
Sweden	5.3	7.4	21.9	4.9	7.1	22.1
Switzerland*	4.0	5.8	19.6	2.4	3.9	18.0
United Kingdom	4.5	6.2	19.9	4.2	5.5	16.7
United States	5.6	7.5	19.7	4.8	6.7	20.0
World	4.6	5.9	16.5	4.3	5.4	14.6

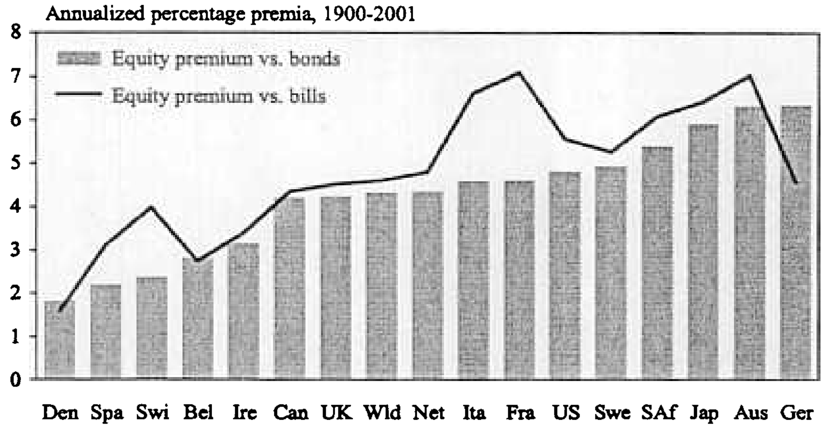
Equity risk premia (percent per year)

**Germany excludes 1922-23. Switzerland commences in 1911.*
Source: Dimson, Marsh, and Staunton, Triumph of the Optimists, Princeton University Press, 2002.

Over the entire 102-year period, the annualized equity risk premium, relative to bills, was 5.6 percent for the United States and 4.5 percent for the United Kingdom. Averaged across all sixteen countries, the risk premium relative to bills was 4.8 percent, while the risk premium on the world equity index was 4.6 percent. Relative to long bonds, the story is similar. The annualized US equity risk premium relative to bonds was 4.8 percent, and the corresponding figure for the United Kingdom was 4.2 percent. Across all sixteen countries, the risk premium relative to bonds averaged 4.3 percent, while for the world index it was also 4.3 percent.

The annualized equity risk premia are plotted in Figure 2. In this figure, countries are ranked by the equity premium relative to bonds, displayed as bars. The line-plot presents each country's risk premium relative to bills. It can be seen that the United States does indeed have a historical risk premium that is above the world average, but it is by no means the country with the largest recorded premium. The equity premium for the United Kingdom is closer to the worldwide average. While US and UK equities have performed well, both countries are towards the middle of the distribution of worldwide equity premia. Commentators have suggested that survivor bias may have given rise to equity premia for the United States and the United Kingdom that are unrepresentative. While legitimate, these concerns are somewhat overstated. Investors may not have been materially misled by a focus on the US and UK experiences. Rather, the critical factors are the period over which the risk premium is estimated, together with the quality of the index series.

FIGURE 2
WORLDWIDE ANNUALIZED EQUITY RISK PREMIA 1900–2001



Germany excludes 1922-23. Switzerland commences in 1911.
Source: Dimson, Marsh, and Staunton, Triumph of the Optimists, Princeton University Press, 2002.

Avoiding bias

There are noteworthy differences between the premia reported in this paper and those put forward, prior to publication of our research, by Ibbotson Associates in the United States, and by Barclays Capital and CSFB in the United Kingdom. Indeed, the premia estimated in this paper are around 1½ percent lower than those reported in these earlier studies. The differences arise from previous biases in index construction for the United Kingdom and, for both countries, from the choice of time frame, which in our case extends back to 1900[8]. We thus include the pre-1926 period for the United States (and pre-1919 for the United Kingdom) when returns were lower, partly due to events in the period leading up to, and including, World War I. Moreover, as noted above, prior perceptions about the risk premium have been dominated by the widely cited US estimates. Yet Table 1 and Figure 2 show that the premia for two-thirds of the other countries in our sample were lower than for the United States[9].

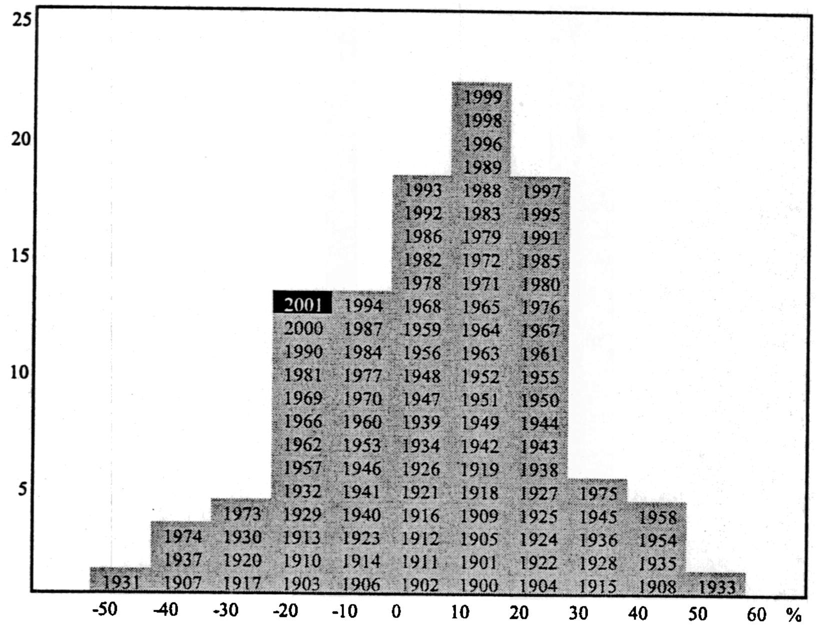
It is thus clear that the 102-year historical estimates of equity premia reported here are lower than was previously thought and other studies suggest. Even then, however, the historical record may overstate expectations. First, even if we have been successful in avoiding survivor bias within each index, we still focus on markets that survived, omitting countries such as Poland, Russia or China whose compound rate of return was –100 percent. Although these markets were relatively small in 1900[10], their omission probably leads to an overestimate of the

8. Interestingly, after publication of our research, Barclays Capital (but not CSFB) corrected their pre-1955 estimates of UK equity returns for bias and extended their index series back to 1900.

9. Table 1 shows that the annualized world equity risk premium relative to bills was 4.6 percent compared with 5.6 percent for the United States. Part of this difference, however, reflects the strength of the dollar over the period 1900–2001. The world risk premium is computed here from the world equity index expressed in dollars, in order to reflect the perspective of a US-based global investor. Since the currencies of most other countries depreciated against the dollar over the twentieth century, this lowers our estimate of the world equity risk premium relative to the (weighted) average of the local-currency based estimates for individual countries.

10. See Rajan, R and L Zingales, "The Great Reversals: The Politics of Financial Development in the 20th Century", Working paper No. 8178, Cambridge MA: National Bureau of Economic Research and Dimson, E, P R Marsh, and M Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press, 2002.

FIGURE 3
HISTOGRAM OF US EQUITY RISK PREMIUM RELATIVE TO BILLS, 1900–2001



Source: Dimson, Marsh and Staunton, Triumph of the Optimists, Princeton University Press, 2002

worldwide risk premium.[11] Second, our premia are estimated relative to bills and bonds, which in a number of countries gave markedly negative real returns. Since these "risk-free" returns likely fell below investors' expectations, the corresponding equity premia are probably overstated.[12]

Although there is room for debate, we do not consider market survivorship to be the most important source of bias when inferring expected premia from the historical record. There are cogent arguments for suggesting that investors expected a lower premium than they actually received. However, this is more to do with a failure to fully anticipate improvements in business and investment conditions during the second half of the last century, an issue that we will return to below.

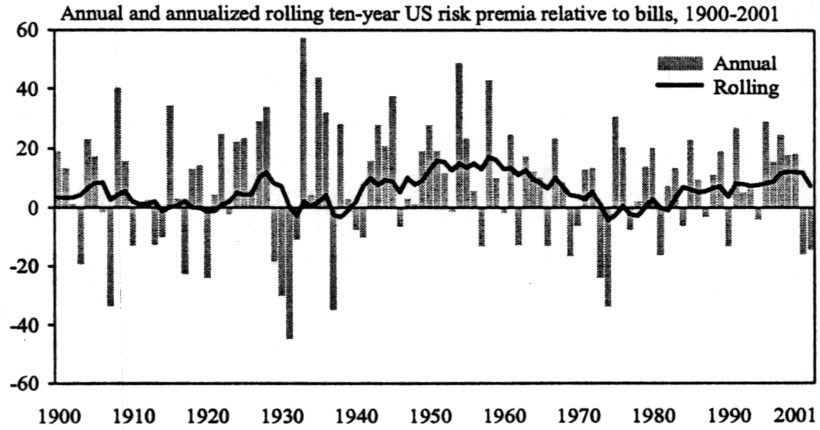
VARIATION IN RISK PREMIA OVER TIME

The historical equity premia shown in Figure 2 are the geometric means of 102 separate one-year premia that vary a great deal. In Figure 3 we show the year-by-year premia on US equities relative to bills. The lowest excess return was –45 percent in 1931, when equities returned –44 percent and treasury bills 1.1 percent; the highest was 57 percent in 1933, when equities gave 57.6 percent and bills 0.3 percent. Figure 3 shows that, for the United States,

11. We say omitting non-surviving markets "probably" gives rise to overestimated risk premia because of the possibility that some defaulting countries have returns of −100 percent on bonds, while equities retain some residual value. For such countries, the ex post equity premium would be positive.

12. We again say low risk-free rates probably give rise to overstated risk premia because equity returns would presumably have been higher if economic conditions had not given rise to markedly negative real fixed-income returns. If economic conditions had been better, it is possible that the equity premium would then have been larger.

FIGURE 4
ROLLING AND ANNUAL TEN-YEAR US PREMIA RELATIVE TO BILLS, 1900–2001



Source: Dimson, Marsh and Staunton, Triumph of the Optimists, Princeton University Press, 2002

the distribution of annual excess returns is roughly symmetrical with a mean of 7.5 percent and a standard deviation of 19.7 percent. On average, therefore, US investors received a positive, and quite large, reward for exposure to equity market risk.

Because the range of excess returns encountered on a year-to-year basis is very broad, it can be misleading to label them "risk premia." As already noted, investors cannot have expected, let alone required, a negative risk premium from investing in equities, otherwise they would simply have avoided them. All the negative and many of the very low premia plotted in the histogram must therefore reflect nasty surprises. Equally, investors could not have required premia as high as 57 percent in 1933. Such numbers are implausible as a required reward for risk, and the high realizations must therefore reflect pleasant surprises. To avoid confusion, many writers choose not to refer to annual excess returns as "risk premia". They simply clarify that excess returns are ex post returns in excess of the risk free interest rate.

As we noted above, because one-year excess returns are so variable, we need to examine much longer periods, in the hope that good and bad luck might then cancel out. A common choice of time frame is a decade. In Figure 4, we show the US equity risk premium, measured over a sequence of rolling ten-year periods, superimposed on the annual returns since 1900.

Even over ten-year periods, the historical risk premium was sometimes negative, most recently in the 1970s and early 1980s. Again, since investors cannot have required a negative reward for risk, these must reflect unpleasant surprises. Figure 4 also reveals several cases of double-digit ten-year premia. These must have been pleasant surprises, as they are too high to reflect prior expectations. Clearly, a decade is still too short a period for good and bad luck to cancel out, and for drawing inferences about investors' expectations. Over a decade, like a single year, all we are plotting is the excess return that was realised over a period in the past.

Imprecise estimates

Prior to our research, studies for countries other than the United States and United Kingdom used the longest stock return series available, typically covering an interval of up to half a century. Sadly, even such a long research period does not yield an answer that is invariant to the choice of period. Taking the United Kingdom as an illustration, the arithmetic mean annual excess return for the first half of the twentieth century was only 3.1 percent, as compared to 9.2 percent from 1950 to date.

Even with a full century of data, market fluctuations have an impact. All we can state with confidence is what the excess return was in the past. This is why some writers restrict the term "risk premium" to denote the expected reward from equity investment. To avoid confusion, we make it clear when we are looking to the future by referring to the expected or "prospective" risk premium. When we measure the excess return over a period in the past we generally refer to this as the "historical" risk premium.

With 102 years of data, the potential inaccuracy in historical risk premia is high. The standard error measures this inaccuracy. It is approximately equal to one-tenth of the annual standard deviation of returns reported in Table 1. The standard error for the United States is 1.9 percent, and the range runs from 1.7 percent (Australia and Canada) to 3.5 percent (Germany). This means that while the US arithmetic mean premium (relative to bills) has a best estimate of 7.5 percent, we can be only two-thirds confident that the true mean lies within one standard error of this, namely within the range 7.5 ± 1.9 percent, or 5.6 to 9.4 percent. Similarly, there is a nineteen-out-of-twenty probability that the true mean lies within two standard errors, namely 7.5 ± 3.8 percent, or 3.7 to 11.3 percent. These high standard errors are why the longest possible series of stock market data should in general be used for estimating risk premia.

FROM THE PAST TO THE FUTURE

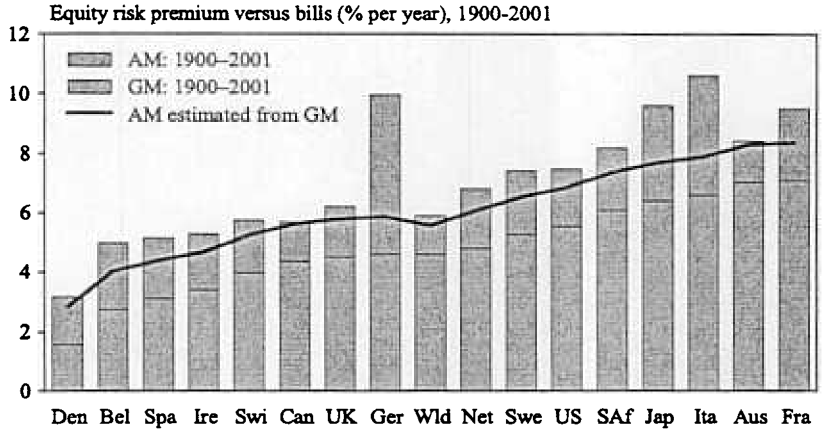
To estimate the equity risk premium to use in discounting future cash flows, we need the expected future risk premium, i.e., the arithmetic mean of the possible premia that may occur. Suppose the returns that may happen in the future are drawn from the same distribution as those that occurred in the past. If so, the expected risk premium is the arithmetic mean (or simple average) of the one-year historical premia. Whenever there is some variability in annual premia, the arithmetic mean will always exceed the geometric mean (or annualized) risk premium.[13]

In Figure 5, the full height of the bars shows the historical arithmetic mean premium relative to bills for each country. The US equity premium is 7.5 percent, while the world equity risk premium is 5.9 percent. The arithmetic mean premia are noticeably higher than the geometric mean premia shown by the light blue portion of each bar. They are at their largest (in both absolute terms and relative to the geometric mean) for the countries that experienced the greatest volatility of returns over the last century (see Table 1).

In looking to the future, let us assume for the moment that investors in each country expect the same annualized (geometric mean) risk premium as they have received in the past. The bar and line plots in Figure 5 can then be interpreted as forecasts of the prospective arithmetic risk premia under alternative assumptions about future volatility. If there were no volatility in future annual returns, the expected arithmetic risk premia would be equal to their (historical)

13. For example, the arithmetic mean of two equally likely returns of +25 percent and –20 percent is (+25 – 20)/2 = 2½ percent, while their geometric mean is zero since (1 + 25/100) x (1 – 20/100) – 1 = 0.

FIGURE 5
ARITHMETIC MEAN EQUITY RISK PREMIA RELATIVE TO BILLS, 1900–2001



Source: Dimson, Marsh and Staunton, Triumph of the Optimists, Princeton University Press, 2002

geometric mean premia shown by the height of the light blue portion of the bars in Figure 5. On the other hand, if future volatility were equal to the long-term historical volatility, the expected risk premia would be equal to the historical arithmetic mean risk premia, shown by the full height of the bars. However, the long-term historical standard deviation is a poor predictor of future volatility, especially since some sources of extreme volatility (such as hyperinflation) are unlikely to recur. We therefore need estimates of expected future risk premia that are conditional on current predictions for market volatility.

When returns are distributed lognormally, the geometric and arithmetic means are linked by the standard deviation (or volatility) of returns. We therefore estimate the expected future arithmetic mean premium for each country, replacing the historical difference between the arithmetic and geometric means with a difference based on contemporary risk estimates. For expositional simplicity, even though the volatility of one stock market is not in reality the same as another, we assume a current volatility level for all sixteen national markets of 16 percent, and for the world index of 14 percent. The resulting estimates of the arithmetic mean premia relative to bills are shown by the dark blue line-plot in Figure 5.

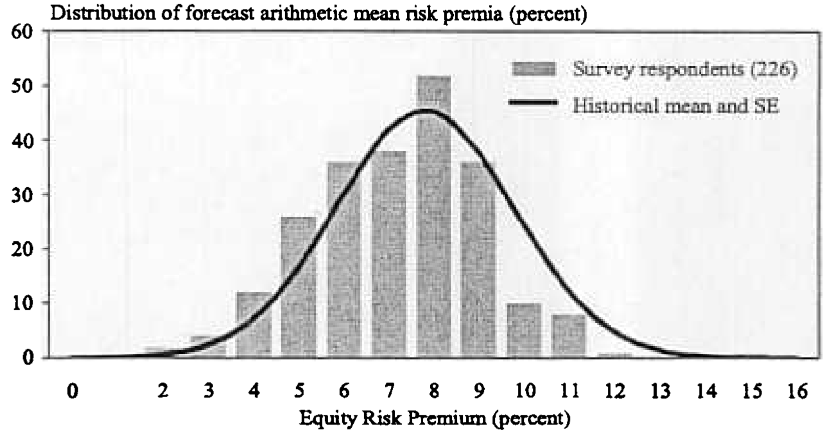
For those wishing to forecast future arithmetic mean risk premia by extrapolating from the long-run historical annualized premia, the premia illustrated by the line plot in Figure 5 are the ones to use. The historical equity risk premium, adjusted to current levels of market volatility, is estimated as 6.8 percent for the United States, and 5.6 percent for the world index.

THE EXPERTS' CONSENSUS

In refocusing on the expected future risk premium, however, we must do more than extrapolate from the past. The question of what equity premium we can expect has, for years, been a source of controversy. In late 1998 Ivo Welch studied the opinions of 226 financial economists who were asked to forecast the thirty-year arithmetic mean equity risk premium [14].

14. Welch, I, "Views of Financial Economists on the Equity Premium and Other Issues," *Journal of Business*, Vol. 73, 2000, pp. 501-537.

FIGURE 6
FINANCIAL ECONOMISTS' RISK PREMIUM FORECASTS AND MARKET HISTORY



The bars in Figure 6 show the distribution of the responses. The mean forecast was 7.1 percent; the median was 7.0 percent, and the range ran from 1 to 15 percent.

While the bars in Figure 6 show the distribution of survey responses, the curved line represents the normal distribution based on the mean over approximately a century and the associated standard error for the US equity risk premium. The spread in both distributions indicates that the uncertainty across financial experts about the risk premium is as large as the uncertainty that arises from statistical analysis of historical returns.

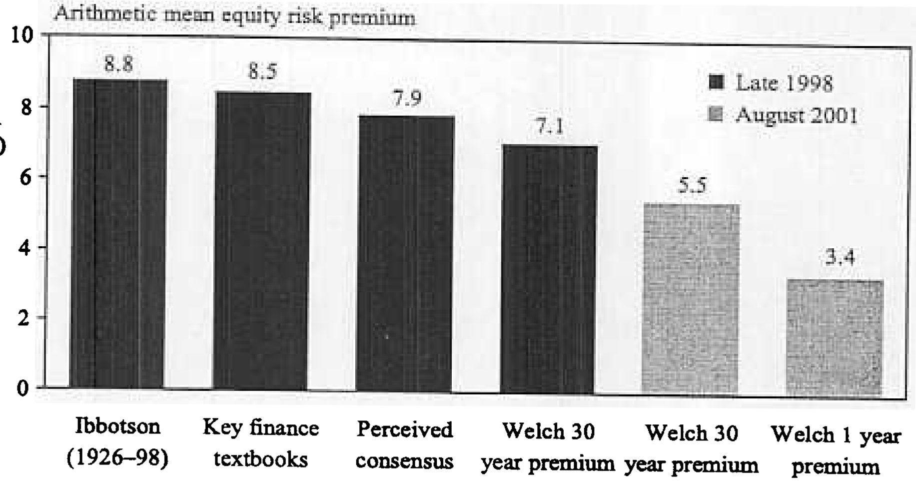
Most respondents to the Welch survey would have regarded the Ibbotson Associates yearbook as the definitive study of the historical US risk premium. The first bar of Figure 7 shows that the 1926-98 arithmetic risk premium computed from Ibbotson data was 8.8 percent per year. The second bar shows that the key finance textbooks were on average suggesting a premium of 8.5 percent, a little below the Ibbotson figure. The textbook authors may have based their views on earlier, slightly lower, Ibbotson estimates, or else they were shading the Ibbotson estimates downward. The Welch survey mean is in turn lower than the textbook figure, but since respondents claimed to lower their forecasts when the equity market rises, this difference may be attributed to the market's strong performance in the 1990s. Interestingly, the third and fourth bars of Figure 7 show that the survey respondents also perceived the profession's consensus to be higher than it really was. That is, they thought the mean was around 0.8 percent higher than the 7.1 percent average revealed in the survey.

These survey and textbook figures represent what was being taught at the end of the 1990s in the world's leading business schools and economics departments in the United States and around the world. As such, these estimates were also widely used by investors, finance professionals, corporate executives, regulators, lawyers and consultants. Their influence extended from the classroom to the dealing room, to the boardroom, and to the courtroom.

New opinions

Whether Welch's survey mean of 7.1 percent was appropriate is another matter. A large number of respondents were calibrating their forecasts relative to the longest-run historical benchmark available from Ibbotson, and then shading the historical number downward based on subjective factors, including their judgement of the impact of strong market performance in the late 1990s. By 2001, longer-term estimates of the US arithmetic mean equity premium

FIGURE 7
ESTIMATED ARITHMETIC RISK PREMIA RELATIVE TO BILLS, 1998 AND 2001



were gaining publicity. Including pre-1926 data, and extending the period through the start of the new millennium, the 1900-2000 mean premium was 1.1 percent lower than the Ibbotson estimate on the left-hand side of Figure 7. At the same time, survey respondents who sought to predict a premium below the consensus might have been encouraged by publication of the survey to further reduce their estimates.

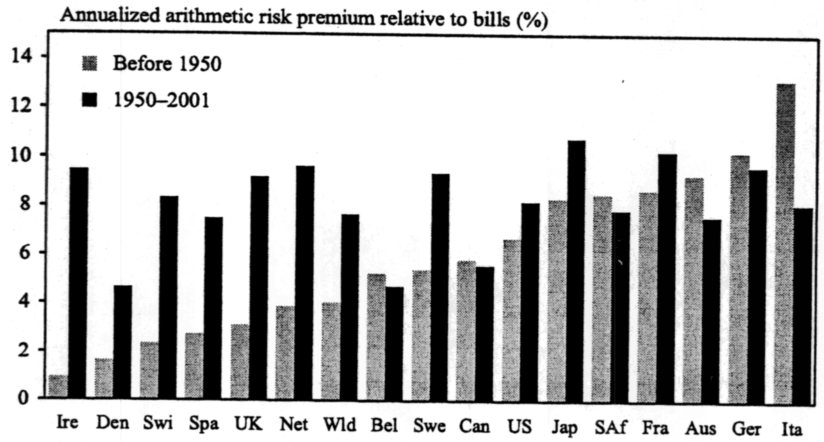
In August 2001, Welch updated his earlier survey, receiving responses from 510 finance and economics professors[15]. He found that respondents to the follow-up questionnaire had revised downward their estimates of the long-term arithmetic mean risk premium by an average of 1.6 percent. Over a thirty-year horizon they now estimated an equity premium averaging 5.5 percent, and over a one-year horizon, an equity premium averaging 3.4 percent (see Figure 7). The mean premia were the same for those who had previously participated in the earlier survey and those who were taking part for the first time. Although respondents to the earlier survey had indicated that, on average, a bear market would raise their equity premium forecast, Welch (2001) reports that "This is in contrast with the observed findings: it appears as if the recent bear market correlates with lower equity premium forecasts, not higher equity premium forecasts".

Predictions of the long-term equity premium should not be so sensitive to short-term stock market fluctuations, especially in the direction and magnitude revealed by Welch's follow-up survey in 2001. While it is possible that one-year required rates of return fluctuate markedly, it is unlikely that thirty-year expectations can be so volatile. The changing consensus may, however, reflect the new approaches to estimating the premium and /or new facts about long-term stock market performance, such as evidence that other countries have typically had historical premia that were lower than the United States.

---

15. Welch, I, "The Equity Premium Consensus Forecast Revisited," Working paper, Yale School of Management, September 2001.

FIGURE 8
PREMIA RELATIVE TO BILLS, FIRST 50 YEARS VERSUS THE NEXT 52 YEARS



Germany excludes 1922-23. Switzerland commences in 1911.
Source: Dimson, Marsh and Staunton, Triumph of the Optimists, Princeton University Press, 2002

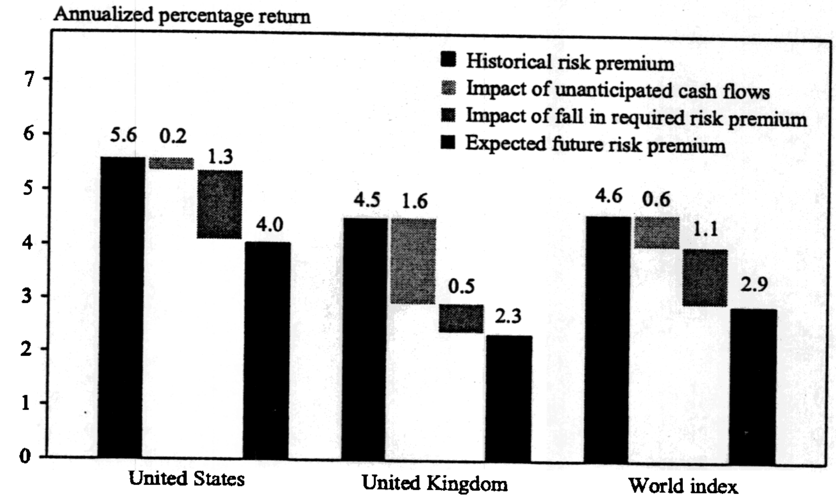
REVISITING HISTORY

The wide dispersion of estimates, together with the dramatic decline in the consensus premium between 1998 and 2001, reinforces the need to better understand the historical record. However, since history may have been kind to (or harsh on) stock market investors, there are coherent arguments for going beyond raw historical estimates. First, the whole idea of using the achieved risk premium to forecast the required risk premium depends on having a long enough period to iron out good and bad luck, yet as we noted earlier, even with 102 years of data our estimates are imprecise. Second, the expected equity risk premium could for good reasons vary over time. Third, we must take account of the fact that stock market outcomes are influenced by many factors, some of which (like removal of trade barriers) may be non-repeatable, which implies projections for the premium that deviate from the past.

A comparison between the first and second halves of our 102-year period makes the point. Over the first half of the twentieth century, the arithmetic average world equity risk premium relative to bills was 4.1 percent, whereas over the period 1950–2001, it was 7.7 percent. Figure 8 shows that most of the sixteen countries had lower mean premia in the first half-century, with Australia, Italy, Belgium, and South Africa being the exceptions. The sixteen-country (unweighted) mean of the arithmetic risk premia in the first half of the twentieth century was 6.0 percent, versus 8.2 percent in the next fifty-two years. The pattern for the equity premium relative to bonds (not shown in Figure 8) is similar: a pre-1950 mean of 5.5 percent as compared to 7.1 percent over the following fifty-two years.

The large risk premia achieved during the second half of the twentieth century are attributable to three factors. First, there was unprecedented growth in productivity and efficiency, accelerating technological change, and enhancements to the quality of management and corporate governance. As Europe, North America, and the Asia-Pacific region emerged from the turmoil of the Second World War, expectations for improvement were limited to what could be imagined. Reality almost certainly exceeded investors' expectations. Corporate cash flows grew faster than investors anticipated, and this higher growth is now known to the market and built into higher stock prices.

FIGURE 9
INFERRING EXPECTATIONS FROM THE HISTORICAL PREMIUM



Source: Dimson, Marsh and Staunton, Triumph of the Optimists, Princeton University Press, 2002

Second, stock prices have also risen because of a fall in the required rate of return due to diminished business and investment risk. Business risk diminished as the economic and political lessons of the twentieth century were learned, international trade flows increased, and the Cold War ended. Investment risk diminished over time as investors gained the benefits of diversification, both domestically (through a wider range of quoted securities and industries[16], and through intermediaries such as mutual funds) and internationally (with the disappearance of impediments to foreign investment). Diversification allows investors to lower their risk exposure without detriment to expected return. Finally, transaction and monitoring costs are also lower now than a century ago. Factors such as these, which led to a reduction in the required risk premium, have contributed further to the upward re-rating of stock prices.

To convert from a pure historical estimate of the risk premium into a forward-looking projection, we need to reverse-engineer the factors that drove up stock markets over the last 102 years. The simplest idea would be to infer the impact on returns of the historical changes in dividend yield. But we can go beyond this, as shown in Figure 9. The left-hand panel of Figure 9 relates to the US equity market, the centre panel to the UK market, and the right-hand panel to the world market. Within each panel, the first bar portrays the historical annualized risk premium of the equity market. This includes the contribution from unanticipated growth in cash flows and the gain from falls in the required risk premium. We therefore deduct the impact of these two factors. What remains in the right-hand bar of each panel is an estimate of the prospective risk premium demanded by investors as compensation for the risks of equity investment. We explain below how we quantify the deductions in the two centre bars of each panel, but the key qualitative point is that the prospective risk premium is lower than the raw historical risk premium.

---

16. At the start of our research period in 1900, US domestic investors would have found it much harder than today to construct a well-diversified portfolio. At the start of 1900, there were just 123 stocks listed on the New York Stock Exchange, and a single industry, railroads, accounted for 63 percent of their total market value. See Chapter 2, Dimson, E, P R Marsh, and M Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press, 2002.

Unanticipated growth

To apply this framework, we need some notion of when cash flows (proxied here by equity dividends) have exceeded or fallen short of expectations. A simple approach that is commonly used today for forecasting the long-run dividend growth rate is to extrapolate from previous long-term dividend growth. The long-term real dividend growth rate is then used to make a naive projection of future real growth. That is, we estimate the product of *1 + Year 1 annual growth* multiplied by *1 + Year 2 annual growth* and so on to year n. We then compute the n^{th} root of this product, which is equal to *1 + Projected growth*. To summarize, we calculate the annualized real dividend growth rate to each year-end, over periods that start in 1900.

We assume that at every December 31st, investors compare the year's real dividend growth to the real growth rate that would have been projected as at January 1st of that year. The difference is defined as *1 + Annual dividend growth* divided by *1 + Projected growth*, minus *1*. This error in projecting dividend growth may be thought of as the unanticipated growth rate in dividends. The unanticipated changes in dividend growth are compounded together to produce an estimate of their annualized impact over the last century. This is clearly a rather ad hoc measure of unanticipated real dividend growth, but it suffices to illustrate the general idea. Defined this way, Figure 9 shows that the stock price impact of unanticipated dividend growth over the period from 1900 to 2001 is 0.2 percent per year for the United States, 1.6 percent per year in the United Kingdom, and 0.6 percent per year for the world equity market.

Since 1900, there has also been a dramatic change in the valuation basis for equity markets. The price/dividend ratio (the reciprocal of the dividend yield) at the start of 1900 was twenty-three in both the United States and the United Kingdom, but by the start of 2002, the US ratio had risen to eighty-one and the UK ratio to thirty-nine. Undoubtedly, this change is in part a reflection of expected future growth in real dividends, so we could in principle decompose the impact of this valuation change into both an element that reflects changes in required rates of return, and an element that reflects enhanced growth expectations.

To keep things simple, we assume that the increase in the price/dividend ratio is attributable solely to a long-term fall in the required risk premium for equity investment. Given this assumption, Figure 9 shows that the stock price impact of the fall in the required risk premium since 1900 is 1.6 percent per year in the United States and 0.5 percent per year in the United Kingdom. This, together with the impact of unanticipated dividend growth, must be deducted from the historical risk premium.

To estimate the expected future risk premia, we must deduct the impact of both unanticipated cash flows and the fall in the required risk premium from our historical premia. The first of these adjustments can be thought of as the impact of good luck, while the second can be viewed as the effect of re-rating. Figure 9 shows quite large differences in the relative importance of these factors between the United States and the United Kingdom. In particular, for the US market, good luck appears to have had a smaller impact, and re-rating a larger influence. This arises partly from our using dividends as a proxy for unexpected cash flows and changes in the dividend price ratio as a proxy for re-rating. In the United States, the rapid growth of stock repurchases and the trend toward "disappearing dividends"[17] makes it harder

---

17. See Fama, E. F. and K. R. French, "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay", *Journal of Financial Economics*, Vol. 60, 2001, pp.3-43.

to disentangle these effects. The United States is the outlier among our sixteen countries[18], and in judging the relative contribution of unanticipated cash flows versus the impact of the fall in the required risk premium, the UK pattern may be more informative (see Figure 9).

The net effect of deducting the two adjustments from the historical risk premia is shown in the final bar of each of the three panels in Figure 9. These indicate an expected future geometric risk premium of 4.0 percent for the United States, 2.3 percent for the United Kingdom, and 2.9 percent for the world equity market. Our estimates for the United States are similar to those obtained recently by Fama and French using a related approach[19]. Also based on dividend yields and dividend growth estimates, Fama and French use the Gordon model to compute the US equity premium from 1872–1999. They find a premium of 3.8 percent before 1949, and a premium of 3.4 percent for the subsequent period. They argue that the difference between these estimates and the larger ex post risk premium based on historical realized returns is attributable to a reduction since 1949 in investors' required rate of return.

EXPECTED RISK PREMIA

If they are to be used as prospective risk premia, our annualized figures need to be converted into arithmetic means, as explained earlier. Using a projected standard deviation for US and UK equities of 16 percent, the prospective arithmetic risk premia for the United States is 5.3 percent, while the premium for the United Kingdom is 3.6 percent. Using a slightly lower standard deviation for the world index of 14 percent, the prospective arithmetic risk premium for the world index is 3.9 percent. Whichever country one focuses on, our forward-looking predictions for the equity risk premium are lower than the historically based projections reviewed earlier.

A literal interpretation of historical averages might suggest that France has a higher equity risk premium, while Denmark's is lower. While there are obviously differences in risk between markets, this is unlikely to account for cross-sectional differences in historical premia. Indeed, much of the cross-country variation in historical equity premia is attributable to country-specific historical events that will not recur. When making future projections, there is a strong case, particularly given the increasingly international nature of capital markets, for taking a global rather than a country-by-country approach to determining the prospective equity risk premium.

However, just as there must be some true differences across countries in their riskiness, there must also be variation over time in the levels of stock market risk. It is well known that stock market volatility wanders over time, and it is likely that the "price" of risk—namely the risk premium—also fluctuates over time. In the days following September 11, 2001 for example, financial market risk was high, and it is likely that the equity premium demanded by investors was also high. This depressed the market. If the terror had escalated further, the market may have collapsed; but Armageddon did not arrive and the market bounced back.

---

18. Compared with the United States, stock repurchases have been far less prevalent in the other countries. In Europe, the United Kingdom has the highest level of buybacks, but even UK repurchases are small compared with the United States. See section 11.6 of Dimson, E, P R Marsh, and M Staunton, *Triumph of the Optimists: 101 Years of Global Investment Returns*, Princeton University Press, 2002.

19. Fama, E. F. and K. R. French, "The Equity Premium", *Journal of Finance*, Vol. 57, 2002, pp.637-59.

There were similar considerations a generation earlier during the Cuban missile crisis—another Armageddon that was averted. Clearly, at such times risk premia are above average. However, it is difficult to predict premia from the rolling ten-year averages depicted earlier in Figure 4. Indeed, it is difficult to infer expected premia from any analysis of historical excess returns. It may be better to use a "normal" equity premium most of the time, and to deviate from this prediction only when there are compelling economic reasons to suppose expected premia are unusually high or low.

CONCLUSION

The equity premium is the difference between the return on risky stocks and the return on safe bonds. The equity risk premium is central to corporate finance and investment. It is often described as the most important number in finance. Yet it is not clear how big the equity premium has been in the past, or how large it is today.

This paper has presented new evidence on the historical risk premium for sixteen countries over 102 years. Our estimates are lower than frequently quoted historical averages such as the Ibbotson Associates' figures for the United States and the earlier Barclays Capital and CSFB studies for the United Kingdom. The differences arise from previous bias in index construction for the United Kingdom, and, for both countries, from our choice of a longer time frame from 1900–2001, which incorporates the earlier part of the twentieth century, as well as the opening years of the new millennium. In addition, our global focus results in somewhat lower risk premia than hitherto assumed, since prior views have been heavily influenced by the experience of the United States, yet we find that the US risk premium has been somewhat higher than the average for the other fifteen countries.

The historical equity premium is often presented in the form of an annualized rate of return, which summarizes past performance in one number. For the future, what is required is the arithmetic mean of the distribution of equity premia, which is larger than the geometric mean. For markets that have been particularly volatile, the arithmetic mean of past equity premia may exceed the geometric mean premium by several percentage points.

In forecasting the future arithmetic mean premium, investors or companies who believe they can expect the same annualized risk premium as they have received in the past still need to adjust for the differences between historical market volatility and the volatility that we might anticipate today. More fundamentally, however, we have argued that past returns have been flattered by the impact of good luck and re-rating. Since the middle of the last century, equity cash flows almost certainly exceeded expectations, and the required rate of return doubtless fell as investment risk declined and the scope for diversification increased. Stock markets rose for reasons that are unlikely to be repeated. This means that when seeking forecasts for the future, historical risk premia should be adjusted downward for the impact of these factors.

We have illustrated one approach that can be used to make such adjustments. The result is a set of forward-looking, geometric mean risk premia for the United States, United Kingdom and for the world all falling within a range of around 2½ to 4 percent, and a corresponding set of arithmetic mean risk premia falling in a range of around 3½ to 5¼ percent. These estimates are not only far lower than the historical premia quoted in most textbooks, but they are also lower than those cited in surveys of finance academics.

TO BE AUGMENTED WITH DATA FROM FUTURE SURVEYS

Expectations of equity risk premia, volatility and asymmetry from a corporate finance perspective

John R. Graham,
Fuqua School of Business, Duke University, Durham, NC 27708, USA

Campbell R. Harvey*
Fuqua School of Business, Duke University, Durham, NC 27708, USA
National Bureau of Economic Research, Cambridge, MA 02912, USA

ABSTRACT

We present new evidence on the distribution of the ex ante risk premium based on a multi-year survey of Chief Financial Officers (CFOs) of U.S. corporations. Currently, we have responses from surveys conducted from the second quarter of 2000 through the third quarter of 2001. The results in this paper will be augmented as future surveys become available. We find direct evidence that the one-year risk premium is highly variable through time and 10-year expected risk premium is stable. In particular, after periods of negative returns, CFOs significantly reduce their one-year market forecasts, disagreement (volatility) increases and returns distributions are more skewed to the left. We also examine the relation between ex ante returns and ex ante volatility. The relation between the one-year expected risk premium and expected risk is negative. However, our research points to the importance of horizon. We find a significantly positive relation between expected return and expected risk at the 10-year horizon.

---

*Draft: November 30, 2001. Corresponding author, Telephone: +1 919.660.7768, Fax: +1 919.660.8030, E-mail address: cam.harvey@duke.edu. We thank the Financial Executives International (FEI) executives who took the time to fill out the surveys. We thank Alon Brav, Magnus Dahlquist, Ron Gallant, Jim Smith, Paul Soderlind and Bob Winkler for their helpful comments and the participants at the NBER Corporate Finance Summer Workshop. Krishnamoorthy Narasimhan provided research assistance. This research is partially sponsored by FEI but the opinions expressed in the paper are those of the authors and do not necessarily represent the views of FEI. Graham acknowledges financial support from the Alfred P. Sloan Research Foundation.

1. Introduction

The current market capitalization of U.S. equities is approximately $10 trillion. A shift in the equity risk premium by just one percent could add or subtract $1 trillion in market value. In addition, corporate investment decisions hinge on the expectations of the risk premium (via the cost of capital) as do both U.S. and international asset allocation decisions. Therefore, it is important for financial economists to have a thorough understanding of the expected risk premium and the factors that influence it.

The expected market risk premium has traditionally been estimated using long-term historical average equity returns. Using this approach, in December 1999, the arithmetic average return on the S&P 500 over and above the U.S. Treasury bill was reported by Ibbotson Associates (2000) to be 9.32%. This is an extraordinarily high risk premium – though it seems to have influenced the views of a great many academics [Welch (2000)]. Fama and French (2001) conclude that average realized equity returns are in fact higher than ex ante expected returns over the past half century because realized returns included "large unexpected capital gains". If this is true, then using historical averages to estimate the risk premium is misleading.

We use a different approach to estimate the expected risk premium and offer a number of new insights. We base our estimate on a multiyear survey of Chief Financial Officers (CFOs), designed to measure their expectations of risk premia over both short and long horizons. Our survey is unique in that we obtain a measure of the entire risk premium distribution, rather than just the expected value (mean). That is, our survey captures both market volatility and asymmetries implicit in the respondents' probability distributions. In addition, we shed light on how recent stock market performance impacts the ex ante risk premium, volatility and asymmetries. We also study the relation between expected risk and expected return.

There are many methods to estimate the equity risk premium and we can not tell which method is the best – because the variable of interest is fundamentally unobservable. The average of past returns is the method with the longest tradition. However, there are other time-series methods that use measures like dividend yields to forecast and short-horizon premia. These models are difficult to estimate and often structurally unstable [see Garcia and Ghysels (1999)].

There is considerable recent interest in what might be referred to as the implied method. There are two streams of this research. The original is based on the work of Black and Litterman (1990, 1991) and French and Poterba (1991). They argue that one can use investment weights to determine the equilibrium expected returns on equities as well as other assets. Graham and Harvey (1996) use a variant of this method to study the time-series behavior of equity risk premia implicit in the asset allocation recommendations of investment advisors.

A second approach uses fundamental data to deduce risk premia. Gebhardt, Lee and Swaminathan (2000) use firm level cash flow forecasts to derive an internal rate of return, or cost of capital, given the current stock price. Fama and French (2001) study the risk premia on the S&P 500 from 1872-2000 using fundamental data. They argue that the ex ante risk premia is much lower than the historical average, between 2.55% and 4.32% for 1951-2000. Ibbotson and Chen (2001) estimate a long-term risk premium between 4 and 6%.

The final approach to estimate the equity risk premium category directly measures investor's and analyst expectations using survey methods. For example, Welch (2000) analyzes the views of financial economists. Fraser (2001) and Harris and Marston (2001) consider the evidence from financial analysts.

We, instead, survey CFOs. We think that this approach has several advantages. First, one could argue that the financial economists are not directly connected to the allocation decisions in the economy - either capital allocation (financial investment decisions) or real allocation (choosing real investment projects).

CFOs, in contrast, are directly involved their firms' financial and real allocation decisions.

Second, biases in analysts' earnings expectations are well documented. Claus and Thomas (2001) use analysts' earnings expectations to derive an estimated market risk premium of 3.4%. However, to obtain a risk premium this low they dampen the analysts' earnings growth projections for earnings more than five years in the future. When growth is not dampened, Harris and Marston (2001) find an implicit risk premium of 9.2% in 1998. More to the point, Brav and Lehavy (2001) show that analysts' target stock prices are also biased upward. Brav and Lehavy find that analysts' target prices predict a 22% average annual increase in stock prices from 1997-1999, while realized returns average only 15%. In contrast, there is no reason to think that CFOs are biased in their view of the market equity premium.

The CFOs determine the hurdle rate for their firm's investments, and presumably, the equity risk premium plays an important role. Indeed, the evidence in Graham and Harvey (2001) indicates that three-fourths of firms use the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) to establish their cost of capital. The equity risk premium is a critical input into the CAPM.

Our paper offers much more than a survey of CFO's expectations for the market. Our survey is multiyear and rich with additional information. We ask CFOs about their expectations of market performance over both one and 10-year horizons. We ask questions designed to determine their assessment of market volatility. These questions allow us to deduce each CFO's view about the distribution for the market risk premium, and we can observe how the shape and location of these distributions vary with market conditions.

The temporal dimension distinguishes our work from most previous survey work. We are able to address issues such as whether volatility and the risk premium are positively correlated through time. We are able to determine whether recent stock market performance changes expected returns. The interplay of recent equity performance and volatility expectations allows us to say something about asymmetric volatility. Our survey even allows us to deduce a measure of ex ante skewness.

While the surveys are anonymous, we have information on each respondent's industry, size by revenue, number of employees, headquarters location, ownership and percentage of foreign sales. We use this information to see if there are systematic differences in expectations based on firm characteristics.

Importantly, this is on-going research. We have conducted surveys representing over 1,100 total responses, from the second quarter of 2000 through the third quarter of 2001. We plan to update this paper as new surveys are conducted.

The results indicate that the one-year risk premium averages between 0.1 and 2.5 percent depending on the quarter surveyed. The 10-year premium is much less variable and ranges between 3.6 and 4.7 percent. We find that the CFOs' assessment of market volatility is much lower than popular alternative measures, strongly suggesting that CFOs are very confident in their opinions (i.e., their individual distributions for the market risk premium are tight).

We also find that the recent performance of the S&P 500 has a significant effect on the short-term expected risk premium as well as forecasted volatility. Recent stock market performance also has a pronounced effect on CFO's ex ante skewness. In general, when recent stock market returns have been low, the expected risk premium is low, its distribution has a relatively fat left tail, and expected market volatility is high. Finally, we document a negative ex ante relation between expected returns and expected volatility at the one-year horizon and a positive relation at the 10-year horizon. Our results support the notion of a positive tradeoff between risk and expected return – but only at longer horizons.

The paper is organized as follows. The second section details the methodology and the sampling procedure. The results are presented in the third section. An analysis conditional on firm characteristics is outlined in the fourth section. Some concluding remarks are offered in the final section.

2. Methodology

2.1 Design

The quarterly survey project is a joint effort with the Financial Executives International (FEI). FEI has approximately 14,000 members that hold policy-making positions as CFOs, treasurers, and controllers at 8,000 companies throughout the U.S. and Canada. Every quarter, Duke University and the FEI poll these financial officers with a one-page survey on important topical issues (Graham, 1999). The usual response rate for the quarterly survey is 5%-8%.

The history of the survey instrument appears on the Internet at the address http://www.duke.edu/~charvey/Research/indexr.htm. Exhibit 1 details the exact questions that we asked regarding the equity premium.

2.2 Delivery and response

The survey is administered by a third-party data processing firm (Office Remedies Inc.). FEI faxes out approximately 4,000 surveys to a sample of their membership. The executives return their completed surveys by fax to the third-party data vendor. Using a third party ensures that the survey responses are anonymous. Although we do not know the identity of the survey respondents, as mentioned previously, we do know a number of firm-specific characteristics, as discussed below.

The surveys analyzed in this paper were distributed on the following days: June 6, 2000; September 7, 2000; December 4, 2000; March 12, 2001; June 7, 2001 and September 10, 2001. In each case, the survey contained information about the yield on the 10-year Treasury bond at the close of the previous business day, and the respondents were given approximately five business days to return the survey. The date and time the survey is received is recorded on the survey. This allows us to examine if recent equity returns impact the CFOs' responses when they fill out the survey. Two-thirds of the surveys are usually returned within two business days.

We also conducted a survey at the North Carolina CFO Symposium (also sponsored by FEI) on August 22, 2000. In this case, we were able to obtain a response from nearly every executive in the room. By comparing these responses with the faxed quarterly survey responses, we are able to examine whether the response rate on the quarterly survey affects the CFO predictions about the equity market risk premium. (For example, perhaps predominantly "optimists" respond to the quarterly survey.) The North Carolina CFO survey also gathered some additional information about the 10-year risk premium not found on the quarterly surveys. We find that the responses for the North Carolina CFO survey are consistent with those from the quarterly survey. We integrate the responses from this survey into our main results. In our graphical analysis, we highlight this particular survey with a different symbol.[1]

2.3 The survey instrument and summary statistics

The risk premium questions are a subset of a larger set of questions in the Duke-FEI quarterly survey of CFOs. Copies of the surveys can be found on the Internet.

We ask respondents for their one- and 10-year forecasts of the S&P500 given the current 10-year Treasury bond rate (see Exhibit 1). The CFOs also complete the following statement: "During the next year, there is a 1-in-10 chance that the S&P 500 return will be higher than ___%" as well as the analogous question for the "lower" equity return. This allows us to examine each respondent's distribution of expected returns. We can recover a measure of volatility as well as skewness from each individual's responses.

While the survey is anonymous, we ask questions about the firms' characteristics. Fig. 1 presents summary information about the firms in our sample. For this figure, we do not include the characteristics of the firms that participated in the North Carolina CFO Symposium – but concentrate on the quarterly survey participants. We examine three characteristics: industry, revenue, and number of employees.

3. The market risk premium and volatility

3.1 Risk premium

Fig. 2 and 3 present histograms of the ex ante one-year and 10-year risk premia. In Fig. 2, the average one-year risk premium ranges from 0.1% (September 10, 2001 survey) to 3.0% (December 4, 2000). Each of the graphs contains the previous week and previous month's S&P 500 return. Note that the market return was negative preceding the September 10, 2001 survey, and that the average risk premium is the lowest for this survey, 0.1%. Also, for this survey we only include observations that faxed before September 11, 2001.

In Fig. 3, the 10-year risk premium is much more stable ranging from 3.6% (September 10, 2001) to 4.7% (September 7, 2000). Even after the large negative returns in the first quarter of 2001, the survey for the March 12, 2001 shows a 4.5% risk premium.

[1] Later in our analysis, using the non-CFO Symposium data, we test whether headquarters location explains variation in the risk premium across respondents. We find no evidence of a headquarters effect which provides another justification for integrating the CFO Symposium into our results.

Fig. 4 examines whether the past quarter's market performance affects the average one-year and ten-year risk premium.[2] In panel A of Fig. 4, there is a significant relation between the average risk premium and the previous quarter's return. Note that the data for the North Carolina CFO survey is presented with a different symbol, a circle. The results of this survey do not appear unusual. Panel B shows that there is no obvious relation between recent quarterly returns and the 10-year risk premium. While CFOs' assessments of the one-year risk premium appear strongly influenced by recent returns, there is no impact on the 10-year premium.[3]

Table 1 presents regressions that use all of the data (rather than the means of the surveys which are presented in Fig. 4). We estimate weighted least squares regressions where the weights are the inverse of each quarter's variance. Consistent with the graphical analysis, recent realized returns significantly impact the respondents' forecasts of the one-year premium.[4] There is an insignificant relation between the previous return and the 10-year premium. Our one-year results might be capturing an expectational momentum effect. Momentum occurs when future returns are related to past returns. We find that *expected* future returns are related to past returns.

3.2 Volatility and disagreement

We use Davidson and Cooper's (1976) method to recover the probability distribution:

$$\text{Variance} = ([x(0.90)\text{-}x(0.10)]/2.65)^2$$

[2] We also examined the past month. The results are broadly similar and are available on request.

[3] Given that we know the day that the survey was returned, we also investigate whether the past day's return affects the forecasted risk premium. We find evidence that the past day's return has an impact on the one-year forecast and little impact on the 10-year forecast. These results are available on request.

where $x(0.90)$ and $x(0.10)$ represent the 90^{th} and 10^{th} percentile of the respondent's distribution. Keefer and Bodily (1983) show that this simple approximation is the preferred method of estimating the variance of a probability distribution of random variables, given information about the 10^{th} and 90^{th} percentiles. Note that this method allows us to estimate the market variance for each individual survey response.

The distribution of the individual volatilities is presented in Fig. 5. In all cases, the mean volatility is less than seven percent on an annual basis. This is sharply lower than other benchmark measures of volatility, such as the implied volatility on S&P100 index options (VIX). During this time period, the VIX trades between 21 and 35%. However, the VIX roughly measures the standard deviation of daily returns over the next month whereas we are looking for a longer-term volatility. But even if we examine the historical standard deviation of one-year returns (13.0% 1980-2000; 20.1% 1926-2000), the difference between this benchmark and the individual responses suggests that there is a large gap between the individual and market's assessments of volatility. Because the CFO's distributions are very tight, another interpretation is that the CFOs are very confident in their risk premium assessments.

While many studies have econometrically documented a relation between the past returns and volatility, to the best of our knowledge is the first research to examine the relation in the context of survey evidence. Panel A of Fig. 6 shows a somewhat negative relation between the average of the individual ex ante volatilities and the previous quarter's return. However, the regression evidence in

[4] This is also consistent with Welch (2001) who shows in a survey of economists that the mean one-year premium in 1998 was 5.8% (near the peak of the stock market) and only 3.4% in 2001 (after a sizable retreat in the market).

Table 2 that uses all the observations[5] is much weaker. The slope coefficient is not significantly different from zero.

Importantly, market volatility is not the average of individual volatilities. To see this, consider the extreme situation in which everybody has highly confident forecasts (low individual volatility) but considerable disagreement exists across individuals (high cross-sectional dispersion in the risk premium forecasts).[6]

Panel B of Fig. 6 explores this second component of market volatility -- the notion of disagreement. The evidence suggests a sharp negative relation between disagreement and recent returns. That is, large negative returns are associated with a lot of disagreement. The effect is robust to using the previous month instead of the previous quarter's return (unreported).

The final panel in Fig. 6 examines disagreement over the 10-year risk premium and past returns. With this longer horizon forecast, there is not a strong relation between disagreement and past returns.

3.3 Asymmetry in distributions

The survey also captures information on skewness in the individual distributions, which we call asymmetry. We employ a simple metric of asymmetry. We look at the difference between each individual's 90% tail and the mean forecast and the mean minus the 10% tail. Hence, if the respondent's forecast of the risk premium is 6% and the tails are -8% and +11%, then the distribution is negatively skewed with a value of -9%.

Fig. 7 presents histograms of this asymmetry measure for the quarterly surveys. There is substantial asymmetry in the expectations of the risk premium. Indeed,

[5] There are fewer observations in Tables 2 and 3 than Table 1 because a number of respondents did not fill in the range questions.

asymmetric distributions are the rule not the exception. The average asymmetry is generally positive (e.g., panels A, B, C and D). The ex ante asymmetry is quite negative in both the March 12, 2001 and the September 10, 2001 surveys. These are the quarters where the previous three months' stock market returns are very negative.

Fig. 7 suggests a relation between recent return performance and expected asymmetry in the returns distribution. Fig. 8 combines the information from all the surveys and finds a strong positive relation between recent returns and asymmetry. Large negative returns are associated with negative asymmetry in the respondents' distribution of the ex ante risk premium.

Table 3 confirms the highly significant positive relation. Both the lagged one-month and one-quarter returns significantly positively influence the measure of asymmetry. All the coefficients are more than four standard errors from zero.

3.4 The relation between expected returns and volatility

Our results offer some new insights on the modeling of volatility. We have already demonstrated that low or negative realized returns are associated with higher expected volatility and more negative asymmetry in the ex ante returns distributions. This is consistent with the statistical evidence of asymmetry in GARCH modeling (e.g., Nelson (1992) and Glosten, Jagannathan and Runkle (1994)). The statistical evidence usually relies on the leverage hypothesis of Black (1976) and Christie (1982). We refer to this work as statistical evidence because the volatility is measured statistically from past returns data.[7] We offer corroboration by linking past returns to a survey-based ex ante measure of volatility.

---

[6] The variance of returns is the sum of the average of the forecasters' variances and the variance of the forecasters' means. In terms of conditional expectations, Var[r]= E[Var(r|Z)] + Var(E[r|Z)], where r represents returns and Z is the conditioning information that forecasters use.

Given that we have new measures of expected (rather than realized) returns and the ex ante volatility, we can say something about the link between expected returns to expected risk – a fundamental component of asset pricing theory. Indeed, there is a considerable research on this topic which exclusively relies on statistical measures of both the mean and volatility based on historical data. However, the literature is evenly split on whether there is a positive relation or a negative relation between the mean and volatility.

For example, using a GARCH framework, French, Schwert and Stambaugh (1987) and Campbell and Hentshel (1992) estimate a positive relation while Campbell (1987), Breen, Glosten, and Jagannathan (1989), Nelson (1991) and Glosten, Jagannathan and Runkle (1993) find a negative relation between the realized mean and volatility. Harrison and Zhang (1999) use a semi-nonparametric method and find a positive relation. Brandt and Kang (2001) use a latent VAR technique and document a strong negative correlation. Harvey (2001) uses a combination of nonparametric density estimation and GARCH models and finds that the relation depends on the instrumental variables chosen. Both Harvey (2001) and Brandt and Kang (2001) document a distinct counter-cyclical variation in the ratio of mean to volatility.

While our sample is limited in size, we are able to document the relation between a survey-based ex-ante mean and volatility over our surveys. Fig. 9 presents the evidence for three different measures of volatility: the average the respondents' volatilities, disagreement (standard deviation of risk premium forecasts) and a combined measure. The combined measure considers the variation in the location of the individual distributions in addition to considering

---

[7] Figlewski and Wang (2001) re-examine the leverage effect using options implied volatility as an alternative to volatility estimated from past returns.

the volatility of each distribution (aggregate volatility is the mean of the variances plus the variance of the means).[8]

There is a mildly negative relation between the one-year mean and the average volatility in panel A of Fig. 9. In comparison, there is a sharp negative relation between the one-year mean and disagreement in panel B. While R-squares with so few data points can be misleading, the fit here is extraordinary, 93%. The combined measure of volatility also shows a very strong negative relation (panel C in Fig. 9).

Almost all of the past research focuses on short-horizon forecasts of the risk premium and volatility. Our results link well to this past research. However, we also offer some insights on longer-term forecasts. While we only have a measure of disagreement for the one-year forecasts (we do not ask respondents about the 10^{th} and 90^{th} percentiles of the 10-year distribution and, therefore, cannot deduce 10-year volatility), our evidence suggests a strongly significant positive relation between the mean and volatility (panel D). That is, the ex ante relation between mean and volatility appears to be sensitive to the time horizon.

It is possible that the difference between the short-horizon and long-horizon provides some resolution to the conflicting findings in the literature. It seems reasonable that short-horizon expected returns could move around substantially producing either a positive or negative expected returns. Longer horizon returns, on the other hand, are more stable, as we document.

Pástor and Stambaugh have recently presented a Bayesian analysis of long-horizon risk premia. They find that the risk premium in the 1990s is 4.8% which is consistent with our results. However, a critical component of their analysis is the tying of their prior to a positive relation between the premium and volatility. If Pástor and Stambaugh instead chose a diffuse prior relation between volatility and

[8] We appreciate the insights of Bob Winkler on this particular point.

the premium, their estimate of the risk premium in June 1999 rises dramatically to 27.7%. The lower risk premium in the 1990s in the face of high ex post average returns is a result of lower volatility in the market.[9] Our results support the prior they impose.

As a robustness check, we obtain data from the Federal Reserve Board of Philadelphia's Survey of Professional Forecasters. Once a year, the quarterly survey asks a question about the respondent's expected 10-year return on the S&P 500 index. The analysis of this relation is contained in panel E of Fig. 9. We present the risk premium and disagreement for ten surveys beginning in 1992.

Consistent with panel D, there is a positive relation between the expected premium and the expected volatility using these alternative data. There are also differences. There is a much greater variation in disagreement and the risk premium tends to be smaller in the Fed survey. However, these surveys were obtained over a 10-year period where as panel D represents a shorter sample. Nevertheless, the positive relation using long-horizon returns appears to be robust to at least one additional survey.

3.5 Do firm characteristics impact expectations?

Our survey collects information on six firm characteristics: industry, revenue, number of employees, headquarters location, ownership and percentage of sales from foreign sources. It is possible that expectations of market-wide measures like the risk premium might depend on firm characteristics. For example, we have established that the one-year premium depends on past market returns. Is the premium significantly different across the respondents' industries? Given that a

[9] Pastor and Stambaugh show the volatility is 12.8% in the 1990s compared to 17.0% in their full sample.

market-wide measure is being forecasted, our null hypothesis is that there are no significant differences across firm characteristics.

In unreported results, we estimate six regression models (one for each of the characteristics). We regress the risk premium on a series of indicator variables representing fixed effects for each firm characteristic. We also include an indicator variable for each survey date. In all six regressions, the coefficients on the characteristic indicators are not significant at the usual levels of confidence. As a result, we do not reject the null hypothesis that firms' characteristics have no impact on market-wide expectations.

3.6 The September 11, 2001 crisis

Our survey was faxed to CFOs at 8:00am on September 10, 2001. The results in the tables and figures only include data through September 10. However, we have responses that were returned after the crisis. Although the post-crisis sample is small, it is interesting to examine the impact of what we consider a shock to systematic risk because terrorism is undiversifiable in world markets.

Table 4 presents summary statistics for both the September 10 and the post-September 11 sample. We exclude September 11 because some of the surveys we received may have been completed the day before.

The first panel examines the one-year premium which decreases from 0.05% to –0.70% even though both measures of volatility increase substantially. The second panel shows a sharp increase in the 10-year premium from 3.63% to 4.82%. Consistent with the one-year analysis, the volatility increases. While these differences are economically interesting, they are not significantly different because of the small number of observations in the post-September 11 sample.

The differences between the one-year premium and the 10-year premium are consistent with our other analysis. The responses to the one-year premium are likely what the CFOs think will happen near-term in the market – not necessarily what they would require to make a capital investment. However, the 10-year premium more likely represents both expected returns and required returns. In this case, what appears to be a shock to systematic risk, has led to perceptions of higher required returns in equity markets.

6. Conclusions

While surveys of the risk premium are not new, we provide a number of new insights. First, we survey Chief Financial Officers of U.S. corporations and argue that they are uniquely well suited to assess the risk premium given that they routinely use this input in their capital allocation decisions. In addition, we are not particularly concerned that the CFOs are biased in their assessment of the premium – a concern that we have for surveys of financial analysts.

Our survey is designed to look at different horizons (one-year versus 10-year) and, most importantly, to recover the distribution of the risk premium through time. Our survey evidence finds that the one-year premium varies between 0.1 and 2.5% and the 10-year premium falls in the 3.6 to 4.7% range. We find that recent past stock market performance has a large effect on the expected one-year premium and only a small effect on the 10-year premium.

We find that past returns significantly impact volatility as well as the degree of asymmetry in the respondents' distributions. Indeed, we find convincing evidence that recent low returns are associated with higher volatility and more negative asymmetry (i.e., relatively large left tails in the distributions of the expected risk premium). Our evidence supports the statistical evidence that negative return shocks increase volatility.

We have also attempted to shed some light on the relation between the mean and volatility. All previous research has relied on historic data to statistically measure the mean and the variance and this research is split on whether there is a positive relation or negative relation between reward and risk. Our evidence suggests that at the one-year horizon there is a negative relation between the mean and the variance. This poses a challenge to asset pricing theory which implies a positive tradeoff between risk and expected returns. However, at the 10-year horizon, there is evidence of a significantly positive relation. As a robustness experiment, we examine the relation between the ten-year risk premium and dispersion from a Federal Reserve Bank of Philadelphia survey from 1992-2001 and confirm the positive relation between mean and volatility.

Finally, let us emphasize that our work is ongoing. While we have over 1,100 survey responses, much of the analysis presented relies on seven aggregated observations. Indeed, this is the reason that we have mainly presented the data graphically. By viewing these data, each reader can judge the influence of particular observations. Our goal is to continue the survey and dynamically augment this research as new results arrive.

References

Asness, C. S., 2000, Stocks vs. bonds: Explaining the equity risk premium, *Financial Analysts Journa,l* May/June.

Black, F. and R. Litterman, 1990, Asset allocation: Combining investors' views on market equilibrium, Goldman Sachs Fixed Income Research, September.

Black, F. and R. Litterman, 1991, Global asset allocation with equities, bonds and currencies, Goldman Sachs Fixed Income Research, October.

Brandt, M. W. And Q. Kang, 2001, On the relation between the conditional mean and volatility of stock returns: A latent VAR approach, Unpublished working paper, University of Pennsylvania, Philadelphia, PA.

Brav, A., and R. Lehavy, 2001, An empirical analysis of analysts' target prices: Short-term informativeness and long term dynamics, Unpublished working paper, Duke University.

Breen, W., L. R. Glosten and R. Jagannathan, 1989, Economic significance of predictable variation in stock index returns, Journal of Finance 44, 1177-1189.

Campbell, J. Y. and L. Hentschel, 1992, No news is good news: An asymmetric model of changing volatility in stock returns, Journal of Financial Economics 31, 281-318.

Campbell, J. Y., 1987, Stock returns and the term structure, Journal of Financial Economics 18, 373-400.

Chen, N.-f., R. Roll, and S. A. Ross, 1986. Economic forces and the stock market. Journal of Business 59, 383-404.

Claus, J. and J. Thomas, 2001, Equity premia as low as three percent: Evidence from analysts' earnings forecasts for domestic and international stock markets, Journal of Finance 56, 1629-1666.

Christie, A. A., 1982, The stochastic behavior of common stock variances: Value, leverage, and interest rate effects, Journal of Financial Economics 10, 407-432.

Davidson, L. B., and D. O. Cooper, 1976, A simple way of developing a probability distribution of present value, Journal of Petroleum Technology, September, 1069-1078.

Fama, E. F. and French, K. R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427-465.

Fama, E. F. and French, K. R., 2001, The equity premium, Unpublished working paper, University of Chicago, Chicago, IL.

Ferson, W. E. and Harvey, C. R., 1991. The variation of economic risk premiums. Journal of Political Economy 99, 285-315.

Ferson, W. E. and Harvey, C. R., 1993. The risk and predictability of international equity returns. Review of Financial Studies 6, 527-566.

Figlewski, S. and X. Wang, 2001, Is the 'leverage effect' a leverage effect?, Unpublished working paper, New York University, New York, NY.

French, K. R. and J. Poterba, 1991, Investor diversification and international equity markets, American Economic Review 81, 222-226.

French, K. R., G. W. Schwert and R. F. Stambaugh, 1987, Expected stock returns and volatility, Journal of Financial Economics 19, 3-30.

Fraser, P., 2001, How do U.S. and Japanese investors process information and how do they form their expectations of the future? Evidence from quantitative survey based data, Unpublished working paper, University of Aberdeen.

Gebhardt, W. R., C. M. C. Lee, and B. Swaminathan, 2000, Toward an implied cost of capital, Unpublished working paper, Cornell University, Ithaca, NY.

Giordani, P., and P. Soderlind, 2000, Inflation forecast uncertainty, Unpublished working paper, Stockholm School of Economics, Sweden.

Glosten, L. R., R. Jagannathan and D. Runkle, 1993, On the relation between the expected value and the volatility of the nominal excess returns on stocks, Journal of Finance 48, 1779-1802.

Goyal, A. and I. Welch, 1999, Predicting the risk premium, Unpublished working paper, University of California at Los Angeles, Los Angeles, CA.

Graham, J. R., 1999b. Quarter 2, 1999 FEI Survey. http://www.duke.edu/~jgraham.

Graham, J. R. and C. R. Harvey, 1996, Market timing ability and volatility implied in investment newsletters' asset allocation recommendations, Journal of Financial Economics, 397-422.

Graham, J. R. and C. R. Harvey, 2001, Theory and practice of corporate finance: Evidence from the field, *Journal of Financial Economics*, 60, 187-243.

Harris, R. S. and F. C. Marston, The market risk premium: Expectational estimates using analysts' forecasts, Unpublished working paper, University of Virginia, Charlottesville, VA

Harrison, P. and H. Zhang, 1999, An investigation of the risk and return relation at long horizon, Review of Economics and Statistics 81, 399-408.

Harvey, C. R., 2001, The specification of conditional expectations, Journal of Empirical Finance, forthcoming.

Heaton, J. and D. Lucas, 2000, Stock prices and fundamentals, in Ben S. Bernanke and Julio Rotemberg, Eds., NBER Macroeconomics Annual 1999, The MIT Press.

Heaton, J. B., 2000, Managerial optimism and corporate finance. Unpublished working paper, University of Chicago.

Ibbotson, R. G. and P. Chen, 2001, The supply of stock returns, Unpublished working paper, Yale University.

Ibbotson Associates, 1998 Stocks, bonds, bills and inflation, 1998 yearbook, Chicago, IL.

Jagannathan, R. and Meier, I., 2001, Do we need CAPM for capital budgeting? Unpublished working paper, Northwestern University, Evanston, IL.

Jagannathan, R. and Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. Journal of Finance, 51, 3-53.

Jagannathan, R., E. R. McGrattan and A. Scherbina, 2001, The declining U.S. equity premium, Quarterly Review, Federal Reserve Bank of Mineapolis.

Keefer, D. L. and S. E. Bodily, 1983, Three-point approximations for continuous random variables, *Management Science* 29, 5 595-609.

Lamont, O. Earnings and expected returns, Journal of Finance 53, 1563-1587.

Lintner, J., 1965, The valuation of risky assets and the selection of risk investments in stock portfolios and capital budgets, Review of Economics and Statistics 9, 1033-1059.

Nelson, D. B., 1991, Conditional heteroskedasticity in asset returns: A new approach, Econometrica 59, 347-370.

Pástor, L. And R. Stambaugh, 2001, The equity premium and structural breaks, Journal of Finance, 56, 1207-1239.

Poterba, J. M. and Summers, L. H., 1995, A CEO sruvey of U.S. companies' time horizons and hurdle rates, Sloan Management Review, Fall, 43-53.

Sharpe, W., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, Journal of Finance 19, 425-442.

Siegel, J. J., 1999, The shrinking equity premium, Journal of Portfolio Management, 10-17.

Welch, I., 2000, Views of financial economists on the equity premium and other issues, Journal of Business 73 (October): 501-37.

Welch, I., 2001, The equity premium consensus forecast revisited, Unpublished working paper, Cowles Foundation for Research in Economics, Yale University, New Haven, CT.

Exhibit 1

Survey question regarding the risk premium

4. On June 7th, the annual yield on 10-yr treasury bonds was 5.3%. Please complete the following:*
a) Best Guess: Over the next 10 years, I expect the S&P 500 will average a ______% annual return
b) Best Guess: During the next year, I expect the S&P to return _______%
c) High range: During the next year, there is a 1-in-10 chance the S&P 500 return will be higher than _____%
d) Low range: During the next year, there is a 1-in-10 chance the S&P 500 return will be lower than ______%

*Drawn from the survey of June 7, 2001. The rate on the 10-year Treasury bond changes in each survey.

Table 1

The impact of past returns on risk premium forecasts

	A. Including CFO Symposium				B. Excluding CFO Symposium			
	One-year premium		10-year premium		One-year premium		10-year premium	
	Previous quarter's return	Previous month's return	Previous quarter's return	Previous month's return	Previous quarter's return	Previous month's return	Previous quarter's return	Previous month's return
Intercept	2.34	2.21	4.27	4.22	2.23	2.09	4.26	4.20
T ratio	*15.40*	*15.40*	*41.23*	*47.98*	*12.94*	*13.66*	*39.18*	*46.94*
Previous return	0.089	0.144	0.016	0.013	0.074	0.013	0.014	0.012
T ratio	*3.87*	*3.59*	*1.31*	*0.77*	*2.88*	*3.23*	*1.13*	*0.69*
Adj. R^2	0.012	0.010	0.001	0.000	0.007	0.008	0.000	-0.001
Observations	1147	1147	1154	1154	1109	1109	1116	1116

Table 2

The impact of past returns on forecast volatility

	A. Including CFO Symposium One-year forecast volatility		B. Excluding CFO Symposium One-year forecast volatility	
	Previous quarter's return	Previous month's return	Previous quarter's return	Previous month's return
Intercept	6.71	6.78	7.01	6.98
T ratio	*41.54*	*45.18*	*39.90*	*44.86*
Previous return	-0.029	-0.016	0.008	0.003
T ratio	*-1.33*	*-0.47*	*0.33*	*0.07*
Adj. R^2	0.001	-0.001	-0.001	-0.001
Observations	911	911	877	877

Table 3

The impact of past returns on forecast asymmetry

	A. Including CFO Symposium: One-year forecast asymmetry		B. Excluding CFO Symposium: One-year forecast asymmetry	
	Previous quarter's return	Previous month's return	Previous quarter's return	Previous month's return
Intercept	-0.86	-0.94	-0.88	-1.11
T ratio	*-3.27*	*-3.69*	*-2.80*	*-3.93*
Previous return	0.154	0.234	0.153	0.219
T ratio	*4.88*	*4.90*	*4.22*	*4.46*
Adj. R^2	0.025	0.025	0.019	0.021
Observations	911	911	877	877

Table 4

The impact of the September 11, 2001 crisis on expectations

	Pre-September 11	Post-September 11*
One-year risk premium		
Mean premium	0.05	-0.70
Std. dev. (disagreement)	6.61	7.86
Std. dev. (average of individual volatilities)	6.79	9.76
Asymmetry (disagreement)	-2.24	1.96
Asymmetry (average of individual asymmetries)	-0.82	-0.57
Observations	125	33
10-year risk premium		
Mean premium	3.63	4.82
Std. dev. (disagreement)	2.36	3.03
Asymmetry (disagreement)	-0.36	0.14
Observations	127	33

*Surveys faxed on September 11 were excluded from both samples.

The characteristics of the survey respondents' firms

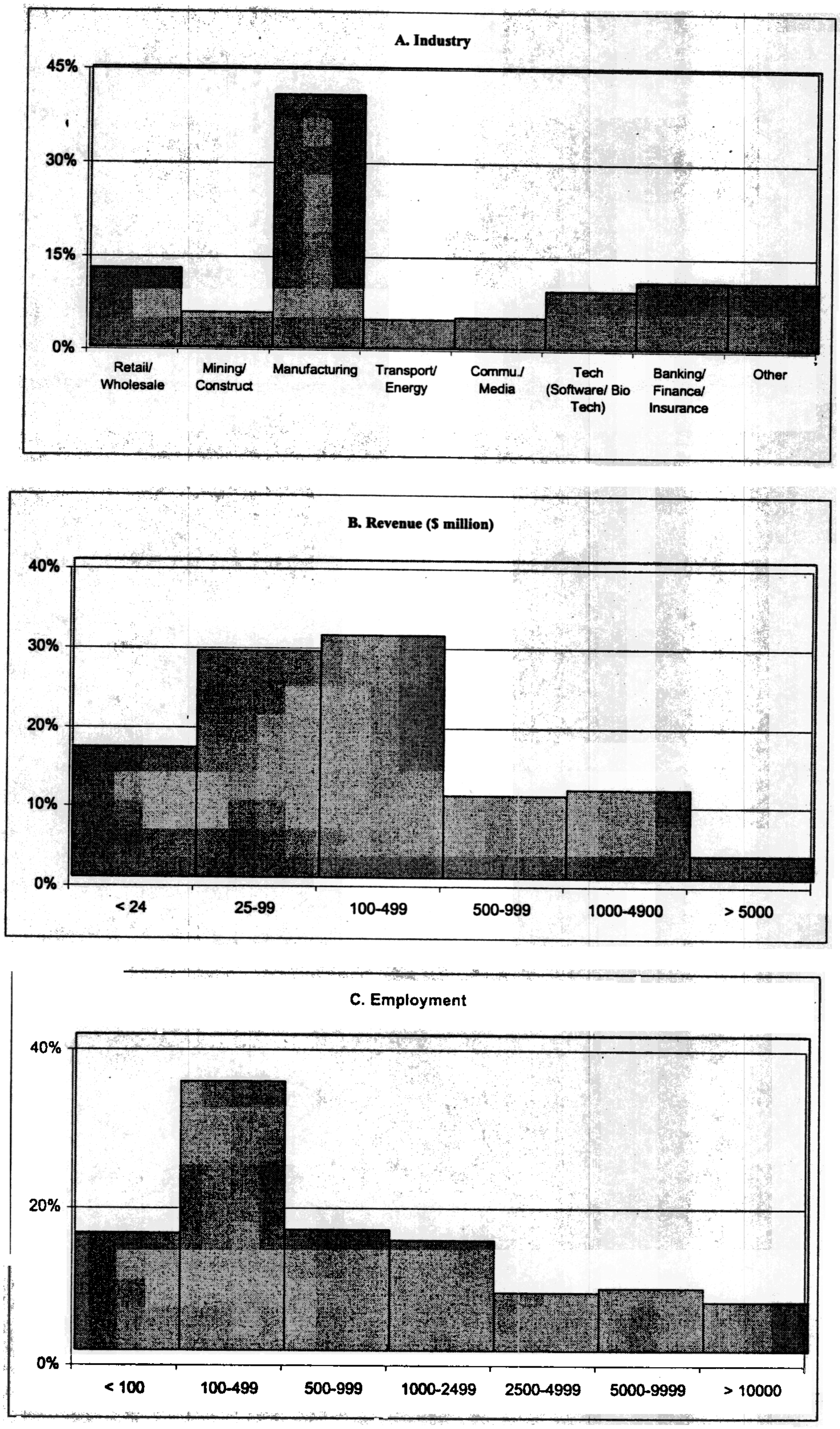


Fig. 1

The distribution of the expected one-year risk premium

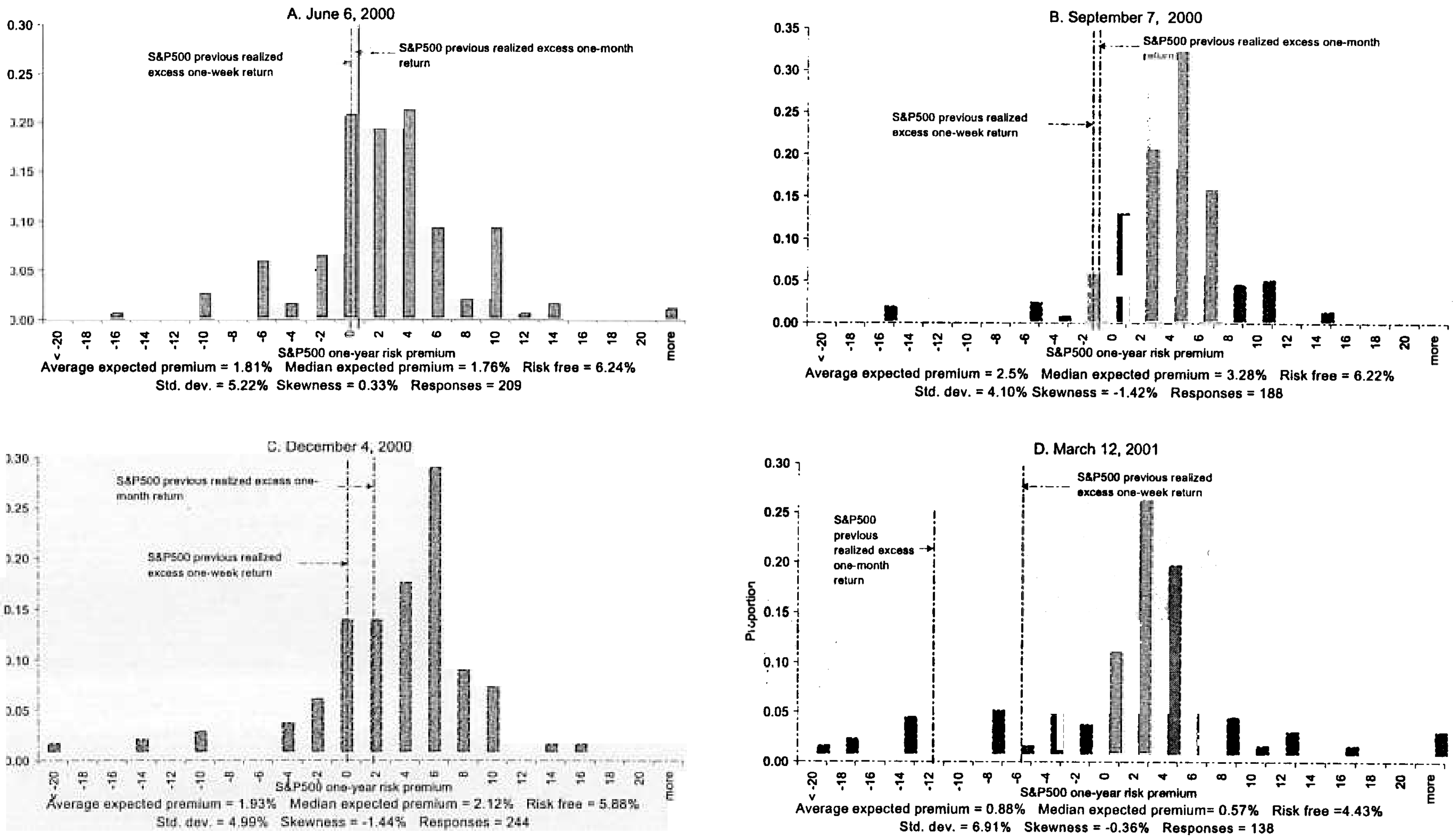


Fig. 2

The distribution of the expected one-year risk premium

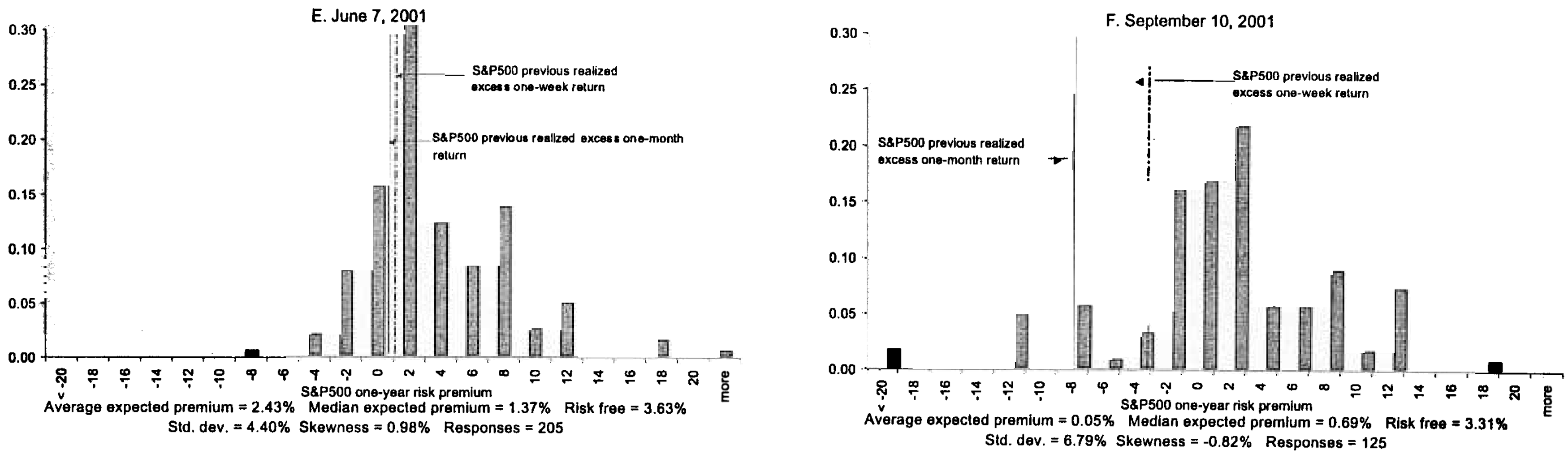


Fig. 2 (continued)

The distribution of the expected 10-year risk premium

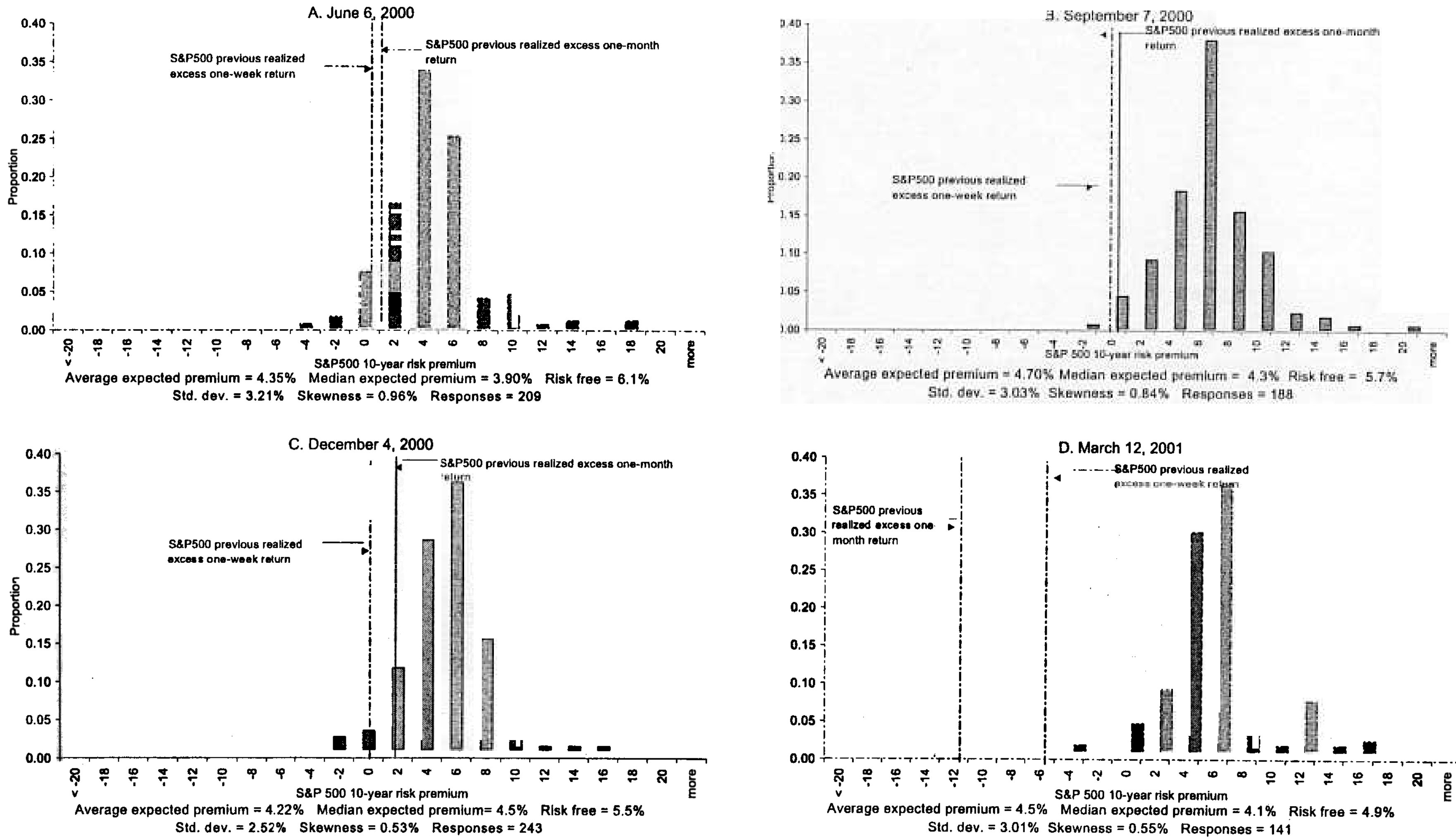
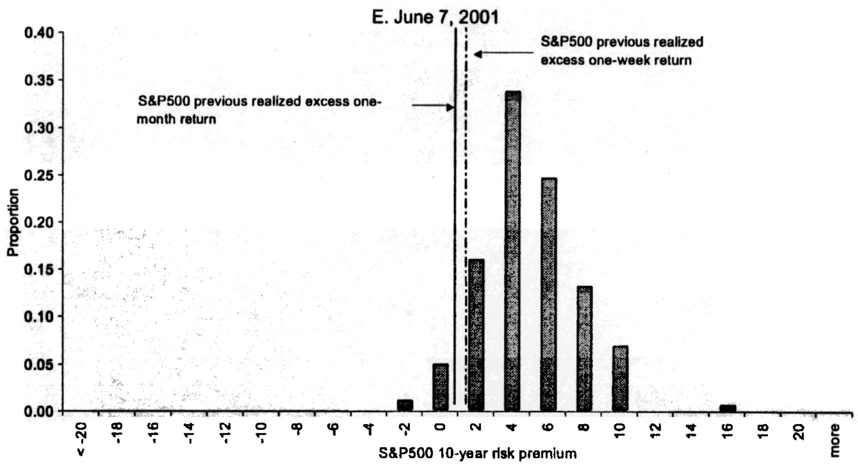
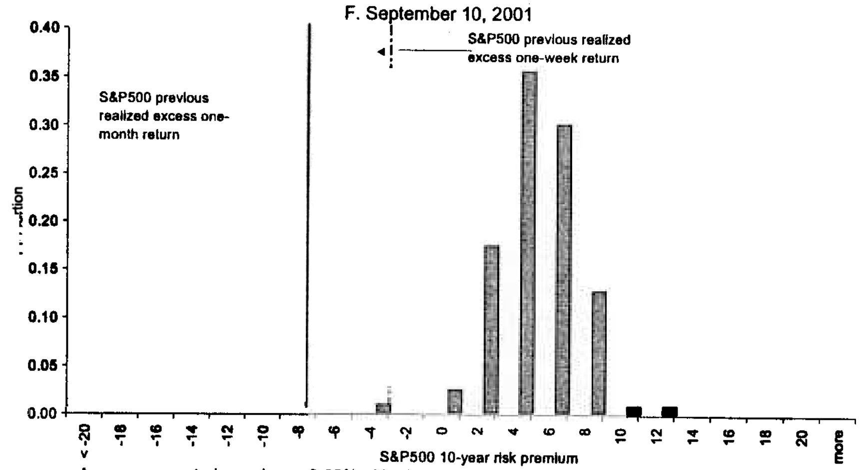


Fig. 3

The distribution of the expected 10-year risk premium



Average expected premium = 3.91% Median expected premium = 3.70% Risk free = 5.3%
Std. dev. = 2.64% Skewness = 0.59% Responses = 208



Average expected premium = 3.63% Median expected premium = 3.20% Risk free = 4.8%
Std. dev. = 2.36% Skewness = -0.36% Responses = 127

Fig. 3 (continued)

Past returns and the one year ex-ante risk premium

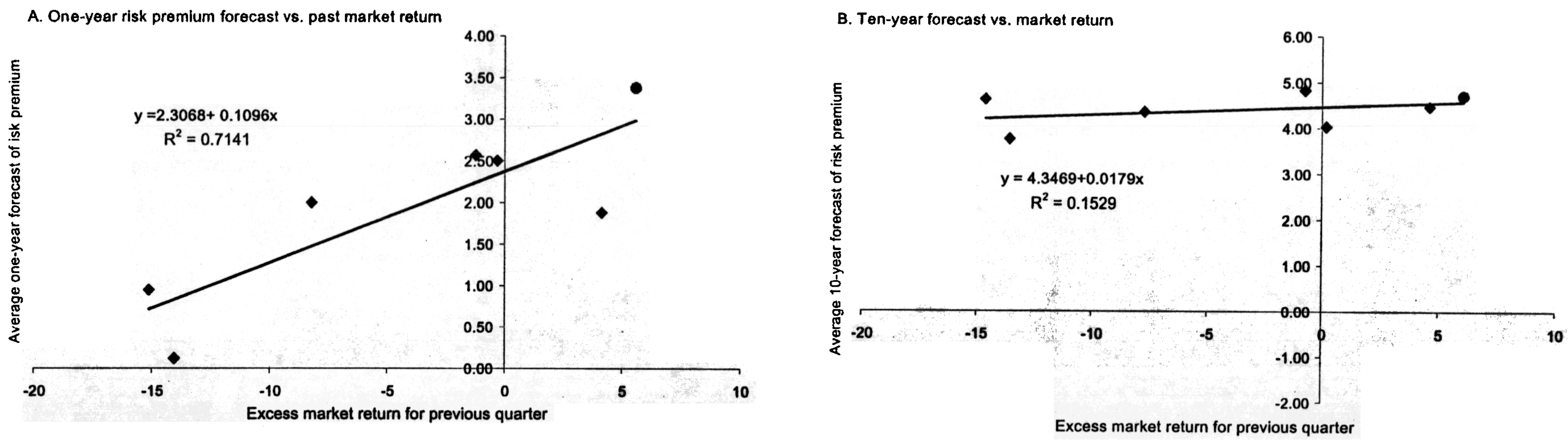


Fig. 4

The distribution of ex ante volatility for one-year return forecasts

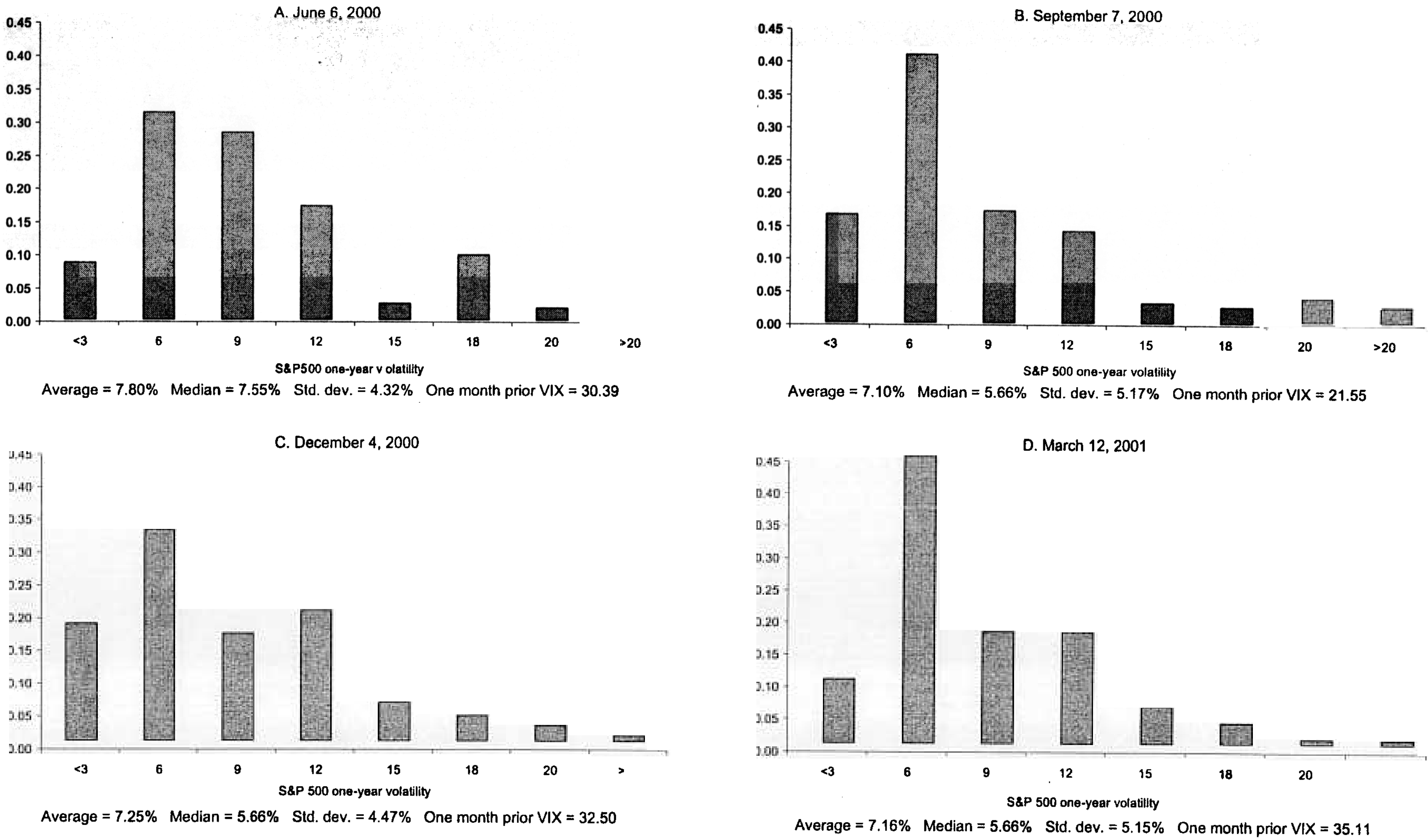
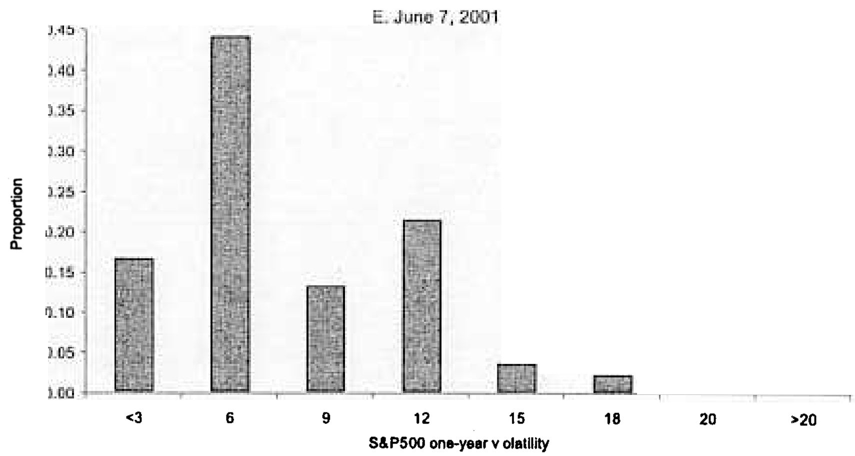
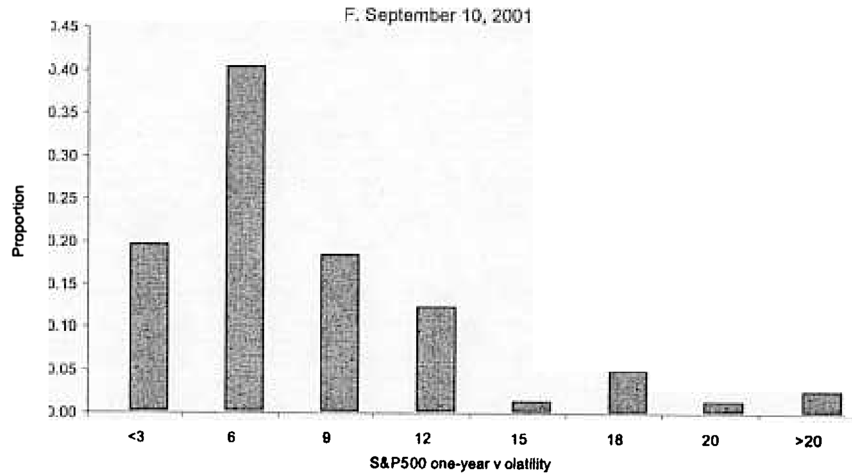


Fig. 5

The distribution of ex ante volatility for one-year return forecasts



Average = 6.14% Median = 5.66% Std. dev. = 3.48% One month prior VIX = 27.54



Average = 6.62% Median = 5.66% Std. dev. = 4.91% One month prior VIX = 23.24

Fig. 5 (continued)

Past returns and volatility

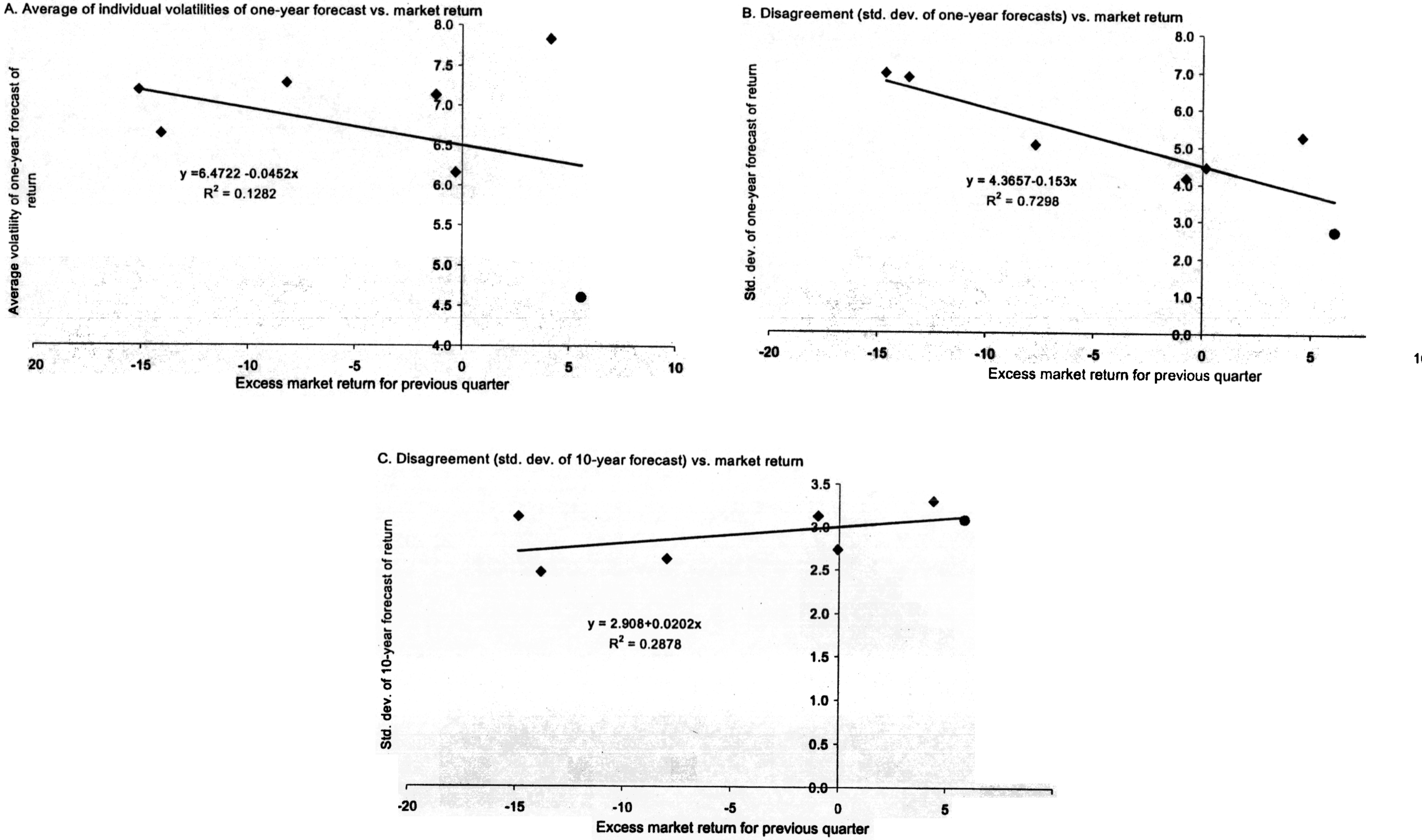


Fig. 6

The distribution of ex-ante asymmetry for one year risk premium forecasts

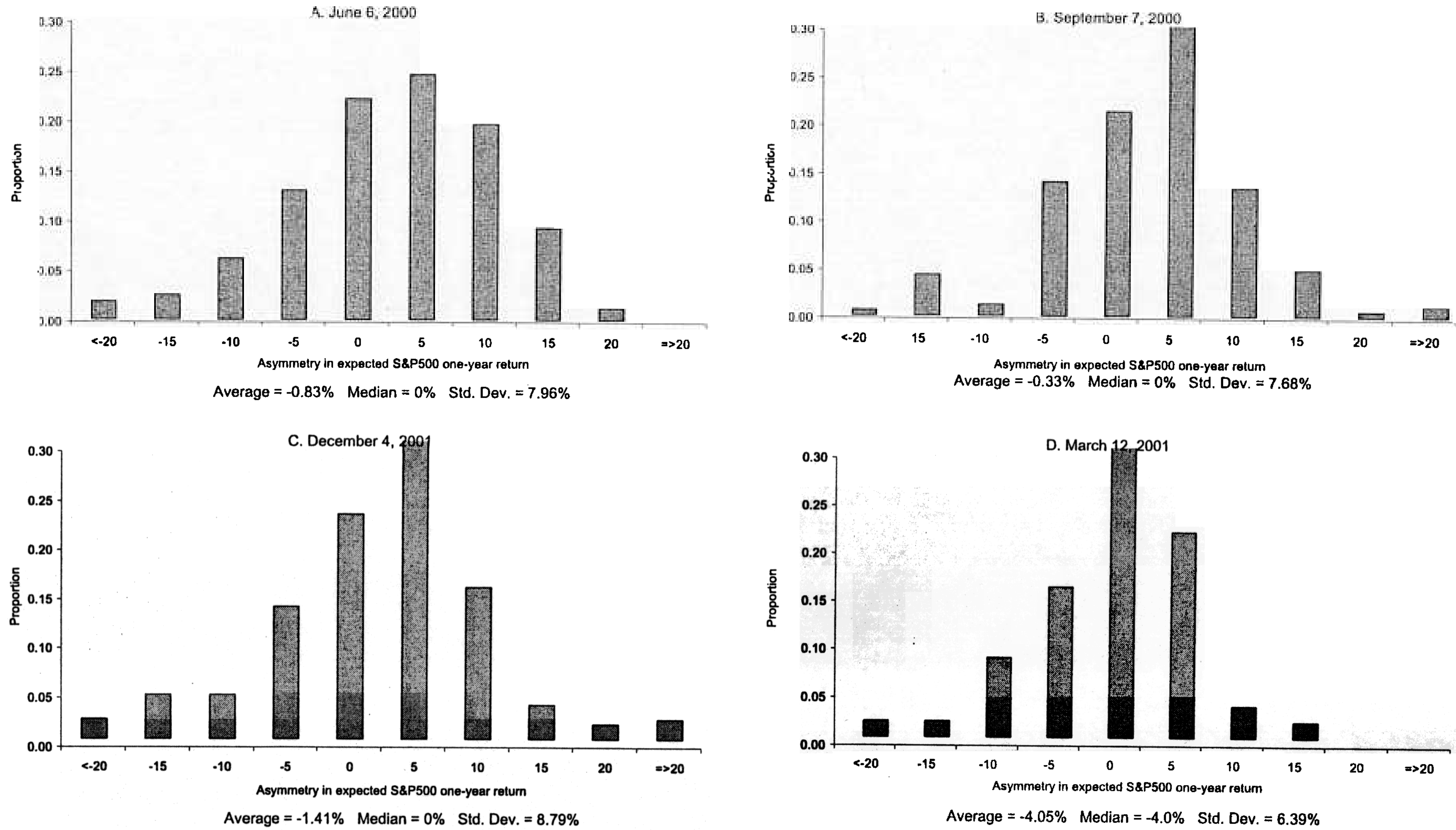
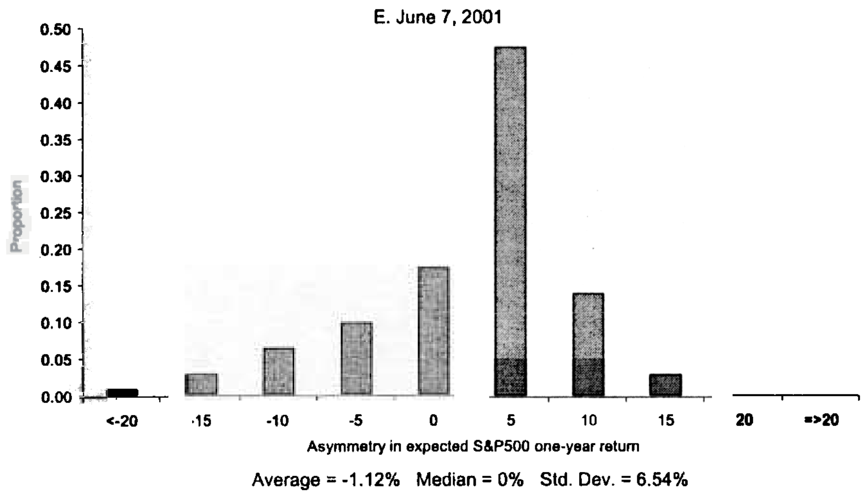
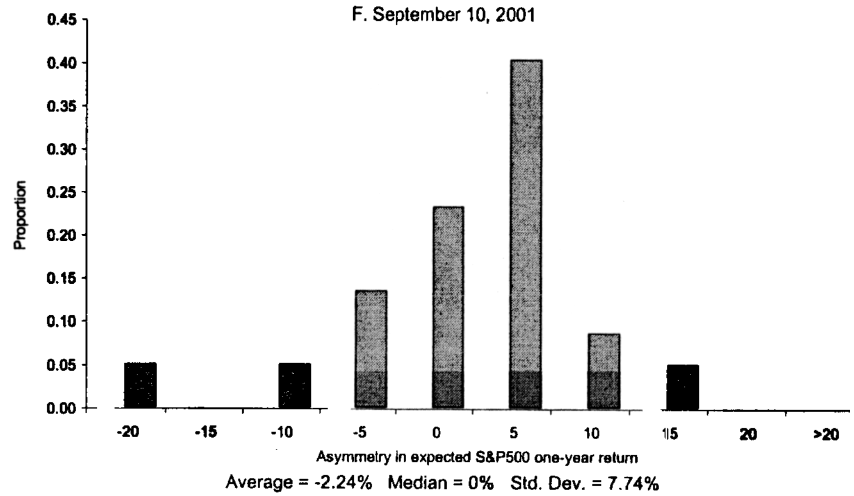


Fig. 7

The distribution of ex-ante asymmetry for one year risk premium forecasts





Recent returns and asymmetry

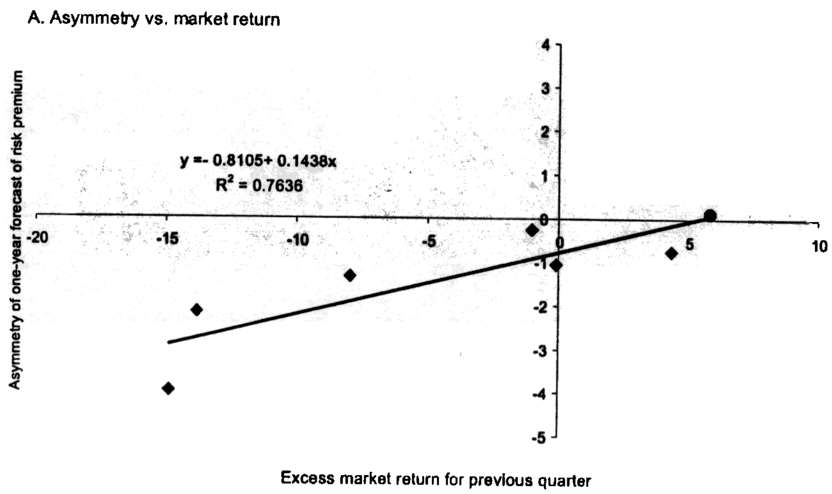


Fig. 8

The relation between the risk premium and ex ante volatility

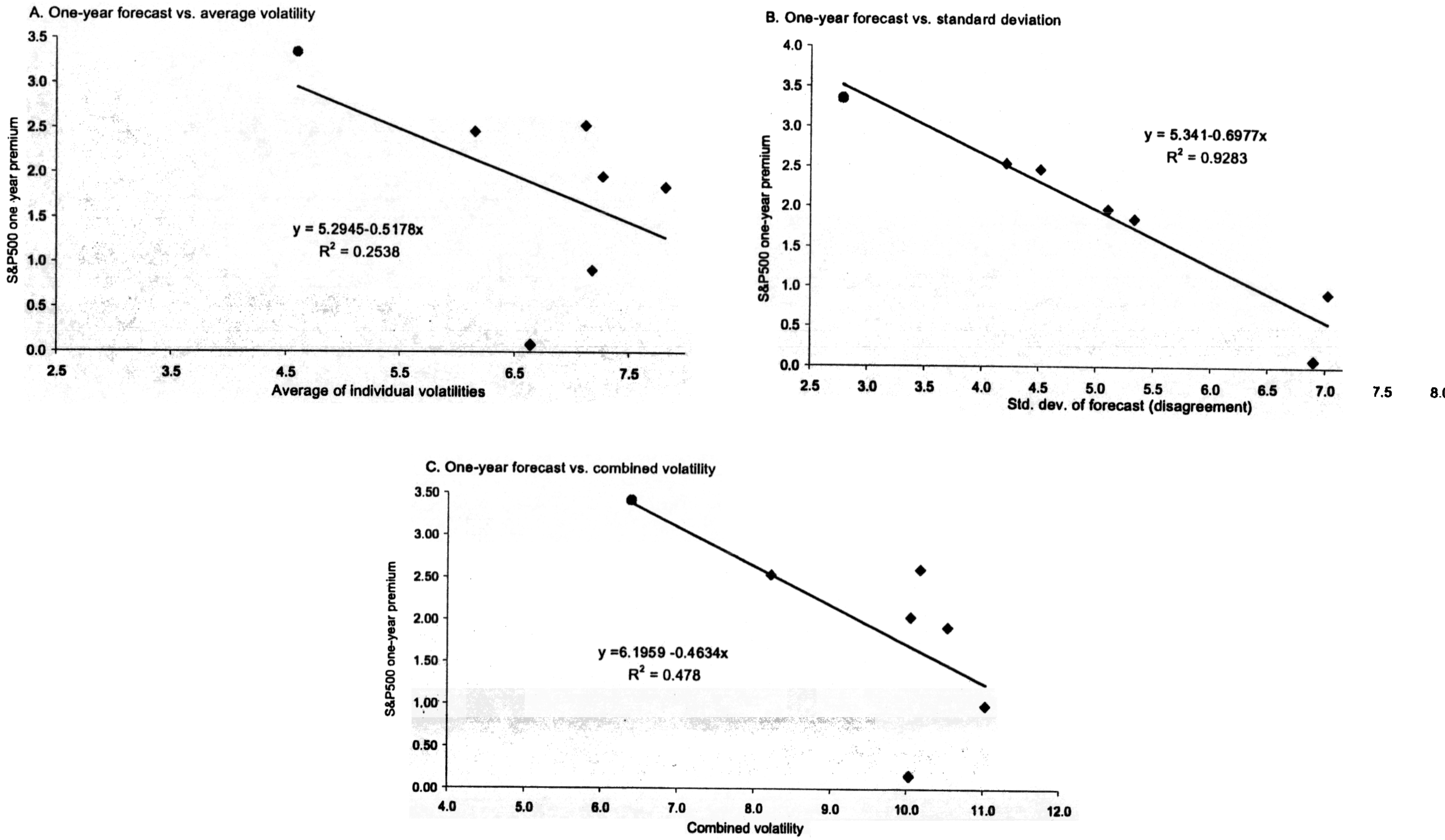
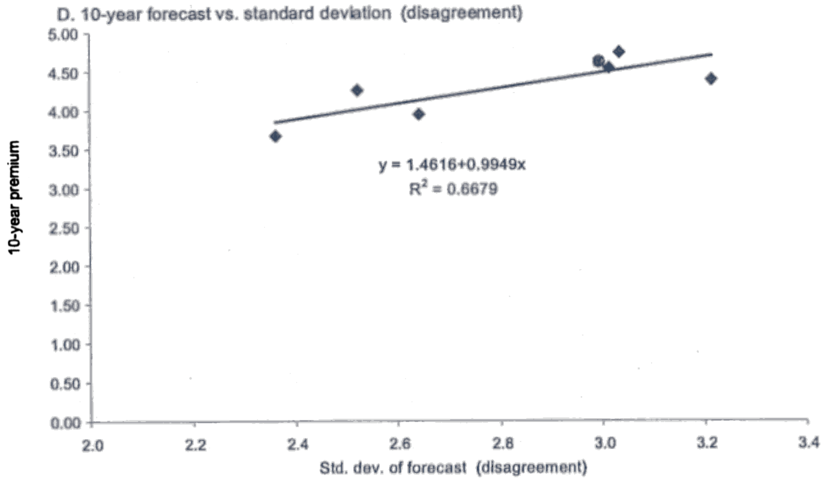
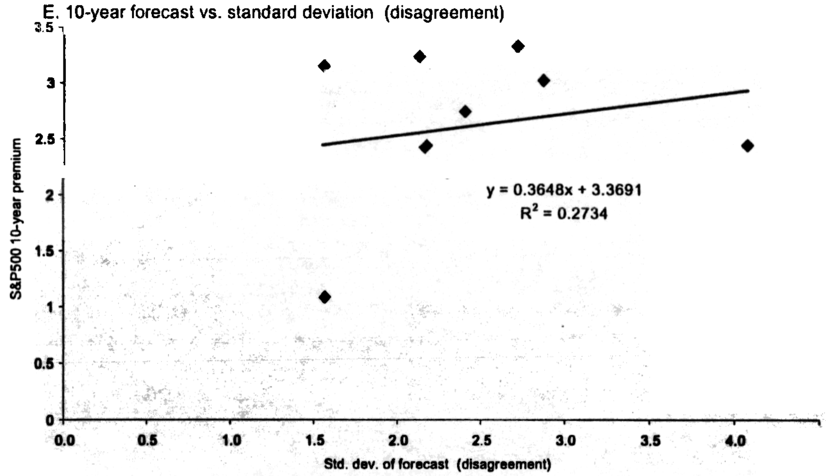


Fig. 9

The relation between the risk premium and ex ante volatility





Data from 1992-2001 from the Federal Reserve of Philadelphia's Survey of Professional Forecasters

Fig. 9 (continued)

PUB-61 NLH
Attachment 4

What Risk Premium Is "Normal"?

Robert D. Arnott and Peter L. Bernstein

The goal of this article is an estimate of the objective forward-looking U.S. equity risk premium relative to bonds through history—specifically, since 1802. For correct evaluation, such a complex topic requires several careful steps: To gauge the risk premium for stocks relative to bonds, we need an expected real stock return and an expected real bond return. To gauge the expected real bond return, we need both bond yields and an estimate of expected inflation through history. To gauge the expected real stock return, we need both stock dividend yields and an estimate of expected real dividend growth. Accordingly, we go through each of these steps. We demonstrate that the long-term forward-looking risk premium is nowhere near the level of the past; today, it may well be near zero, perhaps even negative.

The investment management industry thrives on the expedient of forecasting the future by extrapolating the past. As a consequence, U.S. investors have grown accustomed to the idea that stocks "normally" produce an 8 percent real return and a 5 percent (that is, 500 basis point) risk premium over bonds, compounded annually over many decades.[1] Why? Because long-term historical returns have been in this range with impressive consistency. And because investors see these same long-term historical numbers year after year, these expectations are now embedded in the collective psyche of the investment community.[2]

Both the return and the risk premium assumptions are unrealistic when viewed from current market levels. Few have acknowledged that an important part of the lofty real returns of the past stemmed from rising valuation levels and from high dividend yields, which have since diminished. As we will demonstrate, the long-term forward-looking risk premium is nowhere near the 5 percent level of the past; indeed, today, it may well be near zero, perhaps even negative. Credible studies in and outside the United States are challenging the flawed conventional view. Well-researched studies by Claus and Thomas (2001) and Fama and French (2000) are just two (see also Arnott and Ryan 2001). Similarly, the long-term forward-looking real return from stocks is nowhere near history's 8 percent. We argue that, barring unprecedented economic growth or unprecedented growth in earnings as a percentage of the economy, real stock returns will probably be roughly 2–4 percent, similar to bond returns. In fact, even this low real return figure assumes that current near-record valuation levels are "fair" and likely to remain this high in the years ahead. "Reversion to the mean" would push future real returns lower still.

Furthermore, if we examine the historical record, neither the 8 percent real return nor the 5 percent risk premium for stocks relative to government bonds has ever been a realistic *expectation*, except from major market bottoms or at times of crisis, such as wartime. But this topic merits careful exploration. After all, according to the Ibbotson Associates data, equity investors earned 8 percent real returns and stocks have outpaced bonds by more than 5 percent over the past 75 years. Intuition suggests that investors should not require such outsized returns in order to bear equity market risk. Should investors have expected these returns in the past, and why shouldn't they continue to do so? We examine these questions expressed in a slightly different way. First, can we derive an objective estimate of what investors had good reasons to expect in the past? Second, why should we expect less in the future than we have earned in the past?

The answers to both questions lie in the difference between the *observed* excess return and the *prospective* risk premium, two fundamentally different concepts that, unfortunately, carry the same label—risk premium. If we distinguish between past excess returns and future expected risk premiums, the idea that future risk premiums should be different from past excess returns is not at all unreasonable.[3]

Robert D. Arnott is managing partner at First Quadrant, L.P., Pasadena, California. Peter L. Bernstein is president of Peter L. Bernstein, Inc., New York.

This complex topic requires several careful steps if it is to be evaluated correctly. To gauge the risk premium for stocks relative to bonds, we need an expected real bond return and an expected real stock return. To gauge the expected real bond return, we need both bond yields and an estimate of expected inflation through history. To gauge the expected real stock return, we need both stock dividend yields and an estimate of expected real dividend growth. Accordingly, we go through each of these steps, in reverse order, to form the building blocks for the final goal—an estimate of the objective forward-looking equity risk premium relative to bonds through history.

Has the Risk Premium Natural Limits?

For equities to have a zero or negative risk premium relative to bonds would be unnatural because stocks are, on average over time, more volatile than bonds. Even if volatility were not an issue, stocks are a secondary call on the resources of a company; bondholders have the first call. Because the risk premium is usually measured for corporate stocks as compared with government debt obligations (U.S. T-bonds or T-bills), the comparison is even more stark. Stocks should be priced to offer a superior return relative to *corporate* bonds, which should offer a premium yield (because of default risk and tax differences) relative to T-bonds, which should typically offer a premium yield (because of yield-curve risk) relative to T-bills. After all, long bonds have greater duration—hence, greater volatility of price in response to yield changes—so a capital loss is easier on a T-bond than on a T-bill.

In other words, the current circumstance, in which stocks appear to have a near-zero (or negative) risk premium relative to government bonds, is abnormal in the extreme. Even if we add 100 bps to the risk premium to allow for the impact of stock buybacks, today's risk premium relative to the more relevant corporate bond alternatives is still negligible or negative. This facet was demonstrated in Arnott and Ryan and is explored further in this article.

If zero is the natural minimum risk premium, is there a natural maximum? Not really. In times of financial distress, in which the collapse of a nation's economy, hyperinflation, war, or revolution threatens the capital base, expecting a large reward for exposing capital to risk is not unreasonable. Our analysis suggests that the U.S. equity risk premium approached or exceeded 10 percent during the Civil War, during the Great Depression, and in the wake of World Wars I and II. That said, however, it is difficult to see how one might objectively measure the forward-looking risk premium in such conditions.

A 5 percent excess return on stocks over bonds compounds so mightily over long spans that most serious fiduciaries, if they believed stocks were going to earn a 5 percent risk premium, would not even consider including bonds in a portfolio with a horizon of more than a few years: The probabilities of stocks outperforming bonds would be too high to resist.[4] Hence, under so-called normal conditions—encompassing booms and recessions, bull and bear markets, and "ordinary" economic stresses—a good explanation is hard to find for why expected long-term real returns should ever reach double digits or why the expected long-term risk premium of stocks over bonds should ever exceed about 5 percent. These upper bounds for expected real returns or for the risk premium, unlike the lower bound of zero, are "soft" limits; in times of real crisis or distress, the sky's the limit.

Expected versus "Hoped-For" Returns

Throughout this article, we deal with *expected* returns and *expected* risk premiums. This concept is rooted in objective data and defensible expectations for portfolio returns, rather than in the returns that an investor might *hope* to earn. The distinction is subtle; both represent expectations, but one is objective and the other subjective. Even at times in the past when valuation levels were high and when stockholders would have had no objective reason to expect any growth in real dividends over the long run, hopes of better-than-market short-term profits have always been the primary lure into the game.[5]

When we refer to expected returns or expected risk premiums, we are referring to the estimated future returns and risk premiums that an objective evaluation—based on past rates of growth of the economy, past and prospective rates of inflation, current stock and bond yields, and so forth—might have supported at the time. We explicitly do not include any extrapolation of past returns per se, because past returns are driven largely by changes in valuation levels (e.g., changes in yields), which in an efficient market, investors should not expect to continue into the indefinite future. By the same token, we explicitly do not presume any reversion to the mean, in which high yields or low yields are presumed to revert toward historical norms. We presume that the current yield is "fair" and is an unbiased estimator of future yields, both for stocks and bonds.

Few investors subjectively expect returns as low as the objective returns produced by this sort of analysis. In a recent study by Welch (2000), 236 financial economists projected, on average, a 7.2 percent risk premium for stocks relative to T-bills over the next 30 years. If we assume that T-bills offer the same 0.7 percent real return in the future that they have offered over the past 75 years, then stocks must be expected to offer a compounded geometric average real return of about 6.6 percent.[6] Given a dividend yield of roughly 1.5 percent in 1998–1999, when the survey was being carried out, the 236 economists in the survey were clearly presuming that dividend and earnings growth will be at least 5 percent a year above inflation, a rate of real growth three to five times the long-term historical norm and substantially faster than plausible long-term economic growth.

Indeed, even if ir vestors take seriously the real return estimates and risk premiums produced by the sort of objective analysis we propose, many of them will continue to believe that their own investments cannot fail to do better. Suppose they agree with us that stocks and bonds are priced to deliver 2–4 percent real returns before taxes.[7] Do they believe that *their* investments will produce such uninspired pretax real returns? Doubtful. If these kinds of projections were taken seriously, markets would be at far different levels from where they are. Consequently, if these objective expectations are correct, most investors will be wrong in their (our?) subjective expectations.

What Were Investors Expecting in 1926

Are we being reasonable to suggest that, after a 75-year span with 8 percent real stock returns and a 5 percent excess return over bonds (the Ibbotson findings), an 8 percent real return or a 5 percent risk premium is abnormal? Absolutely. The relevant question is whether the investors of 1926 would have had reason to *expect* these extraordinary returns. In fact, they would not. What they got was different from what they should have expected, which is a normal result in a world of uncertainty.

At the start of 1926, the beginning of the returns covered in the Ibbotson data, investors had no reason to expect the 8 percent real returns that have been earned over the past 75 years nor that these returns would provide a 5 percent excess return over bonds. As we will describe, these outcomes were the consequence of a series of historical accidents that uniformly helped stocks and/or helped the risk premium.

Consider what investors might objectively have expected at the start of 1926 from their long-term investments in stocks and bonds. In January that year, government bonds were yielding 3.7 percent. The United States was on a gold standard, government was small relative to the economy as a whole, and the price level of consumer goods, although volatile, had been trendless throughout most of U.S. history up to that moment; thus, inflation expectations were nil. It was a time of relative stability and prosperity, so investors would have had no reason to expect to receive less than this 3.7 percent government bond yield. Accordingly, the *real* return that investors would have expected on their government bonds was 3.7 percent, plain and simple.

Meanwhile, the dividend yield on stocks was 5.1 percent. We can take that number as the starting point to apply the sound theoretical notion that the real return on stocks is equal to

- the dividend yield
- plus (or minus) any change in the real dividend (now viewed as participation in economic growth)
- plus (or minus) any change in valuation levels, as measured by P/E multiples or dividend yields.

What did the investors expect of stocks in early 1926? The time was the tail end of the era of "robber baron" capitalism. As Chancellor (1999) observed, investors were accustomed to the fact that company managers would often dilute shareholders' returns if an enterprise was successful but that the shareholder was a full partner in any business decline. More important was the fact that the long-run history of the market was trendless. Thoughts of long-term economic growth, or long-run capital appreciation in equity holdings, were simply not part of the tool kit for return calculations in those days.

Investors generally did not yet consider stocks to be "growth" investments, although a few people were beginning to acknowledge the full import of Smith's extraordinary study *Common Stocks as Long-Term Investments*, which had appeared in 1924. Smith demonstrated how stocks had outperformed bonds over the 1901–22 period.[8] His work became the bible of the bulls as the bubble of the late 1920s progressed. Prior to 1926, however, investors continued to follow J.P. Morgan's dictum that the market would fluctuate, a traditional view hallowed by more than 100 years of stock market history. In other words, investors had no *trend* in mind. The effort was to buy low and to sell high, period.

Assuming that markets were fairly priced in early 1926, investors should have expected little or no benefit from rising valuation levels. Accordingly, the real long-term return that stock investors could reasonably have expected on average, or from the market as a whole, was the 5.1 percent dividend yield, give or take a little. Thus, stock investors would have expected roughly a 1.4 percent risk premium over bonds, not the 5 percent they actually earned in the next 75 years. The market exceeded objective expectations as a consequence of a series of historical accidents:

- *Historical accident #1: Decoupling yields from real yields.* The Great Depression (roughly 1929–1939) introduced a revolutionary increase in the role of government in peacetime economic policy and, simultaneously, drove the United States (and just about the rest of the world) off the gold standard. As prosperity came back in a big way after World War II, expected inflation became a normal part of bond valuation. This change created a one-time shock to bonds that decoupled nominal yields from real yields and drove nominal yields higher even as real yields fell. Real yields at year-end 2001 were 3.4 percent (the Treasury Inflation-Indexed Securities, commonly called TIPS, yield[9]), but nominal yields were 5.8 percent. This rise in nominal yields (with real yields holding steady) has cost bondholders 0.4 percent a year over 75 years. That accident alone accounts for nearly one-tenth of the 75-year excess return for stocks relative to bonds.
- *Historical accident #2: Rising valuation multiples.* Between 1926 and 2001, stocks rose from a valuation level of 18 times dividends to nearly 70 times dividends. This fourfold increase in the value assigned to each dollar of dividends contributed 180 bps to annual stock returns over the past 75 years, even though the entire increase occurred in the last 17 years of the period (we last saw 5.1 percent yields in 1984). This accident explains fully one-third of the 75-year excess return.
- *Historical accident #3: Survivor bias.* Since 1926, the United States has fought no wars on its own soil, nor has it experienced revolution. Four of the fifteen largest stock markets in the world in 1900 suffered a total loss of capital, a –100 percent return, at some point in the past century. The markets are China, Russia, Argentina, and Egypt. Two others came close—Germany (twice) and Japan. Note that war or revolution can wipe out bonds as easily as stocks (which makes the concept of "risk premium" less than relevant). U.S. investors in early 1926 would *not* have considered this likelihood to be zero, nor should today's true long-term investor.
- *Historical accident #4: Regulatory reform.* Stocks have gone from passing relatively little economic growth through to shareholders to passing much of the economic growth through to shareholders. This shift has led to 1.4 percent a year growth in real dividend payments and in real earnings since 1926. This accelerated growth in real dividends and earnings, which no one in 1926 could have anticipated, explains roughly one-fourth of the 75-year excess return.[10]

In short, the equity investors of 1926 probably expected to earn a real return little different from their 5.1 percent yield and expected to earn little more than the 140 bp yield differential over bonds. Indeed, an objective investor might have expected a notch less because of the greater frequency with which investors encountered dividend cuts in those days.

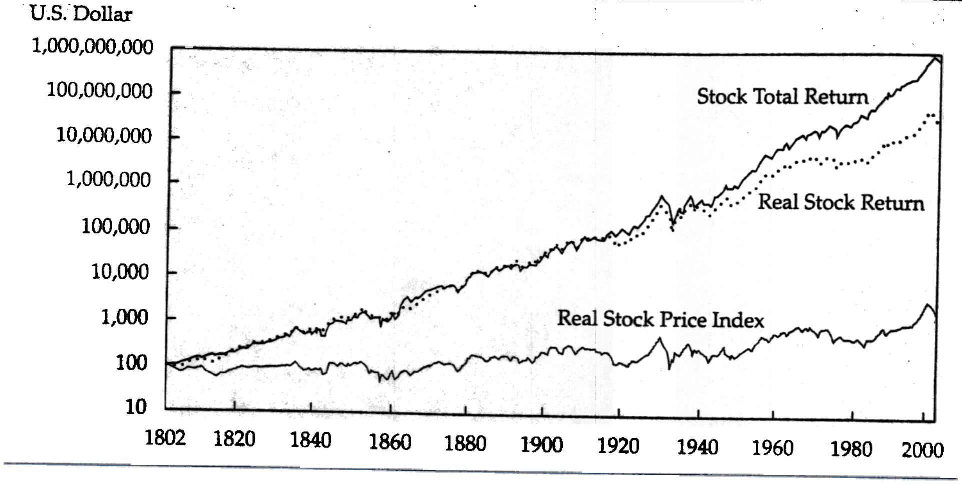
What Expectations Were Realistic in the Past?

To gauge what risk premium an investor might have objectively expected in the longer run past, we need to (1) estimate the real return that investors might reasonably have expected from stocks, (2) estimate the real return that investors might reasonably have expected from bonds, and (3) take the difference. From this exercise, we can gauge what risk premium an investor might reasonably have expected at any point in history, not simply an isolated snapshot of early 1926. A brief review of the sources of stock returns over the past two centuries should help lay a foundation for our work on return expectations and shatter a few widespread misconceptions in the process. The sources of the data are given in Appendix A.[11]

Step I: How Well Does Economic Growth Flow into Dividend Growth? Over the past 131 years, since reliable earnings data became available in 1870, the average earnings yield has been 7.6 percent and the average real return for stocks has been 7.2 percent; this close match has persuaded many observers to the view (which is wholly consistent with finance theory) that the best estimate for real returns is, quite simply, the earnings yield. On careful examination, this hypothesis turns out to be wrong. In the absence of changing valuation levels, real returns are systematically *lower* than earnings yields.

Figure 1 shows stock market returns since 1802 in a fashion somewhat different from that shown in most of the literature. The solid line in Figure 1 shows the familiar cumulative total return for U.S. equities since 1802, in which each $100 invested grows, with reinvestment of dividends, to almost $700 million in 200 years. To be sure, some of this growth came from inflation; as the line "Real Stock Return" shows, $700 million will not buy what it

Figure 1. Return from Inflation and Dividends: Growth of $100, 1801–2001

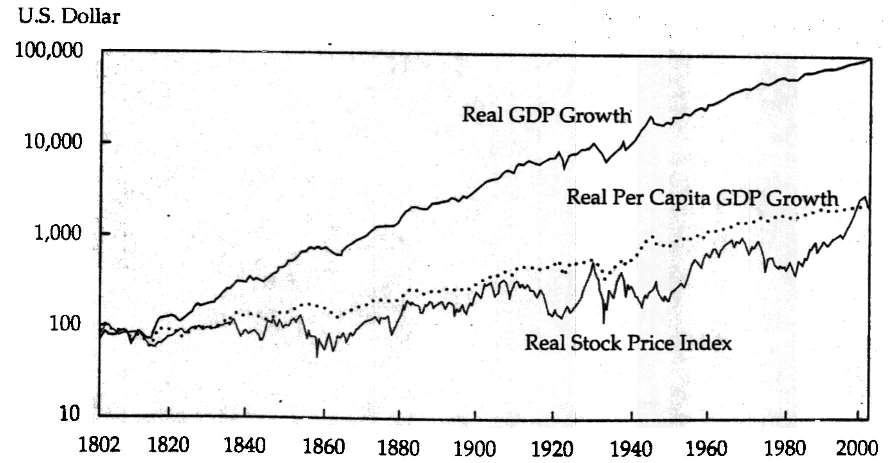


would have in 1802, when one could have purchased the entire U.S. GNP for less than that sum.[12] By removing inflation, we show in the "Real Stock Return" line that the $100 investment grew to "only" $37 million. Thus, adjusted for inflation, our fortune is much diminished but still impressive. Few portfolios are constructed without some plans for future spending, and the dividends that stocks pay are often spent. So, the "Real Stock Price Index" line shows the wealth accumulation from price appreciation alone, net of inflation and dividends. This bottom line (literally and figuratively) reveals that stocks have risen just 20-fold from 1802 levels. Put another way, if an investor had placed $100 in stocks in 1802 and received and spent the average dividend yield of 4.9 percent for the next 200 years, his or her descendants would today have a portfolio worth $2,099, net of inflation. So much for our $700 million portfolio!

Worse, the lion's share of the growth from $100 to $2,099 occurred in the massive bull market from 1982 to date. In the 180 years from 1802 to the start of 1982, the real value of the $100 portfolio had grown to a mere $400. If stocks were priced today at the same dividend yields as they were in 1802 and 1982, a yield of 5.4 percent, the $100 portfolio would be worth today, net of inflation and dividends, just $550. These data put the lie to the conventional view that equities derive most of their returns from capital appreciation, that income is far less important, if not irrelevant.

Figure 2 allows a closer look at the link between equity price appreciation and economic growth. It shows that the growth in share prices is much more closely tied to the growth in real *per capita* GDP (or GNP) than to growth in real GDP per se. The solid line shows that, compounding at about 4 percent in the 1800s and 3 percent in the 1900s, the economy itself delivered an impressive 1,000-fold growth.

Figure 2. The Link between Stock Prices and Economic Growth, 1802–2001



But net of inflation and dividend distributions, stock prices (the same "Real Stock Price Index" line in Figure 1) fell far behind, with cumulative real price appreciation barely 1/50 as large as the real growth in the economy itself.

How can this be? Can't shareholders expect to participate in the growth of the economy? No. Shareholders can expect to participate *only in the growth of the enterprises they are investing in.* An important engine for economic growth is the creation of new enterprises. The investor in today's enterprises does not own tomorrow's new enterprises—not without making a separate investment in those new enterprises with new investment capital.

Finally, the "Real Per Capita GDP Growth" line in Figure 2 shows the growth of the economy measured net of inflation *and population growth.* This growth in real per capita GDP tracks much more closely with the real price appreciation of stocks (the bottom line) than does real GDP itself.

Going one step further, **Figure 3** shows the internal growth of real dividends—that is, the growth that an index fund would expect to see in its own real dividends in the absence of additional investments, such as reinvestment of dividends.[13] Real dividends exhibit *internal* growth that is similar to the growth in real per capita GDP. Because growth in per capita GDP is a measure of productivity growth, the internal growth that can be sustained in a diversified market portfolio should closely match the growth of *productivity* in the economy, not the growth in the economy per se. Therefore, the dotted line traces per capita real GDP growth, the "Real Stock Price Index" line shows real stock prices, and the bottom line shows real dividends (× 10).[14] Figure 3 reveals the remarkable resemblance between real dividend growth and growth in real per capita GDP.

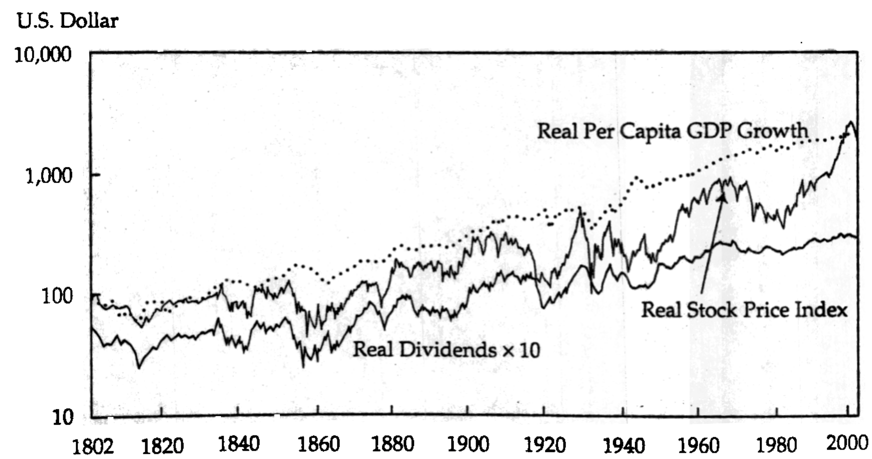
When we measure the internal growth of real dividends as in Figure 3, we see that real dividends have risen a modest fivefold from 1802 levels. In other words, the real dividends for a $100 portfolio invested in 1802 have grown merely 0.9 percent a year net of inflation. To be sure, the price assigned to each dollar of dividends has quadrupled, which leads to the 20-fold real price gain in the 200 years.

Although real dividends have tracked remarkably well with real per capita GDP, they have consistently fallen short of GDP gains. Not only have real dividends failed to match real GDP growth (as many equity investors seem to think is a *minimal* future growth rate for earnings and dividends), they have even had a modest shortfall, at an average of about 70 bps a year, relative to per capita economic growth.

In short, more than 85 percent of the return on stocks over the past 200 years has come from (1) inflation, (2) the dividends that stocks have paid, and (3) the rising valuation levels (rising P/Es and falling dividend yields) since 1982, not from growth in the underlying fundamentals of real dividends or earnings.[15] Furthermore, real dividends and real per capita GDP both grew faster in the 20th century than in the 19th century. Conversely, GDP grew faster in the 19th century than in the 20th century, *unless* we convert to per capita GDP.

Many observers think that earnings growth is far more important than dividend growth. We respectfully disagree. As noted by Hicks (1946), "... any increase in the present value of prospective net receipts must raise profits." In other words, properly stated, earnings should represent a proportional share of the net present value of all future

Figure 3. Dividends and Economic Growth, 1802–2001



profits. The problem is that reported earnings often do not follow this theoretical definition. For example, negative earnings should almost never be reported, yet reported operating losses are not uncommon. Furthermore, the quality of earnings reports prior to the advent of the U.S. SEC is doubtful at best; worse, we were unable to find any good source for earnings information prior to 1870. Accordingly, the dividend is the one reliable aspect of stock ownership over the past two centuries. It is the cash income returned to the shareholders; it is the means by which the long-term investor earns most of his or her internal rate of return. Finally, with earnings growth barely 0.3 percent faster than dividend growth over the past 131 years, an analysis based on earnings would reach conclusions nearly identical to our conclusions based on dividends.

Finance theory tells us that capital is fungible; that is, equity and debt, retained earnings and dividends—all should flow to the best use of capital and should (in the absence of tax-related arbitrages and other nonsystematic disruptions) produce a similar risk-adjusted return on capital. Thus, the retained earnings should deliver a return similar to the return an investor could have earned on that capital had it been paid out as dividends. Consider an example: If a company has an earnings yield of 5 percent (corresponding to a P/E of 20), it can pay out all of the earnings and thereby deliver a 5 percent yield to the shareholder. The real value of the company should not be affected by this full earnings distribution (unless the earnings are themselves being misstated), so the 5 percent earnings yield should also be the expected real return. Now, if the company, instead, pays a 2 percent yield and retains earnings worth 3 percent of the stock price, the company ought to achieve 3 percent real growth in earnings; otherwise, it should have distributed the cash to the shareholders. How does this theory stand up to reality?

Over the past 200 years, dividend yields have averaged 4.9 percent, yet real returns have been far higher, 6.6 percent. Since 1870, earnings yields have averaged 7.6 percent, close to the real returns of 7.2 percent over that span. This outcome is consistent with the notion of fungible capital, that the return on capital reinvested in an enterprise ought to match the return an investor might otherwise have earned on that same capital if it had been distributed as a dividend. However, if we take out the changes in valuation levels since 1982 (regardless of whether dividend yields or P/Es are used for those levels), the close match between earnings yield and real stock returns evaporates.

Moreover, with an average earnings yield of 7.6 percent and an average dividend yield of 4.7 percent since 1871, the average "retained earnings yield" has been nearly 3 percent. This retained earnings yield should have led to real earnings and dividend growth of 3 percent; otherwise, management ought to have paid this money out to the shareholders. Instead, real dividends and earnings grew at annual rates of, respectively, 1.2 percent and 1.5 percent. Where did the money go? The answer is that during the era of "pirate capitalism," success often led to dilution: Company managers issued themselves more stock![16]

Furthermore, retained earnings often chase poor internal reinvestment opportunities. If existing enterprises experienced only 1.2–1.5 percent internal growth of real dividends and earnings in the past two centuries, most of the 3.6 percent economic growth the United States has enjoyed has clearly not come from reinvestment in existing enterprises. In fact, it has stemmed from entrepreneurial capitalism, from the creation of new enterprises. Indeed, dividends on existing enterprises have fallen relative to GDP growth by approximately 100-fold in the past 200 years.[17]

The derring-do of the pirate capitalists of the 19th and early 20th centuries is not the only or even the most compelling explanation for this phenomenon. All the data we used are from indexes, which are a particular kind of sampling of the market. Old companies fading from view lose their market weight as the newer and faster growing companies gain a meaningful share in the economy. The older enterprises often have the highest earnings yield and the worst internal reinvestment opportunities, but the new companies do not materialize in the indexes the minute they start doing business or even the minute they go public. When they do enter the index, their starting weight is often small.

Furthermore, an index need only change the divisor whenever a new enterprise is added, whereas we cannot add a new enterprise to our portfolio without cost. The index changing the divisor is mathematically the same as selling a little bit of *all other* holdings to fund the purchase of a new holding, but when we add a new enterprise to our portfolios, we must commit some capital to effect the purchase. Whether through reinvestment of dividends or infusion of new capital, this new enterprise cannot enter our portfolio through the *internal* growth of an existing portfolio of assets. In effect, we must rebalance out of existing stocks to make room for the new stock—which produces the natural dilution that takes place as a consequence of the creation of new enterprises in a world of entrepreneurial capitalism: The same dollar cannot own an existing enterprise and simultaneously fund a new enterprise.[18]

The dynamics of the capitalist system inevitably lead to these kinds of results. Good business leads to expansion; in a competitive environment, expansion takes place on a wide scale; expansion on a wide scale intensifies the competitive environment; margins begin to decline; earnings growth slows; in time, earnings begin to decline; then, expansion slows, profit margins improve, and the whole thing repeats itself. We can see this drama playing out in the relationship between payout ratios in any given year and earnings growth: Since 1984, the payout ratio has explained more than half of the variation in five-year earnings growth rates with a t-statistic of 9.51.[19]

Few observers have noticed that much of the difference between stock dividend yields and the real returns on stocks can be traced directly to the upward revaluation of stocks since 1982. The historical data are muddied by this change in valuation levels—which is why we find the current fashion of forecasting the future by extrapolating the past to be so alarming. The earnings yield is a better estimate of future real stock returns than any extrapolation of the past. And the dividend yield plus a small premium for real dividend growth is even better, because in the absence of changes in valuation levels, the earnings yield systematically overstates future real stock returns.

If long-term real growth in dividends had been 0.9 percent, real stock returns would have been only 90 bps higher than the dividend yield if it were not for the enormous jump in the price-to-dividend ratio since 1982. Even if we adjust today's 1.4 percent dividend yield sharply upward to include "dividends by another name" (e.g., stock repurchases), making a case for real returns higher than the 3.4 percent currently available in the TIPS market would be a stretch.[20]

Step II: Estimating Real Stock Returns. To estimate the historical equity risk premium, we must compare (1) a realistic estimate of the *expected* real stock return that objective analysis might have supported in past years with (2) the *expected* real bond return available at the time. Future long-term real stock return is defined as[21]

$$RSR(t) = DY(t) + RDG(t) + \Delta PD(t) + \varepsilon, \tag{1}$$

where

$DY(t)$ = percentage dividend yield for stocks at time t

$RDG(t)$ = percentage real dividend growth rate over the applicable span starting at time t

$\Delta PD(t)$ = percentage change in the price assigned to each dollar of dividends starting at time t

ε = error term for sources of return not captured by the three key constituents (this term will be small because it will reflect only compounding effects)

Viewed from the perspective of forecasting future real returns, the $\Delta PD(t)$ term is a valuation term, which we deliberately exclude from our analysis. If markets exhibit reversion to the mean, valuation change should be positive when the market is inexpensive and negative when the market is richly priced. If markets are efficient, this term should be random. We choose not to go down the slippery slope of arguing valuation, even though we believe that valuation matters. Rather, we prefer to make the simplifying assumption that market valuations at any stage are "fair" and, therefore, that the real return stems solely from the dividend yield and real growth of dividends.

That said, the estimation process becomes more complex when we consider a sensible estimate for real dividend growth. For example, what real dividend growth rate might an investor in 1814 have expected on the heels of the terrible 1802–14 bear market and depression, during which real per capita GDP, real dividends, and real stock prices all contracted 40–50 percent? How can we objectively put ourselves in the position of an investor almost 200 years ago? For this purpose, we partition the real growth in dividends into two constituent parts, real economic growth and the growth of dividends relative to the economy.

Why not simply forecast dividend growth directly? Because countless studies have shown that analysts' forecasts are too optimistic, especially at market turning points. In fact, dividends (and earnings) in aggregate cannot grow as fast as the economy on a sustainable long-term basis, in large part because of the secular increase in shares outstanding and introduction of new enterprises. So, long-term dividend growth should be equal to long-term economic growth minus a haircut for dilution or entrepreneurial capitalism (the share of economic growth that is tied to new enterprises not yet available in the stock market) or plus a premium for hidden dividends, such as stock buybacks. So, real dividend growth is given by

$$RDG(t) = RGDP(t) + DGR(t) + \varepsilon, \tag{2}$$

where

$RGDP(t)$ = percentage real per capita GDP growth over the applicable span starting at time t

$DGR(t)$ = annual percentage dilution of real GDP growth as it flows through to real dividends starting at time t

ε = error term for compounding effects (it will be small)

Basically, in Equation 2, we are substituting $RGDP(t) + DGR(t)$ for $RDG(t)$ and rolling the $\Delta PD(t)$ term into the error term (to avoid getting into the debates about valuation and regression to the mean). With these two changes, and converting to an expectations model, our model for expected real stock market returns, *ERSR*, becomes

$$ERSR(t) = EDY(t) + ERGDP(t) + EDGR(t), \quad (3)$$

where

$EDY(t)$ = *expected* percentage dividend yield for stocks at time t

$ERGDP(t)$ = *expected* percentage real per capita GDP growth over the applicable span starting at time t

$EDGR(t)$ = *expected* annual percentage dilution of real per capita GDP growth as it flows through to real dividends starting at time t

A complication in this structure is the impact of recessions. In serious recessions, dividends are cut and GDP growth stops or reverses, possibly leading to a decline in even the long-term GDP growth. The result is a dividend yield that is artificially depressed, real per capita GDP growth that is artificially depressed, and long-term dividend growth relative to GDP growth that is artificially depressed, all three of which lead, in recessionary troughs, to understated expected real stock returns. The simplest way to deal with this issue is to use the last peak in dividends before a business downturn and the last peak in GDP before a business downturn in computing each of the three constituents of expected real stock returns.[22]

We illustrate how we constructed an objective real stock return forecast for the past 192 years in **Figure 4**; Panel A spans 1810 to 2001, and Panel B shows the same data after 1945. To explain these graphs, we will go through them line by line.

The easiest part of forecasting real stock returns, the "Estimated Real Stock Return" line in Figure 4, is the dividend yield: It is a known fact. We have adjusted dividends to correct for the artificially depressed dividends during recessions to get the $EDY(t)$ term shown as the "Dividend Yield" line in Figure 4. This step allows us to avoid understating the equity risk premium in recessions when dividends are artificially depressed. This adjustment boosts the expected dividend yield slightly relative to the raw dividend yield because the deepest recessions are often deeper than the average recessions of the prior 40 years. Against an average dividend yield of 4.9 percent, we found an average *expected* dividend yield of 5.0 percent.

Most long-run forecasts of earnings or dividend growth ignore the simple fact that aggregate earnings and dividends in the economy cannot sustainably grow faster than the economy itself. If new enterprise creation and secondary equity offerings dilute the share of the economy held by the shareholders in existing enterprises, then one sensible way to forecast dividend growth is to forecast economic growth and then forecast how rapidly this dilution will take place.[23] Stated another way, we want to know how much *less* rapidly dividends (and earnings) on existing enterprises can grow than the economy at large. The sum of real economic growth less this shortfall is the real growth in dividends.

The resulting line, "Dilution of GDP Growth in Dividends," in the two graphs of Figure 4 represents the $EDGR(t)$ term in our model (Equation 3). Note the persistent tendency for dividend growth to lag GDP growth: Real dividends have grown at 1 percent a year over the past 192 years, whereas the real economy has grown at 3.8 percent a year, and even real per capita GDP has grown at 1.8 percent a year. Why should real dividends have grown so much more slowly than the economy?

First, much of the growth in the economy has come from innovation and entrepreneurial capitalism. More than half of the capitalization of the Russell 3000 today consists of enterprises that did not exist 30 years ago. The 1971 buy-and-hold investor could not participate in this aspect of GDP growth or market growth because the companies did not exist. So, today's dividends and earnings on the existing companies from 1971 are only part of the dividends and earnings on today's total market.

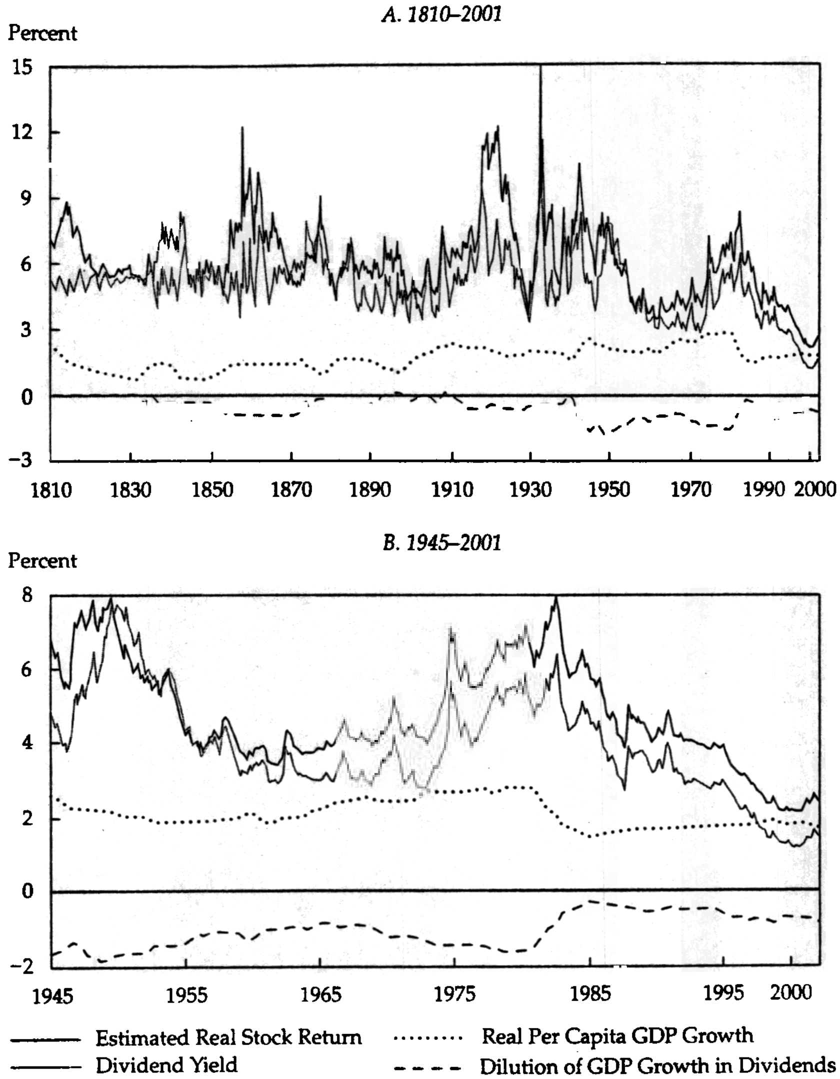
Second, as was demonstrated in Bernstein (2001b), retained earnings are often not reinvested at a return that rivals externally available investments; earnings and dividend growth are faster when payout ratios are high than when they are low, perhaps because corporate managers are then forced to be more selective about reinvestment alternatives.[24]

Finally, as we have emphasized, corporate growth typically leads to more shares outstanding, which automatically imposes a drag on the growth in dividends per share.

As a sensible estimate of the future dividend/GDP shortfall, the rational investor of any day might forecast dividend growth by using the prior 40-year shortfall in dividend growth relative to per capita GDP or might choose to use the cumulative (by now, 200-year) history. We chose the simple expedient of averaging the two.

The dilution effect we found from the 40-year and cumulative data for real dividends and real per capita GDP averages –60 bps. So, in the past 40 years, the dilution of dividend growth is almost

Figure 4. Estimating Real Stock Returns



exactly the same as the long-term average, –80 bps. With a standard deviation of just 0.5 percent, this shortfall of dividend growth relative to economic growth is the steadiest of any of the components of real stock returns or real bond returns. It has never been materially positive on a long-term sustained basis; it has never risen above +10 bps for any 40-year span in the entire history since 1810.

The history of dividend growth shows no evidence that dividends can ever grow materially faster than per capita GDP. Indeed, they almost always grow more slowly. Suppose real GDP growth in the next 40 years is 3 percent a year and population growth is 1 percent a year. These assumptions would appear to put an *upper limit* on real dividend growth at a modest 2 percent a year, far below consensus expectations. If the historical average dilution of dividend growth relative to real per capita GDP growth prevails, then the future real growth in dividends should be only about 1 percent, even with relatively robust, 2.5–3.0 percent, real GDP growth.

Now consider the third part of forecasting real stock returns in this fashion—the forecast of long-term real per capita GDP growth, *ERGDP(t)* in our model. How much real per capita GDP growth would an investor have expected at any time in the past 200 years? Again, a simple answer might come from the most recent 40 years' growth rate; another might come from the cumulative record going back as far as we have dividend and GDP data, to 1802. These historical data are shown in the "Real per Capita GDP Growth" line in Figure 4. And again, we chose the simple expedient of averaging the average of the two. Real per capita GDP growth has been remarkably stable over the past 200 years, particularly if we adjust it to correct for temporary dips during recessions. If we examine truly long-term results, the 40-year real growth rate in real per capita GDP has averaged 1.8 percent with a standard deviation of only 0.9 percent.[25]

Note from Figure 4 that the total economy grew faster during the 19th century than the 20th century whereas stock returns (and the underlying earnings and dividends) grew faster in the 20th century than the 19th. Why would the rapid growth of the 19th century flow through to the shareholder less than the slower growth of the 20th century? We see two possible answers. First, the base from which industrial growth started in the 19th century was so much smaller that much faster new enterprise creation occurred then than in the 20th century. Second, with nearly 3 percent growth in the population from 1800 to 1850, the growing talent and labor pool fueled a faster rate of growth than the 1.25 percent annual population growth rate of the most recent 50 years. It is not surprising that the pace of dilution, both from the creation of new enterprises and from secondary equity offerings, is faster when the population is growing faster. Population growth fuels growth in human capital, in available labor, and in both demand *and* supply of goods and services. As a result, when population growth is rapid, the pace of dilution of growth in the economy (as it flows through to a shareholder's earnings and dividends) is far more stable relative to real per capita GDP than relative to real GDP itself.

The simple framework we have presented for estimating real stock returns reveals few surprises. As Panels A and B of Figure 4 show, the expected stock return is the sum of the three constituent parts graphed in the other lines. We estimate that expected real stock returns for the past 192 years averaged about 6.1 percent with the following constituent parts: an expected yield averaging 5.0 percent plus real per capita GDP growth of 1.7 percent a year minus an expected *shrinkage* in dividends relative to real per capita GDP averaging –0.6 percent. Meanwhile, investors actually earned real returns of 6.8 percent. Most of this 70 bp difference from the 6.1 percent rational expectation over the past 192 years can be traced to the rise in valuation levels since 1982; the rest consists of the other happy accidents detailed previously.

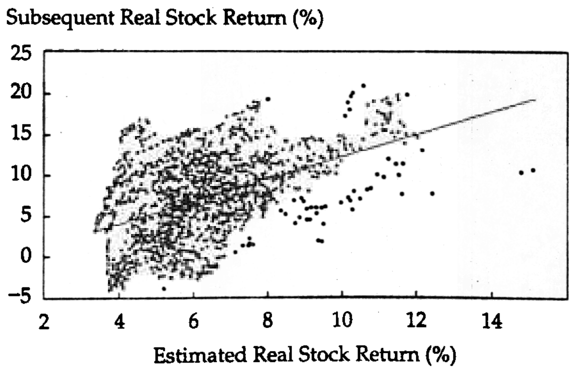
Expectations for real stock returns have soared above 6 percent often enough that many actuaries even today consider 8 percent a "normal" real return for equities. Our estimate for real stock returns, however, exceeds 8 percent only during the depths of the Great Depression, in the rebuilding following the War of 1812, the Civil War, World War I, and World War II, and in the Crash of 1877. In the past 50 years, expected real stock returns above 7 percent have been seen only in the aftermath of World War II, when many investors still feared a return to Depression conditions, and in the depths of the 1982 bear market.

When viewed from the vantage point of this formulation for expected real stock returns, the full 192-year record shows that expected real stock returns fell below 3.5 percent only once before the late 1990s, at the end of 1961 just ahead of the difficult 1962–82 span, real stock prices fell by more than 50 percent. Since 1997, expected real stock returns have fallen well below the 1961 levels, where they remain at this writing.

This formulation for expected real stock returns reveals the stark paradigm shift that took place in the 1950s. Until then, the best estimate for real dividend growth was rarely more than 1 percent, so the best estimate for real stock returns was approximately the dividend yield plus 100 bps—considerably less than the earnings yield! From the 1950s to date, as Panel B of Figure 4 shows, the shortfall of dividends relative to GDP growth improved (perhaps because the presence of the SEC discourages company managers from ignoring shareholder interests) and the real return that one could objectively expect from stocks finally and persuasively rose above the dividend yield. Today, it stands at almost twice the dividend yield, but it is still a modest 2.4 percent.

Figure 5 shows the strong correlation between our formulation for expected real stock returns and the actual real returns that stocks have delivered over the subsequent 10-year span. The correlation is good—at 0.62 during the modern market era after World War II and 0.46 for the full 182 years.[26] If we test the correlation between this simple metric of expected real stock returns and the actual subsequent 20-year real stock returns (not shown), the correlations grow to 0.95 and 0.60 for the post-1945 period and the full 182 years, respectively.

Figure 5. Estimated and Subsequent Actual Real Stock Returns, 1802–2001



The regression results given in Panel A **Table 1** show that the coefficient in the regression is larger than 1.00. So, that 100 bp increase in the expected real stock return, *ERSR*, is worth more than 100 bps in the subsequent 10-year actual real stock return, *RSR*. The implication is that some tendency for reversion to the mean does exist and that it will magnify the effect of unusually high or low expected real stock returns. This suggestion has worrisome implications for the recent record low levels for expected real stock returns.

Because rolling 10-year returns (and expected returns in our model) are highly serially correlated, the *t*-statistics given in Panel A of Table 1 are not particularly meaningful. One way to deal with overlapping data is to eliminate the overlap by using nonoverlapping samples—in this case, examining only our 19 nonoverlapping samples beginning December 1810. The Panel B results, with a coefficient larger than 1.00, confirm the previous results (and approach statistical significance, even with only 17 degrees of freedom).[27] One worrisome fact, in light of the recent large real stock returns, is that the nonoverlapping real stock returns by decades have a –31 percent serial correlation. Although it is not a statistically significant correlation, it is large enough to be interesting: It suggests that spectacular decades or wretched decades may be considerably more likely to reverse than to repeat.

Evaluating the real returns on stocks is clearly a useful exercise if the metric of success for a model is subsequent actual real returns, but we live in a relative world. The future real returns on all assets will rise and fall; so, real returns are an insufficient metric of success. What is of greater import is whether this metric of prospective real stock returns helps us identify the attractiveness of stocks *relative to other assets*.

Step III: Estimating Future Real Bond Returns. On the bond side, real realized returns are equal to the nominal yield minus inflation (or plus deflation) and plus or minus yield change times duration:

$$RBR(t) = BY(t) - INFL(t) + \Delta BY(t)DUR(t) + \varepsilon, \quad (4)$$

where

$BY(t)$ = percentage bond yield at time t

$INFL(t)$ = percentage inflation over the applicable span starting at time t

$\Delta BY(t)DUR(t)$ = annual change in yield over the applicable span times duration at time t (under the assumption that rolling reinvestment is in bonds of similar duration)

ε = error term (compounding effects lead to a small error term in this simple formulation)

As with stocks, we prefer to take current yields as a fair estimate of future bond yields. So, we eliminate the variable that focuses on changes in yields, $\Delta BY(t)DUR(t)$. We also need to shift our focus from measuring past real bond returns to forecasting future real bond returns. Therefore, our model is

$$ERBR(t) = BY(t) - EINFL(t), \quad (5)$$

where $BY(t)$ is the percentage bond yield at time t and $EINFL(t)$ is the expected percentage inflation over the applicable span starting at time t.

Equation 5 is difficult only in the sense that expectations for inflation in past economic environs are difficult to estimate objectively. How, for example, are we to gauge how much inflation an investor in February 1864 would have expected at a time when inflation had averaged 20 percent over the prior three years because of wartime shortages?

Table 1. Regression Results: Estimated Real Stock Return versus Actual 10-Year Real Stock Return
(*t*-statistics in parentheses)

Period	*a*	*b*	R^2	Correlation	Serial Correlation
A. Raw data: RSR(t) = a + b[ERSR(t – 120)]					
1810–2001	–1.51%	1.38%	0.214	0.46	0.992
	(–4.2)	(24.4)			0.990
1945–2001	–7.80	3.15	0.391	0.62	0.996
	(–8.8)	(19.0)			0.995
B. Using 19 nonoverlapping samples, beginning December 1810					
1810–2000	–0.35%	1.22%	0.182	0.430	–0.315
	(–0.1)	(1.9)			0.021

Expectations would depend strongly on the outcome of the war: A victory by the North would have been expected to result in a restoration of the purchasing power of the dollar as wartime shortages disappeared; a victory by the South could have had severe consequences on the ultimate purchasing power of the North's dollar as a consequence of debt that could no longer be serviced. A rational expectation might have been for inflation greater than 0 (reflecting the possibility of victory by the South) but less than the 20 percent three-year inflation rate (reflecting the probability of victory by the North).

We based the estimate for expected future inflation on an *ex ante* regression forecast of 10-year future inflation based, in turn, on recent three-year inflation.[28] **Figure 6** shows how the expected rate of inflation has steadily become more closely tied to recent actual inflation in recent decades. Bond yields responded weakly to bursts of inflation up until the time of the Great Depression; they responded more strongly as inflation became a structural component of the economy in the past four decades.

Until the last 40 years, inflation was generally associated with wars and was virtually nonexistent—even negative—in peacetime. Figure 6 shows a burst of double-digit inflation on the heels of the War of 1812, in the late stages of the Civil War, during World War I, and in the rebuilding following World War II. And more recently, double-digit inflation characterized the "stagflation" of 1978–1981 that followed the Vietnam War and the oil shocks of the 1970s. The most notable changes since the Great Depression, especially since World War II, involve the magnitude and perceived role of government and loss of the automatic brakes once applied by the gold standard. From the end of World War II to the great inflationary crisis at the end of the 1970s, the dread of unemployment that was inherited from the Great Depression was the driving factor in both fiscal and monetary policy.

With the introduction of TIPS in January 1997, we finally have a U.S. government bond that pays a real return, which allows us to simplify the expected real bond returns to be the TIPS yield itself from that date forward; that is,

$$ERBR(t) = YTIPS(t), \tag{6}$$

where $YTIPS(t)$ is the percentage TIPS yield at time t.

Figure 7 shows how the current government bond yield (the "Bond Yield" line) minus expected inflation ("Estimated Inflation") leads to an estimate of the real bond return and hence the long-term expected real bond return ("Estimated Real Bond Yield"), which is the estimate through March of 1998 and the TIPS yield thereafter.[29] From the Equation 5 (or, more recently, Equation 6) formulation, expected real bond returns averaged 3.7 percent over the full period, a very respectable real yield, given the limited risk of government bonds, and good recompense for an investor's willingness to bear some bond-price volatility. Investors may not always have viewed government debt as the rock-solid investment, however, that it is generally considered today.

The 3.7 percent real bond return consists of an average nominal bond yield of 4.9 percent minus an expected inflation rate of 1.2 percent. For comparison, the average actual inflation rate has been 1.4 percent. In the years after World War II, the rate of peacetime inflation embedded in investors' memory banks was essentially zero, perhaps even slightly negative. Consequently, bond investors kept expecting inflation to go away, despite its persistence at a modest rate in the 1950s and early 1960s and an accelerating rate thereafter. As a result, bonds were badly priced for reality during most of

Figure 6. Estimating Future Inflation, 1810–2001

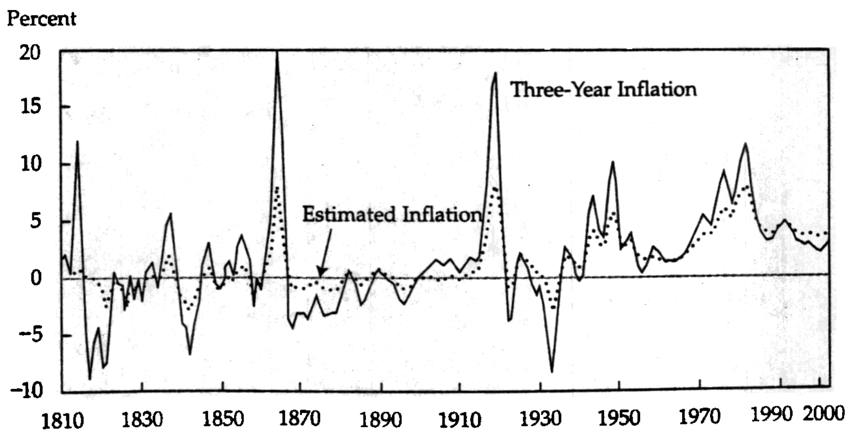
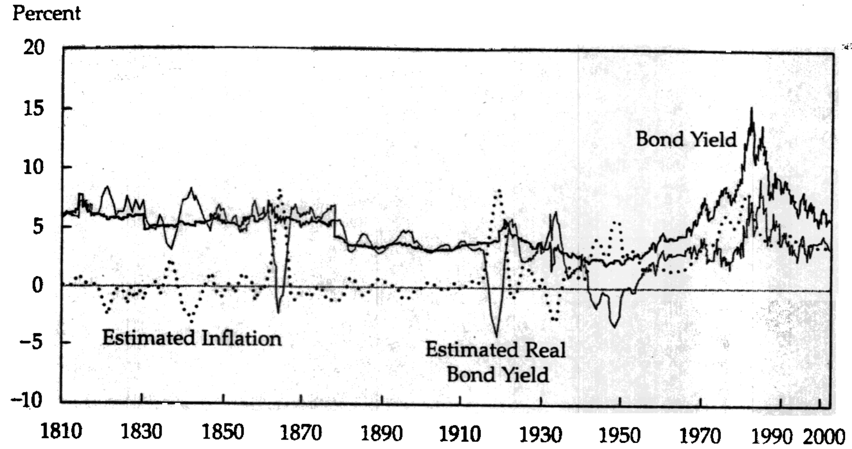


Figure 7. Estimating Real Bond Yields, 1810–2001



these two decades; they turned out to be certificates of confiscation for their holders until people finally woke up in the 1970s and 1980s. Actual inflation exceeded expected inflation with few exceptions from the start of World War II until roughly 1982; as can be seen in Figure 7, our model captures this phenomenon. Expectations are lower than actual outcomes during this span.

Figure 7 also shows several regimes of real yield with distinct structural change from one regime to the next. From the time the United States was in its infancy until the end of Reconstruction in the late 1870s, investors would not have viewed U.S. government bonds as a secure investment. They would have priced these bonds to deliver a 5–7 percent real yield, except during times of war. The overall stability of the yields is impressive: Unlike the history of stock prices, the surprise elements have been small.

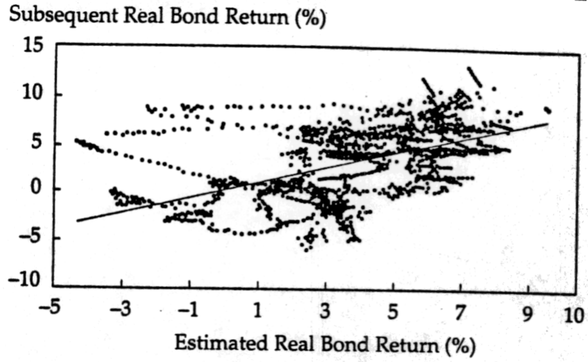
Once the United States had survived the Civil War and the security of U.S. government debt had been demonstrated repeatedly, investors began to price government debt at a 3–5 percent real yield. As Figure 7 shows, this level held, with a brief interruption in World War I, until the country went off the gold standard in 1933. This record is remarkable in view of the high rate of economic growth, but revolutionary technological change in those days, especially in transportation and agriculture, led to such stunning reductions in product costs that inflation was kept at bay except for very brief intervals.

For the next 20–25 years, the nation struggled with the Great Depression, World War II, and the war's aftermath. Investors slowly began to realize that deflationary price drops did not rebound fully after the trough of the Depression and that inflationary price increases did not retreat after the end of the war. The changed role of government plus the end of the gold standard had altered the picture, perhaps irrevocably. During this span, investors priced bonds to offer a 2–4 percent *notional* yield but a rocky −3 percent to +3 percent real yield. As Figure 7 shows, bond investors woke up late to the fact that inflation was now a normal part of life.

From the mid-1950s to date, investors have struggled with more structural inflation and more inflation uncertainty than ever before. Although investors sought to price bonds to deliver a real yield, inflation consistently exceeded their expectations. Only during the down cycle of the inflation roller coaster of 1980–1985 did bonds finally provide real yields to their owners. After this experience, bond investors developed an anxiety about inflation far greater than objective evidence would support. The result was a brief spike in real bond returns in 1984, as Figure 7 shows, with bond yields still hovering at 13.8 percent, even though three-year inflation had fallen to 4.7 percent (and our regression model for future inflation would have suggested expected inflation of 4.6 percent). The "expected" real yield was a most unusual 9.2 percent because investors were not yet prepared to believe that double-digit inflation was a thing of the past.

Another interesting fact is evident in **Figure 8**: The expected real bond returns produced by our formulation are highly correlated with the actual real returns earned over the subsequent decade. For 1810 to 1991, the expected real bond return has a 0.52 correlation with the actual real bond return earned over the next 10 years; from 1945 to date, the correlation rises to an impressive 0.63. Panel A of **Table 2** shows that the coefficient is reliably positive but not reliably more than 1.00, which suggests that, unlike expected real stock returns, no powerful tendency for reversion to the mean is at work in real bond yields. When we used the 19 available nonoverlapping samples (Panel B), we found the resulting correlation to be 0.64, which is a statistically significant relationship.[30]

Figure 8. Estimated and Subsequent Actual Real Bond Yield, 1802–2001



Why is the bond model a better predictor, when raw data are used, than the stock model for the two-century history? Two reasons seem evident. First, stocks have been more volatile than bonds for almost all 200 years of U.S. data. Therefore, any model for expected real stock returns should have a larger error term. Second, stocks are by their very nature longer term than bonds: A 10-year bond expires in 10 years; stocks have no maturity date.

The bond market correlations would be even better were it not for the negative real yields during times of war, when people tend to consider the inflation a temporary phenomenon. These episodes show up as the "loops" to the left of the body of the scatterplot in Figure 8. At these times, many U.S. investors apparently subordinated their own interests in a strong real yield to the needs of the nation: Long Treasury rates were essentially pegged during World War II and up to 1951, but that did not stop investors from buying them.

Step IV: Estimating the Equity Risk Premium. If we now take the difference between the expected real stock return and the expected real bond return, we are left with the expected equity risk premium:

$$ERP(t) = ERSR(t) - ERBR(t), \tag{7}$$

where $ERSR(t)$ is the expected real stock return starting at time t and $ERBR(t)$ is the expected real bond return starting at time t.

Figure 9 shows the results of this simple framework for estimating the risk premium over the past 192 years. Many observers may be startled to see that this estimate of the forward-looking risk premium for stocks has rarely been above 5 percent in the past 200 years; the exceptions are war, its aftermath, and the Great Depression. The historical average risk premium is a modest 2.4 percent, albeit with a rather wide range. The wide range is more a result of the volatility of expected real bond returns than the volatility of expected real stock returns, which are surprisingly steady except in times of crisis.[31]

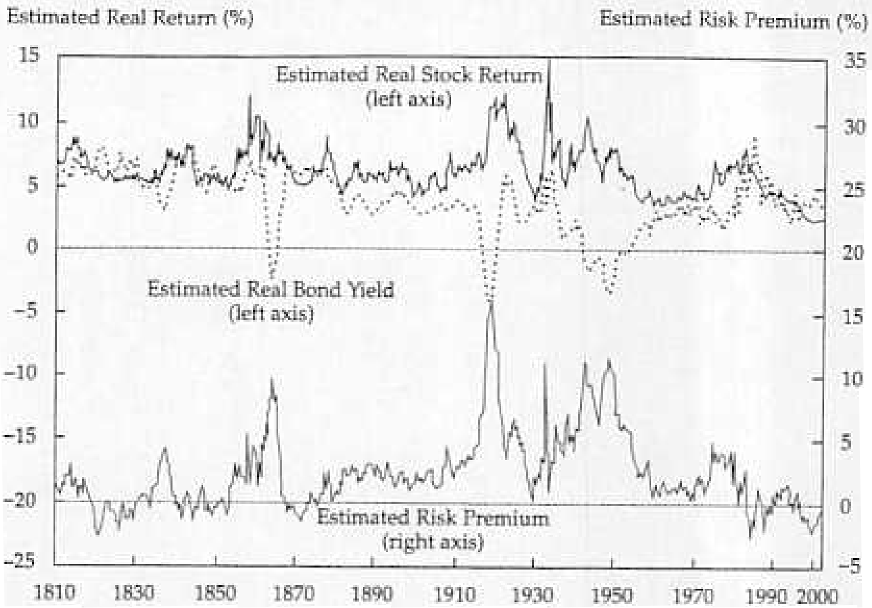
Over the past 192 years, our model (Equation 3) suggests that an objective evaluation would have pegged expected real stock returns at about 6.1 percent on average, only 120 bps higher than the average dividend yield. Investors have earned fully 70 bps more than this objective expectation, but they did not have objective reasons to expect to earn as much as they did. Our model suggests that an objective evaluation would have pegged expected real bond returns at about 3.7 percent. Investors have earned 20 bps less because of the inflationary shocks of the 1960s to 1980s; they expected more than they got.

The difference between the expected real returns for stocks and bonds reveals a stark reality. An objective estimate of the expected risk premium would have averaged 2.4 percent (240 bps) during this history (6.1 percent expected real stock returns minus 3.7 percent expected real bond returns), not the oft-cited 5 percent realized excess return that

Table 2. Regression Results: Estimated Real Bond Return versus Actual 10-Year Real Bond Return
(*t*-statistics in parentheses)

Period		b	R^2	Correlation	Serial Correlation
A. Raw data: $RBR(t) = a + b[ERBR(t - 120)]$					
1810–2001	0.45%	0.81%	0.266	0.52	0.999
	(3.5)	(28.1)			0.997
1945–2001	–0.74	1.05	0.399	0.63	0.997
	(–4.0)	(19.3)			0.980
B. Using 19 nonoverlapping samples, beginning December 1810					
1810–2001	–1.81%	1.31%	0.4120	0.64	0.182
	(–1.1)	(3.5)			0.677

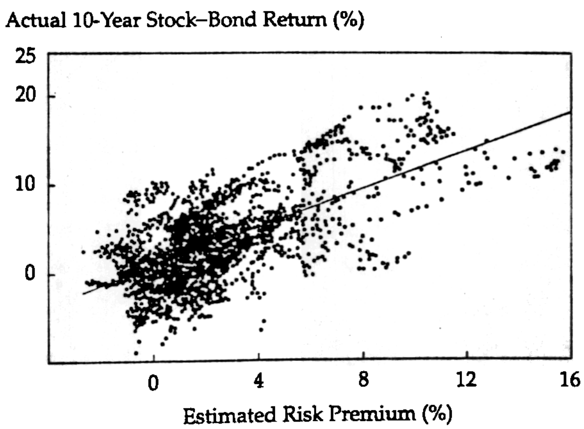
Figure 9. Estimating the Equity Risk Premium, 1810–2001



much of the investment world now depends on. Investors have *earned* a higher 3.3 percent (330 bps) excess return for stocks (6.8 percent actual real stock returns minus 3.5 percent for bonds), but the reason is the array of happy accidents for stocks and one extended unhappy accident for bonds.

All of this analysis is of mere academic interest, however, unless we can establish a link between our estimated risk premium and actual subsequent relative returns. Indeed, such a link does exist. The result of our formulation for the equity risk premium has a 0.79 correlation with the actual 10-year excess return for stocks over bonds since 1945 and a 0.66 correlation for the full span. This strong link is clear in **Figure 10**, for 1810–2001, and **Table 3**

Figure 10. Risk Premium and Subsequent 10-Year Excess Stock Returns: Correlations, 1810–1991



(where, for convenience, we have defined the 10-year excess return of stocks relative to bonds as *ERSB*); each 100 bp change in the equity risk premium is worth modestly more than 100 bps in subsequent annual excess returns for stocks relative to bonds over the next 10 years. As with the expected stock return model (Equation 3), the link for 20-year results is stronger, with correlations over the full span and since 1945 of, respectively, 0.64 and 0.95.

This strong link between objective measures of the risk premium and subsequent stock–bond excess returns is also clear for the 1945–2001 period shown in **Figure 11**, in which every wiggle of our estimate for the risk premium is matched by a similar wiggle in the subsequent 10-year excess return that stockholders earned relative to bondholders. Figure 11 shows that the excess returns on stocks relative to bonds became negative in the late 1960s on a 10-year basis, following low points in the risk premium, and again touched zero 10 years after the 1981 peak in bond yields.

We can also see in Figure 11 how the gap in 10-year results opened up sharply for the 10 years of the 1990s; it opened to unprecedented levels, even wider than in the early 1960s. Prior to this gap opening, the fit between the risk premium and subsequent excess returns is remarkably tight. The question is whether this anomaly is sustainable or is destined to be "corrected." History suggests that such anomalies are typically corrected, especially when the theoretical case to support them is so weak. This reminder should be sobering to investors who are depending on a large equity risk premium.

Table 3. Regression Results: Estimated Equity Risk Premium versus Actual 10-Year Excess Return of Stocks versus Bonds
(t-statistics in parentheses)

Period	a	b	R^2	Correlation	Serial Correlation
A. Raw data: ERSB(t) = a + b[ERP(t – 120)]					
1810–2001	0.91%	1.08%	0.430	0.66	0.993
	(8.8)	(40.6)			0.995
1945–2001	2.85	1.41	0.621	0.79	0.995
	(15.4)	(30.4)			0.996
B. Using 19 nonoverlapping samples, beginning December 1810					
1810–2001	0.84%	1.36%	0.490	0.70	0.055
	(0.8)	(4.0)			0.371

As with the models for real stock returns and for real bond returns, we also used nonoverlapping spans to take out the effect of the strong serial correlation in the estimated risk premium. For the 19 nonoverlapping spans (Panel B of Table 3), the correlation for the full period jumps to 0.70, with a highly significant t-statistic of 4.0.[32]

Conclusions

We have advanced several provocative assertions.

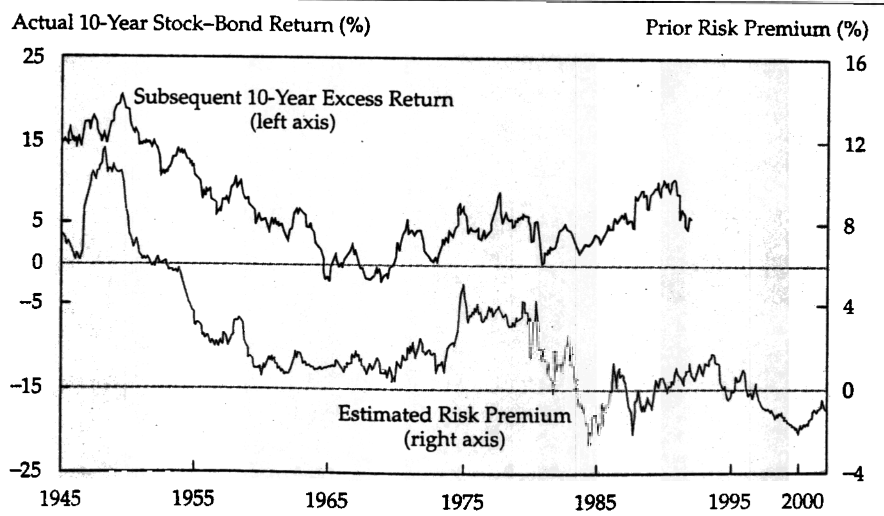
- The observed real stock returns and the excess return for stocks relative to bonds in the past 75 years have been extraordinary, largely as a result of important nonrecurring developments.
- It is dangerous to shape future expectations based on extrapolating these lofty historical returns. In so doing, an investor is tacitly assuming that valuation levels that have doubled, tripled, and quadrupled relative to underlying earnings and dividends can be expected to do so again.

The investors of 75 years ago would not have had an objective basis for expecting the 8 percent real returns or 5 percent risk premium that stocks subsequently delivered. The estimated equity risk premium at the time was above average, however, which makes 1926 a better-than-average starting point for the historical risk premium.

The real internal growth that companies generated in their dividends averaged 0.9 percent a year over the past 200 years, whereas earnings growth averaged 1.4 percent a year over the past 131 years.

Dividends and earnings growth was slower than the increase in real per capita GDP, which averaged 1.6 percent over the past 200 years and 2.0 percent over the past 131 years. This internal growth is far less than the consensus expectations for future earnings and dividend growth.

Figure 11. Risk Premium and Subsequent 10-Year Excess Returns, 1945–2001



- The historical average equity risk premium, measured relative to 10-year government bonds as the risk premium investors might objectively have expected on their equity investments, is about 2.4 percent, half what most investors believe.
- The "normal" risk premium might well be a notch lower than 2.4 percent because the 2.4 percent objective expectation preceded *actual* excess returns for stocks relative to bonds that were nearly 100 bps higher, at 3.3 percent a year.
- The current risk premium is approximately zero, and a sensible expectation for the future real return for both stocks and bonds is 2–4 percent, far lower than the actuarial assumptions on which most investors are basing their planning and spending.[33]
- On the hopeful side, because the "normal" level of the risk premium is modest (2.4 percent or quite possibly less), current market valuations need not return to levels that can deliver the 5 percent risk premium (excess return) that the Ibbotson data would suggest. If reversion to the mean occurs, then to restore a 2 percent risk premium, the difference between 2 percent and zero still requires a near halving of stock valuations or a 2 percent drop in real bond yields (or some combination of the two). Either scenario is a less daunting picture than would be required to facilitate a reversion to a 5 percent risk premium.
- Another possibility is that the modest difference between a 2.4 percent normal risk premium and the negative risk premiums that have prevailed in recent quarters permitted the recent bubble. Reversion to the mean might not ever happen, in which case, we should see stocks sputter along delivering bondlike returns, but at a higher risk than bonds, for a long time to come.

The consensus that a normal risk premium is about 5 percent was shaped by deeply rooted naiveté in the investment community, where most participants have a career span reaching no farther back than the monumental 25-year bull market of 1975–1999. This kind of mind-set is a mirror image of the attitudes of the chronically bearish veterans of the 1930s. Today, investors are loathe to recall that the real total returns on stocks were negative for most 10-year spans during the two decades from 1963 to 1983 or that the excess return of stocks relative to long bonds was negative as recently as the 10 years ended August 1993.[34]

When reminded of such experiences, today's investors tend to retreat behind the mantra "things will be different this time." No one can kneel before the notion of the long run and at the same time deny that such circumstances will occur in the decades ahead. Indeed, such crises are more likely than most of us would like to believe. Investors greedy enough or naive enough to expect a 5 percent risk premium and to substantially overweight equities accordingly may well be doomed to deep disappointments in the future as the realized risk premium falls far below this inflated expectation.

What if we are wrong about today's low equity risk premium? Maybe real yields on bonds are lower than they seem. This chance is a frail reed to rely on for support. At this writing, at the end of 2001, an investor can buy TIPS, which provide government-guaranteed yields of about 3.4 percent, but inflation-indexed bond yields are a relatively recent phenomenon in the United States. So, we could not estimate historical real yields for prior years directly, only through a model such as the one described here. If we compare our model for real stock returns, at 2.4 percent in mid-2001, with a TIPS yield of 3.4 percent, we get an estimate for the equity risk premium of –100 bps.

Perhaps real earnings and dividend growth will exceed economic growth in the years ahead, or perhaps economic growth will sharply exceed the historical 1.6 percent real per capita GDP growth rate. These scenarios are certainly possible, but they represent the dreams of the "new paradigm" advocates. The scenarios are unlikely. Even if they prove correct, it will likely be in the context of unprecedented entrepreneurial capitalism, unprecedented new enterprise creation, and hence, unprecedented dilution of shareholders in existing enterprises.

The recurring pattern of history is that exceptionally poor or exceptionally rapid economic growth is never sustained for long. The best performance that dividend growth has ever managed, relative to real per capita GDP, is a scant 10 bp outperformance. This rate, the best 40-year real dividend growth *ever seen*, fell far short of real GDP growth: Real dividend growth was some 2 percent a year below real GDP growth during those same 40-year spans. So, history does not support those who hope that dividend growth will exceed GDP growth. This evidence is not encouraging for those who wish to see a 1.4 percent dividend yield somehow transformed into a 5 percent (or higher) real stock return.

The negative risk premium that precipitated the writing of "The Death of the Risk Premium" (Arnott and Ryan) in early 2000 was not without precedent, although most of the precedents, until recently, are found in the 19th century. In 1984 and again just before the 1987 market crash, real bond yields rose materially above the estimated real return on stocks. How well did this development predict subsequent relative returns? Stated more provocatively, why didn't our model work? Why didn't bonds beat stocks in the past decade? After all, with the 1984 peak in real bond returns and again shortly before the 1987 crash, the risk premium dipped even lower than the levels seen at the market peak in early 2000. Yet, stocks subsequently outpaced bonds. For an answer, recall that the context was a more than doubling of stock valuations, whether measured in price-to-book ratios, price-to-dividend ratios, or P/E multiples. If valuation multiples had held constant, the bonds would have prevailed.[35]

Appendix A. Estimating the Constituents of Return

An analysis of historical data is only as good as the data themselves. Accordingly, we availed ourselves of multiple data sources whenever possible. We were encouraged by the fact that the discrepancies between the various sources led to compounded rates of return that were no more than 0.2 percent different from one another.

Long Government Bond Yields, *BY*(*t*). Our data sources are as follows: for January 1800 to May 2001, 10-year government bond yields from Global Financial Data of the National Bureau of Economic Research (NBER) (data were annual until 1843 and were interpolated for monthly estimates); for June 2001 to December 2001, Bloomberg; and for January 1926 to December 2000, Ibbotson Associates, long-term government bond yields and returns. In cases of differences, we (1) averaged the yield data and (2) recomputed monthly total returns based on an assumed 10-year maturity standard.

Inflation, *INF*(*t*). We used two sources of inflation and U.S. Consumer Price Index data. For January 1801 to May 2001, NBER (annual until 1950; interpolated for monthly estimates); for June 2001 to December 2001, Bloomberg; and for January 1926 to December 2000, Ibbotson Associates. In cases of differences, we averaged the available data. Ibbotson data were given primary (two-thirds) weighting for 1926–1950 because the NBER data are annual through 1950.

Gross Domestic Product, *GDP*(*t*). For January 1800 to September 2001, NBER GNP data annually through 1920, interpolated July-to-July; for 1921–2001, quarterly GDP data; and for December 2001, *Wall Street Journal* consensus estimates.

Dividend Yield in Month *t*, *DY*(*t*), and Return on Stocks in Month *t*, *RS*(*t*). For January 1802 to December 1925, G. William Schwert (1990); for February 1871 to March 2001, Robert Shiller (2000); for January 1926 to December 2000, Ibbotson Associates (2001); and for April 2001 to December 2001, Bloomberg. In cases of differences, we averaged the available data. In Shiller's data, monthly dividend and earnings data are computed from the S&P four-quarter data for the quarter since 1926, with linear interpolation to monthly figures. Dividend and earnings data before 1926 are from Cowles (1939), interpolated from annual data.

Notes

1. The "bible" for the return assumptions that drive our industry is the work of Ibbotson Associates, building on the pioneering work of Ibbotson and Sinquefield (1976a, 1976b). The most recent update of the annual Ibbotson Associates data (2001) shows returns for U.S. stocks, bonds, bills, and inflation of, respectively, 11.0 percent, 5.3 percent, 3.8 percent, and 3.1 percent. These figures imply a real return for stocks of 7.9 percent and a risk premium over bonds of 5.7 percent (570 bps), both measured over a 75-year span. These data shape the expectations of the actuarial community, much of the consulting community, and many fund sponsors.
2. Fischer Black was fond of pointing out that examining the same history again and again with one new year added each passing year is an insidious form of data mining (see, for example, Black 1976). The past looks best when nonrecurring developments and valuation-level changes have distorted the results; extrapolating the past tacitly implies a belief that these nonrecurring developments can recur and that the changes in valuation levels will continue.
3. We strongly suggest that the investment community draw a distinction between past excess returns (observed returns from the past) and expected risk premiums (expected return differences in the future) to avoid continued confusion and to reduce the dangerous temptation to merely extrapolate past excess returns in shaping expectations for the risk premium. This habit is an important source of confusion that, quite literally, (mis)shapes decisions about the management of trillions in assets worldwide. We propose that the investment community begin applying the label "risk premium" *only* to expected future return differences and apply the label "excess returns" to observed historical return differences.
4. To see the effect of compounding at this rate, consider that if our ancestors could have earned a mere 1.6 percent real return on a $1 investment from the birth of Christ in roughly 4 B.C. to today, we would today have enough to buy more than the entire world economy. Similarly, the island of Manhattan was ostensibly purchased for $24 of goods, approximately the same as an ounce of gold when the dollar was first issued. This modest sum invested to earn a mere 5 percent real return would have grown to more than $20 billion in the 370 years since the transaction. At an 8 percent real return, as stocks earned from 1926 to 2000 in the Ibbotson data, this $24 investment would now suffice to buy more than the entire world economy.

5. No rational investor buys if he or she expects less than 1 percent real growth a year in capital, but objective analysis will demonstrate that this return is what stocks have actually delivered, plus their dividend yield, plus or minus any profits or losses from changes in yields. As Asness pointed out in "Bubble Logic" (2000), few buyers of Cisco would have *expected* a 1 percent internal rate of return at the peak, although the stock was priced to deliver just that, even if the overly optimistic consensus earnings and growth forecasts at the time were used. These buyers were focused on the view that the stock would produce handsome gains, as it had in the past, rather than on pursuing an objective evaluation, by using IRR or similar objective valuation tools, of expected returns. Such a focus plants the seeds of major disappointment.
6. The Welch study investigated an expected *arithmetic* risk premium for stocks relative to *cash*, not bonds. The difference between arithmetic and geometric returns is often illustrated by someone earning 50 percent in one year and –50 percent in the next. The arithmetic average is zero, but the person is down 25 percent (or 13.4 percent a year). Most practitioners think in terms of compounded geometric returns; in this example, practitioners would focus on the 13 percent a year loss, not on the zero arithmetic mean. If stocks have 16 percent average annual volatility (the average since World War II), the result is that the arithmetic mean is 130 bps higher than the geometric mean return (the difference is approximately half the variance, or 16 percent × 16 percent/2). Such a difference might be considered a "penalty for risk." If we add a 70 bp real cash yield (the historical average) plus a 720 bp risk premium minus a 130 bp penalty for risk, we find 6.6 percent to be the implied consensus of the economists for the geometric real stock return.
7. Such a return could easily fall to 0–2 percent net of taxes, especially in light of government's taxes on the inflation component of returns.
8. Smith's work even won a favorable review from John Maynard Keynes (for Keynes' approach, see his 1936 classic).
9. TIPS is the acronym for Treasury Inflation-Protected Securities, which have been replaced by Treasury Inflation-Indexed Securities.
10. In fairness, growth is now an explicit part of the picture. Dividend payout ratios are substantially lower than in the early 1920s and the 19th century as a result, at least in part, of corporate desires to finance growth. That said, our own evidence would suggest that internal reinvestment is not necessarily successful: High payout ratios precede higher growth than do low payout ratios.
11. We are indebted to G. William Schwert and Jeremy Siegel for some of the raw data for this analysis (see also Schwert 1990 and Siegel 1998). Although multiple sources exist for data after 1926 and a handful of sources provide data beginning in 1855 or 1870, Professor Schwert was very helpful in assembling these difficult early data. Professor Siegel provided earnings data back to 1870. We have not found a source for earnings data before 1870.
12. The U.S. Bureau of Labor Statistics maintains GDP data from 1921 to date; the earlier data are for GNP (gross national product). Because the two were essentially the same thing until international commerce became the substantial share of the economy that it is today, we used the GNP data from the Bureau of Labor Statistics for the 19th century and the first 20 years of the 20th century.
13. We stripped out reinvestment in the measure of real dividend growth shown in Figure 3 because investors are already receiving the dividend. To include dividends in the real dividend growth would double-count these dividends. What should be of interest to us is the internal growth in dividends stemming from reinvestment of the retained earnings.
14. We multiplied the real dividends by 10 to bring the line visually closer to the others; the result is that on those few occasions when the price line and dividend line touch, the dividend yield is 10 percent.
15. The fact that growth in real dividends and earnings is closer to per capita GDP growth than it is to overall GDP growth is intuitively appealing on one fundamental basis: Real per capita GDP growth measures the growth in productivity. It is sensible to expect real income, real per share earnings, and real per share dividends to grow with productivity rather than to mirror overall GDP growth.
16. This history holds a cautionary tale with regard to today's stock option practices.
17. This fall in dividends of existing enterprises is not surprising when one considers that the companies that existed in 1802 probably encompass, at most, 1 percent of the economy of 2001. The world has so changed that, at least from the perspective of the dominant stocks, today's economy would be unrecognizable in 1802.
18. Another way to think about this idea is to recognize the distinction between a market portfolio and a market index. The market portfolio shows earnings and dividend growth that are wholly consistent with growth in the overall economy (Bernstein 2001a). But if one were to unitize that market portfolio, the unit values would not grow as fast as the total capitalization and the earnings and dividends per unit (per "share" of the index) would not keep pace with the growth in the aggregate dollar earnings and dividends of the companies that compose the market portfolio. (When one stock is dropped and another added to a market index, typically the added stock is larger in capitalization than the deletion, which increases the divisor for constructing the index.) Precisely the same thing would happen in the management of an actual index fund. When a stock was replaced, the proceeds from the deleted stock would rarely suffice to fund the purchase of the added stock. So, all stocks would be trimmed slightly to fund that purchase; this consequence is implied by the change in the divisor for an index. It is this mechanism that drives the difference between the growth of the aggregate dollar earnings and dividends for the market portfolio, which will keep pace with GDP growth over time, and the growth of the "per share" earnings and dividends for the market *index* that creates the dilution we attribute to entrepreneurial capitalism. After all, entrepreneurial capitalism creates the companies that we must add to the market portfolio, thus changing our divisor and driving a wedge between the growth in market earnings or dividends and the growth in earnings and dividends per share in a market index.
19. See Bernstein (2001b). Over the past 131 years, the correlation between payout ratios and subsequent 10-year growth in real earnings has been 0.39; over the past 50 years, this correlation has soared to 0.66. Apparently, the larger the fraction of earnings paid out as dividends, the faster earnings subsequently grow, which is directly contrary to the Miller–Modigliani maxim (see Miller and Modigliani 1961 and Modigliani and Miller 1958).
20. To produce a 3.4 percent real return from stocks, matching the yield on TIPS, real growth in dividends needs to be 1.9 percent (twice the long-term historical real growth rate) while valuation levels remain where they are. Less than twice the historical growth in real dividends, or a return to the 3–6 percent yields of the past, will not get us there.
21. We have made the simplifying assumption that "long term" is a 10-year horizon. Redefining the long-term returns over a 5-year or 20-year horizon produces similar results.
22. Because this adjusted dividend is always at or above the true dividend, we have introduced a positive error into the average dividend yield. We offset this error by subtracting the 40-year average difference between the adjusted dividend and the true dividend. In this way, $EDY(t)$ is not overstated, on average, over time.

23. Of course, stock buybacks increase the share of the economy held by existing shareholders.
24. Arnott and Asness (2002) have shown that since 1945, the payout ratio has had a 77 percent correlation with subsequent real earnings growth. That is, higher retained earnings have historically led to slower, not faster, earnings growth.
25. Throughout this article, when we refer to a 10-year average or a 40-year average, we have used the available data if fewer years of data were available. For instance, for 1820, we used the 20-year GDP growth rate because 40 years of data were unavailable. We followed a convention of requiring at least 25 percent of the intended data; so, if the analysis was based on a 40-year average, we tolerated a 10-year average if necessary. To do otherwise would have forced us to begin our analysis in about 1840 and lose decades of interesting results. Because data before 1800 are very shaky and we required at least 10 years of data, our analysis begins, for the most part, in 1810.
26. We cannot know the 10-year returns from starting dates after 1991, so 192 years of expected return data lead to 182 years of correlation with subsequent 10-year actual returns.
27. Another way to deal with serially correlated data is to test correlations of differenced data. When we carried out such tests, we found that over the full span, the R^2 actually *rose* to 0.446 from the 0.214 shown in Panel A of Table 1; moreover, since 1945, the differenced results showed a still impressive 46 percent correlation. These results are available from the authors on request.
28. In an *ex ante* regression, the model is respecified for each monthly forecast with the use of all previously available data only.
29. We made the simplifying assumption that "long term" is a 10-year horizon. Redefining the long-term returns over a 5-year or 20-year horizon produced similar results.
30. Even when we considered successive differences to eliminate the huge serial correlation of real bond yields and 10-year real bond returns, the result from 1945 to date (available from the authors) was identical to the result for the raw data—a correlation of 0.63.
31. For investors accustomed to the notion that stock returns are uncertain and bond returns are assured over the life of the bond, this result will come as a surprise. But conventional bonds do *not* assure real returns; their expected real returns, therefore, should be highly uncertain. Stocks do, in a fashion, pass inflation through to the shareholder. So, nominal returns for stocks may be volatile and uncertain, but expected real stock returns are much more tightly defined than expected real bond returns.
32. Differencing caused the correlation for the full 182-year span to fall from 0.66 to 0.61 and, for the span following World War II, caused it to fall from 0.79 to 0.48.
33. For the taxable investor, the picture is worse, of course. In the United States, investors are even taxed on the inflation component of returns. From valuation levels that are well above historical norms, a negative real after-tax return is not at all improbable.
34. The excess return of stocks over bonds was negative also in the decades ended September 1991, November 1990, most 10-year spans ending August 1977 to June 1979, and the spans ending September 1974 to January 1975.
35. Consider the 10 years starting just before the stock market crash in September 1987. This span began with double-digit bond yields. The bond yield of 9.8 percent minus a regression-based inflation expectation of 3.6 percent led to an expected real bond return of 6.2 percent. The stock yield of 2.9 percent plus expected real per capita GDP growth of 1.6 percent minus an expected dividend shortfall relative to per capita GDP of 0.4 percent led to an expected real stock return of 4.0 percent. The risk premium was –2.0 percent. But stocks beat bonds by 4.9 percent a year over the next 10 years ending September 1997. What happened? The dividend yield plunged to 1.7 percent. This plunge in yields contributed 5.8 percent a year to stock returns; in the absence of this revaluation, stocks would have underperformed bonds by –0.9 percent. So, the –2.0 percent forecast was not bad; dividends rose a notch faster than normal, and more importantly, the price that the market was willing to pay for each dollar of dividends nearly doubled.

References

Arnott, Robert D., and Clifford S. Asness. 2002. "Does Dividend Policy Foretell Earnings Growth?" Unpublished working paper (January).

Arnott, Robert D., and James N. von Germeten. 1983. "Systematic Asset Allocation." *Financial Analysts Journal*, vol. 39, no. 6 (November/December):31–38.

Arnott, Robert D., and Ronald Ryan. 2001. "The Death of the Risk Premium: Consequences of the 1990s." *Journal of Portfolio Management*, vol. 27, no. 3 (Spring):61–74.

Asness, Cliff. 2000. "Bubble Logic or How to Learn to Stop Worrying and Love the Bull." AQR working paper (June).

Bernstein, Peter L. 2001a. "Ending Endism." *Economics and Portfolio Strategy* (April 30).

———. 2001b. "The Real Economy and the Unreal Economy." *Economics and Portfolio Strategy* (May 15).

Black, Fisher. 1976. "Studies of Stock Price Volatility Changes." *Proceedings of the 1976 Meetings of the American Statistical Society, Business and Economics Statistics Section*:177–181.

Campbell, John Y., and Robert J. Shiller. 1998. "Valuation Ratios and the Long-Run Stock Market Outlook." *Journal of Portfolio Management*, vol. 24, no. 2 (Winter):11–26.

Chancellor, Edward. 1999. *Devil Take the Hindmost*. New York: Farrar, Straus, and Giroux.

Claus, James, and Jacob Thomas. 2001. "Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stocks." *Journal of Finance*, vol. 56, no. 5 (October):1629–66.

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

Cowles, Alfred. 1939. *Common Stock Indexes: 1871–1937*. Bloomington, IN: Principia.

Fama, Eugene F., and Kenneth R. French. 2000. *The Equity Risk Premium*. Working Paper No. 522, Center for Research in Security Prices (April).

Fisher, Irving. 1930. *The Theory of Interest*. New York: MacMillan.

Hicks, Sir John Richard. 1946. *Value and Capital: An Inquiry into Some Fundamental Principles of Economic Theory*. Oxford, UK: Oxford University Press.

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

Ibbotson Associates. 2001. *Stocks, Bonds, Bills, and Inflation: 2000 Yearbook.* Chicago, IL: Ibbotson Associates.

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.

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money.* New York: Harcourt Brace and World.

Macaulay, Frederick R. 1938. *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States since 1856.* Washington, DC: National Bureau of Economic Research.

Miller, Merton, and Franco Modigliani. 1961. "Dividend Policy, Growth, and the Valuation of Shares." *Journal of Business,* vol. 34, no. 4 (October):411–433.

Modigliani, Franco, and Merton H. Miller. 1958. "The Cost of Capital, Corporation Finance, and the Theory of Investment." *American Economic Review,* vol. 48, no. 3 (June):261–297.

Schwert, G. William. 1990. "Indexes of United States Stock Prices from 1802 to 1987." *Journal of Business,* vol. 63, no. 3 (July):399–426.

Shiller, Robert J. 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review,* vol. 73, no. 3 (June):421–436.

———. *Irrational Exuberance.* 2000. Princeton, NJ: Princeton University Press.

Siegel, Jeremy J. 1998. *Stocks for the Long Run.* 2nd ed. New York: McGraw-Hill.

Smith, Edgar Lawrence. 1924. *Common Stocks as Long-Term Investments.* New York: MacMillan.

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

The Shrinking Equity Premium

Historical facts and future forecasts.

Jeremy J. Siegel

Few conundrums have caught the imagination of economists and practitioners as much as the "Equity Premium Puzzle," the title chosen by Rajneesh Mehra and Edward Prescott for their seminal 1985 article in the *Journal of Monetary Economics*. Mehra and Prescott show that the historical return on stocks has been too high in relation to the return on risk-free assets to be explained by the standard economic models of risk and return without invoking unreasonably high levels of risk aversion.[1] They calculate the margin by which stocks outperformed safe assets — the *equity premium* — to be in excess of 6 percentage points per year, and claim that the profession is at a loss to explain its magnitude.

There have been many attempts since to explain the size of the equity premium by variations of the standard finance model. I shall not enumerate them here, but refer readers to reviews by Abel [1991], Kocherlakota [1996], Cochrane [1997], and Siegel and Thaler [1997].

I review here the estimates of the equity premium derived from historical data, and offer some reasons why I believe that most of the historical data underestimate the real return on fixed-income assets and overestimate the expected return on equities. I shall also offer some reasons why, given the current high level of the stock market relative to corporate earnings, the forward-looking equity premium may be considerably lower than the historical average.

JEREMY J. SIEGEL is the Russell E. Palmer professor of finance at the Wharton School of the University of Pennsylvania in Philadelphia (PA 19104).

REAL RETURNS ON "RISK-FREE" ASSETS

From 1889 through 1978, Mehra and Prescott estimate the real return on short-dated fixed-income assets (commercial paper until 1920 and Treasury bills thereafter) to have been 0.8%. In 1976 and again in 1982, Roger Ibbotson and Rex Sinquefield formally estimated the real risk-free rate to be even lower — at zero, based on historical data analyzed from 1926. This extremely low level of the short-term real rate is by itself puzzling, and has been termed the "real rate puzzle" by Weil [1989]. The essence of this puzzle is that, given the historical growth of per capita income, it is surprising that the demand to borrow against tomorrow's higher consumption has not resulted in higher borrowing rates.

The low measured level of the risk-free rate may in fact be in part an artifact of the time period examined. There is abundant evidence that the real rate both during the nineteenth century and after 1982 has been substantially higher. Exhibit 1, based on Siegel [1998], indicates that over the entire period from 1802 through 1998, the real compound annual return on Treasury bills (or equivalent safe assets) has been 2.9%, while the realized return on long-term government bonds has been 3.5%. Exhibit 2 presents the historical equity premium for selected time periods for both bonds and bills based on the same data.[2]

EXHIBIT 1
COMPOUND ANNUAL REAL RETURNS (%)
U.S. DATA, 1802-1998

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

Source: Siegel [1998] updated.

The danger of using historical averages — even over long periods — to make forecasts is readily illustrated by noting Ibbotson and Sinquefield's long-term predictions made in 1976 and again in 1982 on the basis of their own analysis of the historical data. In 1976, they made predictions for the twenty-five-year period from

EXHIBIT 2
EQUITY PREMIUMS (%) — U.S. DATA, 1802-1998

	Equity Premium with Bonds		Equity Premium with Bills	
	Geometric	Arithmetic	Geometric	Arithmetic
1802-1998	3.5	4.7	5.1	5.5
1802-1870	2.2	3.2	1.9	2.9
1871-1925	2.9	4.0	3.4	4.6
1926-1998	5.2	6.7	6.7	8.6
1946-1998	6.5	7.3	7.2	8.6

Source: Siegel [1998] updated.

1976 through 2000, and in 1982 they made predictions for the twenty-year period from 1982 through 2001. Their forecasts are shown in Exhibit 3. Since we now have data for most of these forecast periods, it is of interest to assess their estimates.

The last two decades have been extremely good for financial assets, so it is not surprising that Ibbotson and Sinquefield underestimate all their real returns. But their most serious underestimation is for fixed-income assets, where they forecast the real bill rate to average essentially zero and the real return on bonds to be less than 2%. Given the standard deviation of estimates, realized annual real bond and bill returns have been 9.9% and 2.9%, respectively, significantly above their estimates. Since negative real returns on fixed-income assets persisted between the two surveys, Ibbotson and Sinquefield more seriously underestimate long-term real bill rates in their 1982 forecasts than they did in 1976.[3]

My purpose here is not to highlight errors in Ibbotson's and Sinquefield's past forecasts. Their analysis was state-of-the-art, and their data have rightly formed the benchmark for the risk and return estimates used by both professional and academic economists. I bring these forecasts to light to show that even the fifty-year history of financial returns available to economists at that time was insufficient to estimate future real fixed-income returns.

It is not well understood why the real rate of returns on fixed-income assets was so low during the 1926-1980 period. The bursts of unanticipated inflation following the end of World War II and during the 1970s certainly had a negative effect on the realized real returns from long-term bonds. Perhaps the shift from a gold standard to a paper monetary standard had a negative effect on these real returns until investors fully adjusted to the inflationary bias inherent in the new monetary standard.[4]

Whatever the reasons, the current yields on the Treasury inflation-protected securities, or TIPS, first issued in 1997 support the assertion that the future real returns on risk-free assets will be substantially above the level estimated over the Ibbotson-Sinquefield period. This is so even when the estimating period includes the higher real rates of the past two decades. In August 1999, the ten- and thirty-year TIPS bond yielded 4.0%, nearly twice the realized rate of return on long-dated government bonds over the past seventy-five years.[5]

The market projects real returns on risk-free assets to be substantially higher in the future than they have been over most of this century. It is also likely that the expected returns in the past are substantially greater than they have turned out ex post, especially for longer-dated securities. If one uses a 3.5% real return on fixed-income assets, the geometric equity premium for a 7.0% real stock return falls to 3.5%.

EXHIBIT 3
LONG-TERM FORECASTS OF REAL RETURNS — COMPOUND ANNUAL RATES OF RETURN

Forecast Period		Stocks	Bonds	Bills	Inflation
1976-2000	Forecast	6.3 (23.5)	1.5 (8.0)	0.4 (4.6)	6.4 (4.8)
	Actual*	11.0	5.3	2.1	4.8
1982-2001	Forecast	7.6 (21.9)	1.8 (8.3)	0.0 (4.4)	12.8 (5.1)
	Actual*	14.6	9.9	2.9	3.3

*Data through 1998.
Standard deviations of annual returns in parentheses.
Source: Ibbotson and Sinquefield [1976, 1982].

HISTORICAL EQUITY RETURNS AND SURVIVORSHIP BIAS

The real return on stocks, as I have emphasized [1998], has displayed a remarkable long-term stability. Over the entire 196-year period that I examine, the long-term after-inflation geometric annual rate of return on equity averages 7.0%. In the 1926-1998 period, the real return has been 7.4%, and since 1946 (when virtually all the thirteenfold increase in the consumer price index over the past two hundred years has taken place) the real return on equity has been 7.8%. The relative stability of long-term real equity returns is in marked contrast to the unstable real returns on fixed-income assets.

Some economists believe the 7% historical real return on equities very likely overstates the true expected return on stocks. They claim that using the ex post equity returns in the United States to represent returns expected by shareholders is misleading. This is because no investor in the nineteenth or early twentieth century could know for certain that the United States would be the most successful capitalist country in history and experience the highest equity returns.

This "survivorship bias" hypothesis, as it has been called, is examined by Jorion and Goetzmann [1999] in "Global Stock Markets in the Twentieth Century." They conclude that of thirty-nine equity markets that existed in 1921, none of them show as high a real capital appreciation as the United States, and most of them have had substantial disruptions in their operations or have disappeared altogether. They report that the median real capital appreciation of non-U.S. markets has been only 0.8% per year as opposed to 4.3% in the U.S.[6]

But this evidence may be misleading. Total returns of a portfolio, especially over long periods of time, are a very non-linear function of the returns of the individual components. Mathematically it can be shown that if individual stock returns are lognormal, the performance of the *median* stock is almost always worse than the market portfolio performance.[7]

So, it is not surprising that the median performance of individual countries will not match the "world portfolio" or the returns in the dominant market. Jorion and Goetzmann recognize this near the end of their study when they show that compound annual real return on a GDP-weighted portfolio of equities in all countries falls only 28 basis points short of the U.S. return. In fact, because of the real depreciation of the dollar over this time, the compound annual *dollar* return on a GDP-weighted world is actually 30 basis points *higher* than the return on U.S. equities.[8]

But examining international stock returns alone does not give us a better measure of the equity premium. The equity premium measures the *difference* between the returns on stocks and safe bonds. Although stock returns may be lower in foreign countries than the U.S., the real returns on foreign bonds are substantially lower. Almost all disrupted markets experienced severe inflation, in some instances wiping out the value of fixed-income assets. (One could say that the equity premium in Germany covering any period including the 1922-1923 hyperinflation is over 100%, since the real value of fixed-income assets fell to zero while equities did not.)

Even investors who purchased bonds that promised precious metals or foreign currency experienced significant defaults. It is my belief that if one uses a world portfolio of stocks and bonds, the equity *premium* will turn out higher, not lower, than found in the U.S.[9]

TRANSACTION COSTS AND DIVERSIFICATION

I believe that 7.0% per year does approximate the long-term real return on equity indexes. But the return on equity *indexes* does not necessarily represent the *realized* return to the equityholder. There are two reasons for this: transaction costs and the lack of diversification.[10]

Mutual funds and, more recently, low-cost "index funds" were not available to investors of the nineteenth or early twentieth century. Prior to 1975, brokerage commissions on buying and selling individual stocks were fixed by the New York Stock Exchange, and were substantially higher than today. This made the accumulation and maintenance of a fully diversified portfolio of stocks quite costly.

The advent of mutual funds has substantially lowered the cost of maintaining a diversified portfolio. And the cost of investing in mutual funds has declined over the last several decades. Rea and Reid [1998] report a decline of 76 basis points (from 225 to 149) in the average annual fee for equity mutual funds from 1980 to 1997 (see also Bogle [1999, p. 69]). Index funds with a cost of less than 20 basis points per year are now available to small investors.

Furthermore, the risk experienced by investors unable to fully diversify their portfolios made the risk-return trade-off less desirable than that calculated from stock indexes. On a risk-adjusted basis, a less-than-fully diversified portfolio has a lower expected return than the total market.

Given transaction costs and inadequate diversification, I assume that equity investors experienced real returns more in the neighborhood of 5% to 6% over most of the nineteenth and twentieth century rather than the 7% calculated from indexes. Assuming a 3.5% real return on bonds, the historical equity premium may be more like 1.5 to 2.5 percentage points, rather than the 6.0 percentage points recorded by Mehra and Prescott.

PROJECTING FUTURE EQUITY RETURNS

Future stock returns should not be viewed independently of current fundamentals, since the price of stocks is the present discounted value of all expected future cash flows. Earnings are the source of these cash flows, and the average price-to-earnings (P-E) ratio in the U.S. from 1871 through 1998 is 14 (see Shiller [1989] for an excellent source for this series).

Using data from August 13, 1999, the S&P 500 stock index is 1327, and the mean 1999 estimate for operating earnings of the S&P 500 stock index of fifteen analysts polled by Bloomberg News is $48.47.[11] This yields a current P-E ratio on the market of 27.4. But due to the increased number of write-offs and other special charges taken by management over the last several years, operating earnings have exceeded total earnings by 10% to 15%.[12] On the basis of reported earnings, which is what most historical series report (including Shiller's), the P-E ratio of the market is currently about 32.[13]

There are two long-term consequences of the high level of stock prices relative to fundamentals. Either 1) future stock returns are going to be lower than historical averages, or 2) earnings (and hence other fundamentals such as dividends or book value) are going to rise at a more rapid rate in the future. A third possibility, that P-E ratios will rise continually without bound, is ruled out since this would cause an unstable bubble in stock prices that must burst.

If future dividends grow no faster than they have in the past, forward-looking real stock returns will be lower than the 7% historical average. As is well known from the dividend discount model, the rate of return on stocks can be calculated by adding the current dividend yield to the expected rate of growth of future dividends. The current dividend yield on the S&P 500 index is 1.2%. Since 1871, the growth of real per share dividends on the index has been 1.3%, but since 1946, due in part to a higher reinvestment rate, growth has risen to 2.1%. If we assume future growth of real per share dividends to be close to the most recent average of 2.1%, we obtain a 3.3% real return on equities, less than one-half the historical average.

A second method of calculating future real returns yields a similar figure. If the rate of return on capital equals the return investors require on stocks, the *earnings yield*, or the reciprocal of the price-earnings ratio, equals the forward-looking real long-term return on equity (see Phillips [1999] for a more formal development of this proposition). Long-term data support this contention; a 14 price-to-earnings ratio corresponds to a 7.1% earnings yield, which approximates the long-term real return on equities. The current P-E ratio on the S&P 500 stock index is between 27 to 32, depending on whether total or operating earnings are considered. This indicates a current earnings yield, and hence a future long-term and real return, of between 3.1% to 3.7% on equities.

One way to explain these projected lower future equity returns is that investors are bidding up the price of stocks to higher levels as the favorable historical data about the risks and returns in the equity market become incorporated into investor decisions.[14] Lower transaction costs further enable investors to assemble diversified portfolios of stocks to take advantage of these returns. The desirability of stocks may be further reinforced by the perception that the business cycle has become less severe over time and has reduced the inherent risk in equities.[15]

If these factors are the cause of the current bull market, then the revaluation of equity prices is a one-time adjustment. This means that future expected equity returns should be lower, not higher, than in the past. During this period of upward price adjustment, however, equity returns will be higher than average, increasing the historical measured returns in the equity market.

This divergence between increased historical returns and lower future returns could set the stage for some significant investor disappointment, as survey evidence suggests that many investors expect future returns to be higher, not lower, than in the past (see "PaineWebber Index of Investor Optimism" [1999]).

SOURCES OF FASTER EARNINGS GROWTH

Although the increased recognition of the risks and returns to equity may be part of the explanation for the bull market in stocks, there must be other reasons. This is because the forward-looking rates of return we derive for equities fall below the current 4.0% yield on inflation-protected government bonds. Although one could debate whether in the long run stocks or *nominal* bonds are riskier in real terms, there should be no doubt that the inflation-protected bonds are safer than equities and should have a lower expected return.

Hence, some part of the current bull market in stocks must be due to the expectations that future earnings (and dividend) growth will be significantly above the historical average. Optimists frequently cite higher growth of real output and enhanced productivity, enabled by the technological and communications revolution, as the source of this higher growth. Yet the long-run relation between the growth of real output and *per share* earnings growth is quite weak on both theoretical and empirical grounds. Per share earnings growth has been primarily determined by the reinvestment rate of the firm, or the earnings yield minus the dividend yield, not the rate of output growth.[16]

The reason why output growth does not factor into per share earnings growth is that new shares must be issued (or debt floated) to cover the expansion of productive technology needed to increase output. Over the long run, the returns to technological progress have gone to workers in the form of higher real wages, while the return per unit of capital has remained essentially unchanged. Real output growth could spur growth in per share earnings only if it were "capital-enhancing," in the growth terminology, which is contrary to the labor-augmenting and wage-enhancing technological change that has marked the historical data (see Diamond [1999] for a discussion of growth and real return).

But there are factors that may contribute to higher future earnings growth of U.S. corporations, at least temporarily. The United States has emerged as the leader in the fastest-growing segments of the world economy: technology, communications, pharmaceuticals, and, most recently, the Internet and Internet technology. Furthermore, the penetration of U.S. brand names such as Coca-Cola, Procter & Gamble, Disney, Nike, and others into the global economy can lead to temporarily higher profit growth for U.S. firms.

Nonetheless, the level of corporate earnings would have to double to bring the P-E ratio down to the long-term average, or to increase by 50% to bring the P-E ratio down to 20. A 20 price-to-earnings yield corresponds to a 5% earnings yield or a 5% real return, a return that I believe approximates realized historical equity returns after transaction costs are subtracted. For per share earnings to temporarily grow to a level 50% above the long-term trend is clearly possible in a world economy where the U.S. plays a dominant role, but it is by no means certain.

CONCLUSION

The degree of the equity premium calculated from data estimated from 1926 is unlikely to persist in the future. The real return on fixed-income assets is likely to be significantly higher than that estimated on earlier data. This is confirmed by the yields available on Treasury inflation-linked securities, which currently exceed 4%. Furthermore, despite the acceleration in earnings growth, the return on equities is likely to fall from its historical level due to the very high level of equity prices relative to fundamentals.[17]

All of this makes it very surprising that Ivo Welch [1999] in a survey of over 200 academic economists finds that most estimate the equity premium at 5 to 6 percentage points over the next thirty years. Such a premium would require a 9% to 10% real return on stocks, given the current real yield on Treasury inflation-indexed securities. This means that real per share dividends would have to grow by nearly 8.0% to 9.0% per year, given the current 1.2% dividend yield, to prevent the P-E ratio from rising farther from its current record levels. This growth rate is more than six times the growth rate of real dividends since 1871 and more than triple their growth rate since the end of World War II.

Unless there is a substantial increase in the productivity of capital, dividend growth of this magnitude would mean an ever-increasing share of national income going to profits. This by itself might cause political ramifications that could be negative for shareholders.

ENDNOTES

This article is adapted from a paper delivered at the UCLA Conference, "The Equity Premium and Stock Market Valuations," and a Princeton Center for Economic Policy Studies Conference, "What's Up with the Stock Market?" both held in May 1999. The author thanks participants in these seminars and particularly Jay Ritter, Robert Shiller, and Peter L. Bernstein for their comments.

[1]A few economists believe these high levels of risk aversion are not unreasonable; see, e.g., Kandel and Stambaugh [1991].

[2]In the capital asset pricing model, equity risk premiums are derived from the *arithmetic* and not geometric returns. Compound annual geometric returns are almost universally used in characterizing long-term returns.

[3]Their wildly high 12.8% long-term inflation estimate in 1982 is derived by subtracting their low historical real yield from the high nominal bond rate. This overprediction has no effect on their estimated *real* returns.

[4]But real rates on *short-dated* bonds, for which unanticipated inflation should have been less important, were also extremely low between 1926 and 1980.

[5]I am very persuaded by the research of Campbell and Viceira [1998], who argue that in a multiperiod world the proper risk-free asset is an inflation-indexed annuity rather than the short-dated Treasury bill. This conclusion comes from intertemporal models where agents desire to hedge against unanticipated changes in the real rate of interest. The duration of such an indexed annuity is closely approximated by the ten-year inflation-indexed bonds.

[6]They are unable to construct dividend series for most foreign countries, but they make a not-unreasonable assumption that dividend yields in the U.S. were at least as high as abroad.

[7]Intuitively, the return of the winners more than compensates for the lower returns of the more numerous losers.

[8]Furthermore, the dollar return on the foreign portfolio is much better measured than the real return. These data are taken from Jorion and Goetzmann [1991], Tables VI and VII.

[9]To avoid the problems with default, gold is considered the "risk-free" alternative in many countries. But gold's long-term real returns are negative in the U.S. even before one considers storage and insurance costs. And precious metals are far from risk-free in real terms. The real return on gold since 1982 has been a negative 7% per year.

[10]I abstract from taxes, which reduce the return on both bonds and stocks.

[11]These data were taken from the Bloomberg terminal on August 16, 1999.

[12]From 1970 through 1989, operating earnings exceeded reported earnings by an average of 2.29%. Since 1990, the average has been 12.93%.

[13]There are other factors that distort reported earnings, some upward (underreporting option costs: see Murray, Smithers, and Emerson [1998]) and some downward (overexpensing R&D; see Nakamura [1999]). No clear bias is evident.

[14]This is particularly true on a long-term, after-inflation basis. See Siegel [1998, Chapter 2].

[15]Bernstein [1998] has emphasized the role of economic stability in stock valuation. Also see Zarnowitz [1999] and Romer [1999]. Other reasons given for the high price of equities rely on demographic factors, specifically the accumulations of "baby boomers." This should, however, reduce both stock and bond returns, yet we see real bond returns as high if not higher than historically.

[16]From 1871 to 1998, the growth of real per share earnings is only 1.7% per year, slightly less than obtained by subtracting the median dividend yield of 4.8% from the median earnings yield of 7.2%.

[17]This should not be construed as predicting that equity prices need fall significantly, or that the expected returns on equities are not higher, even at current levels, than those on fixed-income investments.

REFERENCES

Abel, Andrew B. "The Equity Premium Puzzle." *Business Review*, Federal Reserve Bank of Philadelphia, September/October 1991, pp. 1-14.

Bernstein, Peter L. "The Risk Premium: A Critical Look and Some Pleasant Conclusions." Peter L. Bernstein, Inc., April 1, 1998.

Bogle, John C. *Common Sense on Mutual Funds*. New York: John Wiley & Sons, 1999.

Campbell, John Y., and Luis M. Viceira. "Who Should Buy Long-Term Bonds?" NBER Working Paper No. 6801, November 1998.

Cochrane, John H. "Where is the Market Going? Uncertain Facts and Novel Theories." *Economic Perspectives*, Federal Reserve Bank of Chicago, XXI, 6 (November/December 1997), pp. 3-37.

Diamond, Peter A. "Projecting the Rate of Return on Equities." Working paper, Massachusetts Institute of Technology, June 24, 1999.

Ibbotson, Roger G., and Rex A. Sinquefield. "Stocks, Bonds, Bills, and Inflation: Year-by-Year Historical Returns (1926-74)." *Journal of Business*, 49 (January 1976), pp. 11-45.

——. "Stocks, Bonds, Bills, and Inflation: Simulations of the Future (1976-2000)." *Journal of Business*, 49 (July 1976), pp. 313-338.

——. *Stocks, Bonds, Bills and Inflation: The Past and the Future*. Charlottesville, VA: Financial Analysts Research Foundation, 1982, p. 137.

Jorion, Philippe, and William N. Goetzmann. "Global Stock Markets in the Twentieth Century." *Journal of Finance*, 54, 3 (June 1999), pp. 953-980.

Kandel, Shmuel, and Robert F. Stambaugh. "Asset Returns and Intertemporal Preferences." *Journal of Monetary Economics*, 27 (February 1991), pp. 39-71.

Kocherlakota, Narayana R. "The Equity Premium: It's Still a Puzzle." *Journal of Economics Literature*, 34 (March 1996), pp. 42-71.

Mehra, Rajneesh, and Edward C. Prescott. "The Equity Risk Premium: A Puzzle." *Journal of Monetary Economics*, 15 (March 1985), pp. 145-161.

Murray, Daniel, Andrew Smithers, and John Emerson. "USA: The Impact of Employee Stock Options." Report No. 117, Smithers and Co., Ltd., April 17, 1998.

Nakamura, Leonard. "Intangibles: What Put the New in the New Economy?" *Business Review*, Federal Reserve Bank of Philadelphia, July/August 1999, pp. 3-16.

"PaineWebber Index of Investor Optimism." Gallup Organization, Princeton, July 1999.

Philips, Thomas K. "Why Do Valuation Ratios Forecast Long-Run Equity Returns?" *Journal of Portfolio Management*, Spring 1999, pp. 39-44.

Rea, John D., and Brian K. Reid. "Trends in the Ownership Costs of Equity Mutual Funds." Investment Company Institute, November 1998, pp. 1-12.

Romer, Christina D. "Changes in Business Cycles: Evidence and Explanations." NBER Working Paper No. W6948, February 1999.

Shiller, Robert J. *Market Volatility*. Cambridge, MA: The MIT Press, 1989.

Siegel, Jeremy *Stocks for the Long Run*. New York: McGraw-Hill, 1998.

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

Weil, Philippe. "The Equity Premium Puzzle and the Risk-Free Rate Puzzle." *Journal of Monetary Economics*, 24 (November 1989), pp. 401-421.

Welch, Ivo. "Views of Financial Economists on the Equity Premium and Professional Controversies." Working paper, UCLA, 1999.

Zarnowitz, Victor. "Has the Business Cycle Been Abolished?" NBER Working Paper No. W6367, April 1999.

THE JOURNAL OF FINANCE • VOL. LVI, NO. 5 • OCT. 2001

Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets

JAMES CLAUS and JACOB THOMAS*

ABSTRACT

The returns earned by U.S. equities since 1926 exceed estimates derived from theory, from other periods and markets, and from surveys of institutional investors. Rather than examine historic experience, we estimate the equity premium from the discount rate that equates market valuations with prevailing expectations of future flows. The accounting flows we project are isomorphic to projected dividends but use more available information and narrow the range of reasonable growth rates. For each year between 1985 and 1998, we find that the equity premium is around three percent (or less) in the United States and five other markets.

THE EQUITY RISK PREMIUM LIES at the core of financial economics. Representing the excess of the expected return on the stock market over the risk-free rate, the equity premium is unobservable and has been estimated using different approaches and samples. The estimates most commonly cited in the academic literature are from Ibbotson Associates' annual reviews of the performance of various portfolios of U.S. stocks and bonds since 1926. Those estimates lie in the region of seven to nine percent per year, depending on the specific series examined. This historic evidence is objective and easy to interpret and has convinced many, especially academic financial economists, that the Ibbotson estimates are the best available proxies for the equity premium (Welch (1999)).[1] For discussion purposes, we use "eight percent"

* Barclays Global Investors and Columbia Business School, respectively. We thank I/B/E/S Inc. for their database of earnings estimates and Enrique Arzac and René Stulz for many helpful suggestions and discussions. Useful comments were received from anonymous referees, Bala Dharan, Darin Clay, Ilia Dichev, Ben Esty, Bob Hodrick, Irene Karamanou, S.P. Kothari, Jimmy Liew, Jing Liu, Jim McKeown, Karl Muller, Jim Ohlson, Stephen Penman, Huai Zhang, and workshop participants at AAA annual meetings (San Diego), Columbia University, Copenhagen Business School, University of Michigan, Michigan State University, University of North Carolina-Chapel Hill, Northern Arizona University, Ohio State University, Penn State University, Prudential Securities Quantitative Conference, Syracuse University, and University of Texas-Austin.

[1] The annualized distribution of monthly common stock returns over the 30-day T-bill rate has a mean of 9.12 percent and a standard deviation of 20.06 percent (from data in Table A-16, Ibbotson Associates (1999)). If these 73 observations are independent and identically distributed, the sample mean is a reasonable estimate for the equity premium, and the standard error of 2.35 percent associated with the sample mean allows an evaluation of other hypothesized values of the equity premium.

and "the Ibbotson estimate" interchangeably to represent the historic mean of excess returns earned by U.S. equities since 1926. (Unless noted otherwise, all amounts and rates are stated in nominal, not real, terms.)

Our objective is to show empirically that eight percent is too high an estimate for the equity premium in recent years. Rather than examine observed returns, we estimate for each year since 1985 the discount rate that equates U.S. stock market valuations with the present value of prevailing forecasts of future flows. Subtracting 10-year risk-free rates from these estimated discount rates suggests that the equity premium is only about three percent.[2] An examination of five other large stock markets (Canada, France, Germany, Japan, and the United Kingdom) provides similar results. Despite substantial variation in the underlying fundamentals across markets and over time, observing that every one of our 69 country-year estimates lies well below eight percent suggests that the Ibbotson estimate is too high for our sample period. Examination of various diagnostics (such as implied future profitability) confirms that the projections required to support an eight percent equity premium are unreasonable and inconsistent with past experience.

Some features of our study should be emphasized at the outset. As we only seek to establish a reasonable upper bound for the equity premium, we select long-term growth assumptions that exceed past experience and do not adjust for optimism in the analyst forecasts used.[3] Also, we use the simplest structure necessary to conduct our analysis. Our estimates refer to a long-term premium expected to hold over all future years (whereas historical estimates measure one-period premia), and we assume that the premium is constant over those future years (we do incorporate anticipated variation in risk-free rates). Finally, each annual estimate is conditional on the information available in that year; we do not consider an unconditional equity premium toward which those conditional premia might gravitate in the long run.

We are not the first to question the validity of the Ibbotson estimate. Mehra and Prescott (1985) initiated a body of theoretical work that has examined the so-called "equity premium puzzle." Their model indicates that the variance–covariance matrix of aggregate consumption and returns on stocks and bonds, when combined with reasonable risk-aversion parameters, implies equity premium estimates that are less than one percent. Despite subsequent efforts to bridge this gap (e.g., Abel (1999)), concerns remain about the validity of the Ibbotson estimate (see Kocherlakota (1996), Cochrane (1997), and Siegel and Thaler (1997) for summaries).

[2] Gebhardt, Lee, and Swaminathan (forthcoming) find similar results when estimating firm-specific discount rates, rather than the market-level discount rates considered in this paper.

[3] As described later, analyst optimism has declined systematically over time and a simple adjustment for mean bias is inappropriate. Bayesian adjustments to control for observed analyst optimism are not considered because we focus on an upper bound. In general, we do not use more complex econometric techniques and data refinements that are available to get sharper point estimates (e.g., Mayfield (1999), Vuolteenaho (1999), and Ang and Liu (2000)).

Surveys of institutional investors also suggest an equity premium substantially below eight percent (e.g., Burr (1998)), and there are indications that this belief has been held for many years (e.g., Benore (1983)).[4] Also, the weighted average cost of capital used in discounted cash flow valuations provided in analysts' research reports usually implies an equity premium below five percent. Current share prices appear systematically overpriced if an eight percent equity premium is used on reasonable projections of future flows. This overpricing is more evident when examining mature firms, where there is less potential for disagreement about growth opportunities.

To identify possible reasons why the Ibbotson estimate might overstate the equity premium in recent years, apply the Campbell (1991) decomposition of observed returns (in excess of the expected risk-free rate) for the market portfolio. The four components are: (1) the expected equity premium for that period; (2) news about the equity premium for future periods; (3) news about current and future period real dividend growth; and (4) news about the real risk-free rate for current and future periods. Here, news represents changes in expectations between the beginning and end of the current period (for current period dividend growth and risk-free rates, it represents the unexpected portion of observed values). Summing up both sides of this relation for each year since 1926 indicates that the average excess return observed would exceed the equity premium today if: (1) conditional one-year-ahead equity premia have declined; (2) the conditional long-term equity premium anticipated for future years has declined; (3) news about real dividend growth was positive on average; or (4) the expected real risk-free rate has declined.

The first and second reasons for why the Ibbotson estimate overstates the current equity premium highlight the potential pitfalls of estimating equity premia from observed returns. Holding aside news about dividends and risk-free rates, valuations would exceed expectations if the equity premium has declined (since present values increase when expected rates of return decline). That is, unexpected changes in the equity premium cause historical equity premium estimates to move in the opposite direction. Blanchard (1993) concludes that the equity premium has declined since 1926 to two or three percent by the early 1990s, and speculates that this decline is caused by a simultaneous decline in expected real rates of return on stocks and an increase in expected real risk-free rates. (This increase in expected real risk-free rates is another puzzle, but that puzzle is beyond the scope of this paper.) The remarkable run-up in stock prices during the 1990s, both domestically as well as internationally, is also consistent with a recent decline

[4] While many argue for an equity premium between two and three percent (e.g., Bogle (1999, p. 76)), some suggest that the premium is currently close to zero (e.g., Glassman and Hassett (1998), and Wien (1998)). Surveys of *individual* investors, on the other hand, suggest equity premia even higher than the Ibbotson estimate. For example, the *New York Times* (October 10, 1997, page 1, "High hopes of mutual fund investors"), reported an equity premium in excess of 16 percent from a telephone survey conducted by Montgomery Asset Management.

in the equity premium. Stulz (1999) argues that increased globalization has caused equity premia to decline in all markets.

Examination of historic evidence over other periods and markets suggests that the U.S. experience since 1926 is unusual. Siegel (1992) finds that the excess of observed annual returns for NYSE stocks over short-term government bonds is 0.6, 3.5, and 5.9 percent over the periods 1802 to 1870, 1871 to 1925, and 1926 to 1990, respectively. Jorion and Goetzmann (1999) examine the evidence for 39 equity markets going back to the 1920s, and conclude that the high equity premium observed in the United States appears to be the exception rather than the rule. Perhaps some stock markets collapsed and those markets that survived, like the U.S. exchanges, exhibit better performance than expected (see Brown, Goetzmann, and Ross (1995)). This evidence is consistent with the third reason for the high Ibbotson premium: since 1926, news about real dividend growth for U.S. stocks has been positive on average.

Partially in response to these limitations of inferring equity premia from observed returns, financial economists have considered forward-looking approaches based on projected dividends.[5] Informally, expected rates of return on the market equal the forward dividend yield plus expected growth in dividends (this dividend growth model is discussed in Section I). While dividend yields are easily measured, expected dividend growth in perpetuity is harder to identify. Proxies used for expected dividend growth include observed growth in earnings, dividends, or economy-wide aggregates (e.g., Fama and French (2000)). Unfortunately, the dividend growth rate that can be sustained in perpetuity is a hypothetical rate that is not necessarily anchored in any observable series, leaving considerable room for disagreement (see the Appendix for explanation).

We use a different forward-looking approach, labeled the abnormal earnings (or residual income) model, to mitigate problems associated with the dividend growth model.[6] Recognizing that dividends equal earnings less changes in accounting (or book) values of equity allows the stream of projected dividends to be replaced by the current book value of equity plus a function of future accounting earnings (details follow in Section I). While book values feature prominently in the model, the inclusion of future abnormal earnings makes it isomorphic to the dividend discount model. Relative to the dividend growth model, this approach makes better use of currently

[5] A related approach is to run predictive regressions of market returns or equity premium on dividend yields and other variables (e.g., Campbell and Shiller (1988)). We do not consider that approach because the declining dividend yields in recent years have caused predicted equity premium to turn negative (e.g., Welch (1999)).

[6] The approach appears to have been discovered independently by a number of economists and accountants over the years. Preinreich (1938) and Edwards and Bell (1961) are two early cites. More recently, a large body of analytical and empirical work has utilized this insight (e.g., Penman (1999)). Examples of empirical investigations include market myopia (Abarbanell and Bernard (1999)), explaining cross-sectional variation in returns (Liu and Thomas (2000)), and stock picking (Frankel and Lee (1998a, 1998b)).

available information to reduce the importance of assumed growth rates, and it narrows the range of allowable growth rates by focusing on growth in rents, rather than dividend growth.

If the equity premium is as low as our estimates suggest, required rates of return (used for capital budgeting, regulated industries, and investment decisions) based on the Ibbotson estimate are severely overstated. Second, a smaller equity premium reduces the importance of estimating beta accurately (because required rates of return become less sensitive to variation in beta) and increases the magnitude of beta changes required to explain abnormal returns observed for certain market anomalies. Finally, reducing substantially the magnitude of the equity premium puzzle to be explained might reinvigorate theory-based studies.

In Section I we develop the abnormal earnings approach used in this paper and compare it with the dividend growth model. Section II contains a description of the sample and methodology. The equity premium estimates for the United States are reported in Section III, and those for the five other markets are provided in Section IV. To confirm that our estimates are robust, we conducted extensive sensitivity analyses, which we believe represent an important contribution of our research effort. A summary of that investigation is reported in Section V (details are provided in Claus and Thomas (1999a)) and Section VI concludes.

I. Dividend Growth and Abnormal Earnings Models

The Gordon (1962) dividend growth model is described in equation (1). This relation implies that the expected rate of return on the stock market (k^*) equals the forward dividend yield (d_1/p_0) plus the dividend growth rate in perpetuity (g) expected for the market.

$$p_0 = \frac{d_1}{k^* - g} \Rightarrow k^* = \frac{d_1}{p_0} + g \tag{1}$$

where

p_0 = current price, at the end of year 0,
d_t = dividends expected at the end of future year t,
k^* = expected rate of return on the market, derived from the dividend growth model, and
g = expected dividend growth rate, in perpetuity.

The Gordon growth model is a special case of the general Williams (1938) dividend discount model, detailed in equation (2), where dividend growth is constrained to equal g each year.

$$p_0 = \frac{d_1}{(1+k^*)} + \frac{d_2}{(1+k^*)^2} + \frac{d_3}{(1+k^*)^3} + \dots \tag{2}$$

Research using the dividend growth model has often assumed that g equals forecasted earnings growth rates obtained from sell-side equity analysts, who provide earnings forecasts along with their buy/sell recommendations. These forecasts refer to earnings growth over the next "cycle," which is commonly interpreted to represent the next five years. Consequently, we refer to this earnings growth forecast as g_5. While most studies using g_5 as a proxy for g have focused on the U.S. market alone (e.g., Brigham, Shome, and Vinson (1985)), some have examined other major equity markets also (e.g., Khorana, Moyer, and Patel (1997)). Estimates of the equity premium based on the assumption that g equals g_5 are similar in magnitude to the Ibbotson estimate derived from historical data. For example, Moyer and Patel (1997) estimate the equity premium each year over their 11-year sample period (1985 to 1995) and generate a mean estimate of 9.38 (6.96) percent relative to the 1-year (30-year) risk-free rate.

However, others have balked at using g_5 as a proxy for g (e.g., Malkiel (1996), Cornell (1999)) because it appears unreasonably high at an intuitive level, and have stepped down assumed growth rates. Forecasted values of g_5 for the United States over our sample period, which are close to 12 percent in all years, exceed nominal growth in S&P earnings, which has been only 6.6 percent since the 1920s (*Wall Street Journal*, June 16, 1997, "As stocks trample price measures, analysts stretch to justify buying"). Also, the real growth rate implied by the nominal 12 percent earnings growth rate exceeds both forecast and realized growth in GDP (since 1970, forecasts of expected real growth in GDP have averaged 2.71 percent, and realized real growth has averaged 2.81 percent).

While we show that g_5 is systematically optimistic relative to realized earnings, it is difficult to infer reliably the level of that optimism from the relatively short time-series of forecast errors available (reliable data on analyst forecasts go back only about 15 years). Moreover, the incentives for analysts to make optimistic forecasts vary across firms and over time. For example, the literature on U.S. analysts' forecasts suggests that while analysts tended to make optimistic forecasts early in our sample period (to curry favor with management), more recently, management has tended to guide near-term analyst forecasts downward to be able to meet or beat them when announcing earnings.[7] Even if unbiased estimates of near-term earnings growth (g_5) were available, the Appendix describes why those estimates as well as observed growth rates are conceptually different from g, the hypothetical dividend growth that can be sustained in perpetuity.

[7] Results reported in Table VI offer clear evidence of such a decline in optimism for all horizons. Bagnoli, Beneish, and Watts (1999) document how recent analyst forecasts are systematically below reported earnings for their sample, and also below "whisper" forecasts that are generally viewed as representing the market's true earnings expectations. Matsumoto (1999) offers evidence in support of management guiding analyst forecasts downward, and also investigates factors that explain cross-sectional variation in this propensity to guide analysts.

The abnormal earnings model is an alternative that mitigates many of the problems noted above. Expected dividends can be related to forecasted earnings using equation (3) below, and that relation allows a conversion of the discounted dividends relation in equation (2) to the abnormal earnings relation in equation (4).

$$d_t = e_t - (bv_t - bv_{t-1}) \tag{3}$$

$$p_0 = bv_0 + \frac{ae_1}{(1+k)} + \frac{ae_2}{(1+k)^2} + \frac{ae_3}{(1+k)^3} + \dots, \tag{4}$$

where

e_t = earnings forecast for year t,
bv_t = expected book (or accounting) value of equity at the end of year t,
$ae_t = e_t - k(bv_{t-1})$ = expected abnormal earnings for year t, or forecast accounting earnings less a charge for the cost of equity, and
k = expected rate of return on the market portfolio, derived from the abnormal earnings model.

Equation (3), also known as the "clean surplus" relation, requires that all items affecting the book value of equity (other than transactions with shareholders, such as dividends and share repurchases/issues) be included in earnings. Under U.S. accounting rules, almost all transactions satisfy the clean-surplus assumption. An examination of the few transactions that do not satisfy this relation suggests that these violations occur ex post, and are not anticipated in analysts' earnings forecasts (e.g., Frankel and Lee (1998b)). Since we construct future book values using equation (3), by adding forecast income to and subtracting forecast dividends from beginning book values, clean surplus is maintained and the dividend and abnormal earnings relations in equations (2) and (4) are isomorphic.

Equation (4) shows that the current stock price equals the current book value of equity plus the present value of future expected abnormal earnings. Abnormal earnings, a proxy for economic profits or rents, adjusts reported earnings by deducting a charge for equity capital. Note that the market discount rates estimated from the abnormal earnings and dividend growth approaches are labeled differently: k and k^*. Also, the standard transversality conditions apply to both models: in the limit as t approaches infinity, the present value of future price, p_t (difference between price and book value, $p_t - bv_t$) must tend to zero in equation (2) (in equation (4)).

Financial economists have expressed concerns about accounting earnings deviating from "true" earnings (and book values of equity deviating from market values), in the sense that accounting numbers are noisy and easily manipulated. However, the equivalence between equations (2) and (4) is not impaired by differences between accounting and economic numbers, nor is it affected by the latitude available within accounting rules to report different accounting numbers. As long as forecasted earnings satisfy the clean surplus relation in equation (3) in terms of expectations, equation (4) is simply an algebraic restatement of equation (2), subject to the respective transversality conditions mentioned above.

Since the I/B/E/S database we use does not provide analysts' earnings forecasts beyond year +5, we assume that abnormal earnings grow at a constant rate (g_{ae}) after year +5, to incorporate dates past that horizon. Equation (4) is thus adapted as follows.

$$p_0 = bv_0 + \frac{ae_1}{(1+k)} + \frac{ae_2}{(1+k)^2} + \frac{ae_3}{(1+k)^3} + \frac{ae_4}{(1+k)^4} + \frac{ae_4}{(1+k)^5} + \left[\frac{ae_5(1+g_{ae})}{(k-g_{ae})(1+k)^5}\right]. \tag{5}$$

The last, bracketed term is a terminal value that captures the present value of abnormal earnings after year +5. The terms before are derived from accounting statements (bv_0) and analyst forecasts (e_1 to e_5). Note that there are three separate growth rates in this paper and the different growth rates refer to different streams and periods and arise from different sources. The rate g refers to dividend growth in perpetuity and is assumed by the researcher; g_5 refers to growth in accounting earnings over the first five years and is provided by financial analysts; and g_{ae} refers to abnormal earnings growth past year +5 and is assumed by the researcher.

Whereas expected rates of return are typically viewed as being stochastic (Samuelson (1965)), k^* and k in equations (1) and (5) are nonstochastic discount rates. Barring a few recent exceptions (e.g., Ang and Liu (2000) and Vuolteenaho (1999)), the literature has assumed that expected rates of return can be approximated by discount rates. We make that assumption too. While equation (1) is designed to only reflect a flat k^*, equation (2) can be restated to incorporate predictable variation over time in discount rates. Similarly, equation (5) can be restated to incorporate nonflat discount rates, as shown in Claus and Thomas (1999a). We consider the case when the equity premium is assumed to remain flat but discount rates vary over future periods based on the term-structure of risk-free rates. This restated version of equation (5) is

$$p_0 = bv_0 + \sum_{t=1}^{\infty}\left[\frac{ae_t}{\prod_{s=1}^{t}(1+r_{fs}+rp)}\right], \tag{5a}$$

where

r_{fs} = forward one-year risk-free rate for year s,
rp = equity risk premium, assumed constant over all future years,
ae_t = expected abnormal earnings for year t, equals $e_t - bv_{t-1}(r_{ft} + rp)$ for years +1 through +5, and equals $ae_5(1 + g_{ae})^{t-5}$, from year +6 on.

While the abnormal earnings stream in equation (4) is equivalent to the corresponding dividend stream in equation (2), the abnormal earnings relation in equation (5) (and equation (5a)) offers the following advantages over the dividend growth model in equation (1). First, a substantial fraction of the "value profile" for the abnormal earnings model in equation (5) is fixed by numbers that are currently available and do not need to be assumed by the researcher (current book value and abnormal earnings for years +1 through +5). Value profile is a representation of the fraction of total value captured by each future year's flows. In contrast, the entire value profile for the dividend growth model is affected by the assumed growth rate, g. Since the fraction of value determined by assumed growth rates is lower for the abnormal earnings approach, those risk premium estimates are more reliable.

Second, in contrast to the potential for disagreement about a reasonable range for g, the rate at which rents can grow in perpetuity after year +5, g_{ae}, is less abstract and easier to gauge using economic intuition. For example, to obtain equity premia around 8 percent, rents at the market level would have to grow forever at about 15 percent, on average. It is unlikely that aggregate rents to U.S. equity holders would grow at such high rates in perpetuity because of factors such as antitrust actions, global competition, and pressure from other stakeholders. The historical evidence (e.g., Myers (1999)) is also at odds with such high growth rates in abnormal earnings.

Third, future streams for a number of value-relevant indicators, such as price-to-book ratios (P/B), price-to-earnings ratios (P/E), and accounting return on equity (*roe*), can also be projected under the abnormal earnings approach. This allows one to paint a more complete picture of the future for different assumed growth rates. Analysis of the levels of future P/B and profitability (excess of *roe* over k) implied by growth rates required to obtain equity premium estimates around eight percent are also inconsistent with past experience.

II. Data and Methodology

I/B/E/S provides the consensus of all available individual forecasts as of the middle (the Thursday following the second Friday) of each month. Forecasts and prices should be gathered soon after the prior year-end, as soon as equity book values (bv_0) are available. Rather than collect forecasts at different points in the year, depending on the fiscal year-end of each firm, we opted to collect data as of the same month each year for all firms to ensure that the risk-free rate is the same across each annual sample. Since most firms have December year-ends, and book values of equity can be obtained from the balance sheets that are required to be filed with the SEC within 90 days of the fiscal year-end, we collect forecasts as of April each year.[8] For

[8] For the few firm-years not filing within this 90-day deadline, the book value of equity can be inferred by the market by adding (subtracting) fourth quarter earnings (dividends) from the third quarter book value of equity.

firms with fiscal year-ends other than December, this procedure creates a slight upward bias in estimated equity premium, since the stock prices used (as of April) are on average higher than those near the prior year's fiscal year-end, when bv_0 was released. In addition to earnings forecasts, I/B/E/S also provides data for actual earnings per share, dividends per share, share prices, and the number of outstanding shares. Equity book values are collected from COMPUSTAT's Industrial Annual, Research, and Full Coverage Annual Files, for years up to and including 1997.

The sample includes firms with I/B/E/S earnings forecasts for years +1 and +2 (e_1 and e_2) and a five-year growth forecast (g_5) as well as share prices and shares outstanding as of the I/B/E/S cut off date each April. We also require nonmissing data for the prior year's book value, earnings, and dividends. Explicit forecasts for years +3, +4, and +5 are often unavailable, and are generated by projecting the growth rate g_5 on the prior year's earnings forecast: $e_t = e_{t-1}(1 + g_5)$.[9]

Earlier years in the I/B/E/S database, before 1985, were dropped because they provided too few firms with complete data to represent the overall market. From 1985 on, the number of firms with available data increases substantially. As shown in column 1 of Table I, the number of sample firms increases from 1,559 in 1985 to 3,673 in 1998. Comparison with the total number of firms and market capitalization of all firms on NYSE, AMEX, and Nasdaq each April indicates that, although our sample represents only about 30 percent of all such firms, it represents 90 percent or more of the total market capitalization. Overall, we believe our sample is fairly representative of the value-weighted market, and refer to it as "the market" hereafter.

Firm-level data are aggregated each year to generate market-level earnings, book values, dividends, and capitalization. Actual data for year 0 (the full fiscal year preceding each April when forecasts were collected) is provided in columns 2 through 6 of Table I. Forecasted and projected earnings for years +1 through +5 are reported in columns 7 through 11.

Table I reveals an interesting finding relating to dividend payouts: the ratio of market dividends to earnings is around 50 percent in most years (with a noticeable decline toward the end of the sample period).[10] We use this 50 percent payout ratio to project future dividends from earnings fore-

[9] If any of the explicit earnings forecasts for years +2, +3, +4, or +5 were negative, they were not used to project earnings for subsequent years. For about five percent of our sample, explicit earnings forecasts are available for all five years and do not need to be inferred using g_5. That subsample was investigated to confirm that projections based on five-year growth rates are unbiased proxies for the explicit forecasts for those years.

[10] Although this statistic is well known to macroeconomists, it is higher than average firm-level dividend payouts. Note, however, that aggregate earnings include many loss firms, especially in the early 1990s, when earnings were depressed because of write-offs and accounting changes. This results in a higher aggregate dividend payout than the average firm-level payout ratio, which is computed over profitable firms only (the payout ratio is meaningless for loss firms). Also, since the aggregate payout ratio is a value-weighted average dividend payout, it is more representative of large firms, which tend to have higher dividend payouts than small firms.

Table I
Market Capitalization, Book Values, Dividends, and Actual and Forecast Earnings for U.S. Stocks (1985 to 1998)

The market consists of firms on the I/B/E/S Summary files with forecasts for years +1, +2, and a five-year earnings growth estimate (g_5) as of April each year, and actual earnings per share, dividends per share, number of shares outstanding and share prices as of the end of the prior fiscal year (year 0). Book values of equity for year 0 are obtained from COMPUSTAT. When missing on the I/B/E/S files, forecasted earnings per share for years +3, +4, and +5 are determined by applying g_5, the forecasted five-year growth rate, to year +2 forecasted earnings. All per share numbers are multiplied by the number of shares outstanding to get amounts at the firm level, and these are added across firms to get amounts at the market level each year. All amounts, except for dividend payout, are in millions of dollars.

Forecast as of April	Number of Firms	Actual Values for Year 0: Earnings	Dividends	Payout	Book Value	Market Value	Forecast Earnings for Years +1 to +5: Year +1	Year +2	Year +3	Year +4	Year +5
	1		3			6	7	8	9	10	11
1985	1,559	154,858	71,134	46%	1,191,869	1,747,133	180,945	205,294	228,208	254,181	283,706
1986	1,613	155,201	73,857	48%	1,214,454	2,284,245	178,024	203,677	226,018	251,313	280,035
1987	1,774	146,277	81,250	56%	1,323,899	2,640,743	186,319	220,178	244,174	271,432	302,529
1988	1,735	167,676	86,237	51%	1,430,672	2,615,857	222,497	246,347	273,204	303,642	338,262
1989	1,809	229,070	97,814	43%	1,541,231	2,858,585	261,278	284,616	315,204	349,721	388,776
1990	1,889	228,216	107,316	47%	1,636,069	3,143,879	257,657	295,321	328,803	366,798	410,028
1991	1,939	218,699	108,786	50%	1,775,199	3,660,296	241,760	294,262	328,513	367,521	412,073
1992	2,106	202,275	113,962	56%	1,911,383	4,001,756	252,109	308,567	344,742	386,098	433,552
1993	2,386	247,988	127,440	51%	2,140,668	4,918,359	295,862	356,086	397,969	445,840	501,081
1994	2,784	290,081	129,186	45%	2,168,446	5,282,046	339,694	402,689	450,559	505,315	568,179
1995	2,965	365,079	147,575	40%	2,670,725	6,289,760	444,593	518,600	579,954	650,120	730,648
1996	3,360	446,663	175,623	39%	3,182,952	8,207,274	512,921	588,001	659,732	742,244	837,577
1997	3,797	547,395	201,017	37%	3,679,110	10,198,036	614,932	709,087	800,129	905,787	1,029,061
1998	3,673	526,080	178,896	34%	3,412,303	12,908,495	577,297	682,524	775,707	884,529	1,012,294

casts, as well as to project future book values (using equation (3)). The validity of this assumption is not critical; however, varying the payout ratio between 25 and 75 percent has little impact on the estimated discount rate (results available upon request).

Both short- and long-term risk-free rates have been used in studies that estimate discount rates from flows that extend over many future periods. While one-month or one-year rates are appropriate when inferring the equity premium from historic returns (observed return less risk-free yield for that period), for studies based on forecasted flows, the maturity of risk-free rates used should match that of the future flows (Ibbotson Associates (1999)). Although we allow for expected variation in risk-free rates when estimating the risk premium, using equation (5a), we find almost identical results using a constant risk-free rate in equation (5) equal to the long-term rate. In essence, the shape of the yield curves over our sample period is such that the forward rates settle rather quickly at the long-term rate, and the impact of discounting flows from earlier years in the profile at rates lower than the long-term rate is negligible. For the sensitivity analyses, we find it convenient to use the constant rate structure of equation (5), rather than the varying rate structure of equation (5a). We selected the 10-year risk-free rate for the constant risk-free rate because it is the longest maturity for which data could be obtained for all country-years in our sample. To allow comparisons with other studies that use 30-year risk-free rates, we note that the mean 30-year risk-free rate in April for each year of our U.S. sample period is 31 basis points higher than the mean 10-year risk-free rate we use.

For years beyond year +5, abnormal earnings are assumed to grow at the expected inflation rate, g_{ae}. As explained in the Appendix, the expected nominal inflation rate is higher than values of g_{ae} assumed in the literature, and is an upper bound for expected growth in abnormal earnings. We derive the expected inflation rate from the risk-free rate, based on the assumption that the real risk-free rate is approximately three percent.[11] Since we recognize that this assumption is only an educated guess, we consider in Section V.D other values of g_{ae} also. Fortunately, our estimated risk premium is relatively robust to variation in the assumed growth rate, g_{ae}, since a lower proportion of current market value is affected by g_{ae} in equations (5) and (5a), relative to the impact of g in equation (1).

III. Results

Since k appears in both the numerators (ae_t is a function of k) and denominators of the terms on the right-hand side of equation (5), the resulting

[11] The observed yields on recently issued inflation-indexed government bonds support this assumption. Although estimates of the real risk-free rate vary through time, and have historically been lower than three percent, more recently, the excess of the long-term risk-free rate over inflation forecasts has risen to three or four percent (e.g., Blanchard (1993), and discussion by Siegel).

equation is a polynomial in k with many possible roots. Empirically, however, only one root is real and positive (see Botosan (1997)). We search manually for the value of k that satisfies the relation each year, with the first iteration being close to the risk-free rate. The equity risk premium estimate (rp) that satisfies the valuation relation in equation (5a) is also estimated iteratively.

Table II provides the results of estimating rp, k, and k^*. The annual estimates for rp (in column 13) lie generally between three and four percent and are much lower than the historic Ibbotson estimate. Also, there is little variation over time: each annual estimate is remarkably close to the mean value of 3.39 percent. The annual estimates for k (in column 9) vary between a high of 14.38 percent in 1985 and a low of 8.15 percent in 1998. The corresponding risk-free rates (10-year Government T-bond yields) reported in column 8 vary with the estimated ks, between 11.43 percent in 1985 and 5.64 percent in 1998. As a result, the estimated equity premia (in column 11), equal to k less r_f, exhibit little variation around the time-series mean of 3.40 percent.

While the equation (5a) equity premium estimates (rp) derived from non-flat risk-free rates are in concept more accurate than those derived by subtracting 10-year risk-free rates from the flat k estimated from equation (5), the numbers reported in column 11 are very similar to those reported in column 13. We only consider the equation (5) estimates hereafter because (a) the magnitudes of the discount rates and their relation to risk-free rates are more transparent for the risk premium estimates based on constant risk-free rates, and (b) forward one-year rates for different maturities are not available for the other five markets,.

To understand better the relative magnitudes of the terms in equation (5), we report in the first seven columns of Table II the fraction of market values represented by each term. The fraction represented by book value (column 1) has generally declined over our sample period, from 68.2 percent in 1985 to 26.4 percent in 1998. To compensate, the fraction represented by terminal value (column 7) has increased from 26.6 percent in 1985 to 60 percent in 1998. The fraction represented by abnormal earnings for years +1 to +5 has also increased.

Column 10 of Table II contains our estimates for k^*, the market discount rate based on the dividend growth model described by equation (1), when dividends are assumed to grow in perpetuity at the five-year growth in earnings forecast (g_5). Since g_5 is not available at the aggregate level, we use the forecast growth in aggregate earnings from year +4 to +5 (see column 16 of Table V) to identify g_5 at the market level. To maintain consistency with prior research using the dividend growth model, we estimate d_1 by applying the earnings growth forecast for year 1 on prior year dividends ($d_1 = d_0 * e_1/e_0$). Our estimates for k^* are almost identical to those reported by Moyer and Patel (1997).[12] Note that these estimates of k^* are much larger than the

[12] Similar results are expected because the underlying data is taken from the same source, with minor differences in samples and procedures; for example, they use the S&P 500 index whereas we use all firms with available data.

Table II

Implied Expected Rate of Return on the Market (k and k^*) and Equity Risk Premium (rp and $k - r_f$) for U.S. Stocks (1985 to 1998)

The market is an aggregate of firms on the I/B/E/S Summary files with forecasts for years +1, +2, and a five-year earnings growth estimate (g_5) as of April each year, and actual earnings, dividends, number of shares outstanding and prices as of the end of the prior full fiscal year (year 0). Book values of equity for year 0 (bv_0) are obtained from COMPUSTAT. When missing, forecasted earnings for years +3, +4, and +5 are determined by applying g_5, the forecasted five-year growth rate, to year +2 forecasted earnings. The implied discount rate that satisfies the valuation relation in equation (5) below is k. Abnormal earnings (ae_t) equal reported earnings less a charge for the cost of equity (= beginning book value of equity * k). Assuming that 50 percent of earnings are retained allows the estimation of future book values from current book values and forecast earnings. The terminal value represents all abnormal earnings beyond year +5. Those abnormal earnings are assumed to grow at a constant rate, g_{ae}, which is assumed to equal the expected inflation rate, and is set equal to the current 10-year risk-free rate less 3 percent. The expected rate of return on the market is also estimated using equation (1), and is labeled k^*. Equation (1) is derived from the dividend growth model, and dividend growth in perpetuity, g, is assumed to equal the five-year earnings growth rate, g_5. Subtracting r_f from the discount rates k and k^* generates equity premium estimates. The equity premium (rp) is also estimated using equation (5a), which is based on the same information used in equation (5), except that the constant discount rate k is replaced by forward one-year risk-free rates at different maturities (r_{fs}) plus a constant risk premium (rp). All amounts, except for rates of return, are in millions of dollars.

$$k^* = \frac{d_1}{p_0} + g \tag{1}$$

$$p_0 = bv_0 + \frac{ae_1}{(1+k)} + \frac{ae_2}{(1+k)^2} + \frac{ae_3}{(1+k)^3} + \frac{ae_4}{(1+k)^4} + \frac{ae_5}{(1+k)^5} + \left[\frac{ae_5(1+g_{ae})}{(k-g_{ae})(1+k)^5}\right] \tag{5}$$

$$p_0 = bv_0 + \sum_{t=1}^{\infty}\left[\frac{ae_t}{\prod_{s=1}^{t}(1+r_{fs}+rp)}\right] \tag{5a}$$

Forecast as of April	Book Value as Percent of Market Value	Percent of Market Value Represented by Present Value of											
		ae_1	ae_2	ae_3	ae_4	ae_5	Terminal Value	10-year r_f	k from (5)	k^* from (1)	$k - r_f$	$k^* - r_f$	rp from (5a)
	1			4			7		9	10			13
1985	68.2%	0.5%	0.9%	1.1%	1.3%	1.5%	26.6%	11.43%	14.38%	16.14%	2.95%	4.71%	2.88%
1986	53.2%	1.6%	2.0%	2.1%	2.3%	2.4%	36.3%	7.30%	11.28%	14.90%	3.98%	7.60%	4.03%
1987	50.1%	1.3%	1.9%	2.1%	2.2%	2.3%	40.0%	8.02%	11.12%	15.08%	3.10%	7.06%	3.25%
1988	54.7%	1.7%	1.8%	1.9%	2.0%	2.2%	35.7%	8.72%	12.15%	15.52%	3.43%	6.80%	3.58%
1989	53.9%	2.0%	2.0%	2.0%	2.1%	2.2%	35.7%	9.18%	12.75%	14.85%	3.57%	5.67%	3.54%
1990	52.0%	1.6%	2.0%	2.1%	2.2%	2.3%	37.8%	8.79%	12.33%	15.41%	3.54%	6.62%	3.56%
1991	48.5%	1.1%	1.9%	2.0%	2.2%	2.4%	41.8%	8.04%	11.05%	15.16%	3.01%	7.12%	2.96%
1992	47.8%	1.1%	1.9%	2.1%	2.3%	2.5%	42.4%	7.48%	10.57%	15.55%	3.09%	8.07%	3.06%
1993	43.5%	1.7%	2.3%	2.5%	2.7%	2.9%	44.4%	5.97%	9.62%	15.12%	3.65%	9.15%	3.76%
1994	41.1%	2.1%	2.6%	2.8%	2.9%	3.1%	45.5%	5.97%	10.03%	15.02%	4.06%	9.05%	3.53%
1995	42.5%	2.1%	2.6%	2.7%	2.8%	3.0%	44.3%	7.06%	11.03%	14.96%	3.97%	7.90%	4.02%
1996	38.8%	2.2%	2.5%	2.6%	2.8%	3.0%	48.2%	6.51%	9.96%	14.96%	3.45%	8.45%	3.50%
1997	36.1%	2.2%	2.5%	2.6%	2.8%	3.0%	50.8%	6.89%	10.12%	13.88%	3.23%	6.99%	3.25%
1998	26.4%	2.1%	2.5%	2.7%	3.0%	3.2%	60.0%	5.64%	8.15%	13.21%	2.51%	7.57%	2.53%
Mean								7.64%	11.04%	14.98%	3.40%	7.34%	3.39%

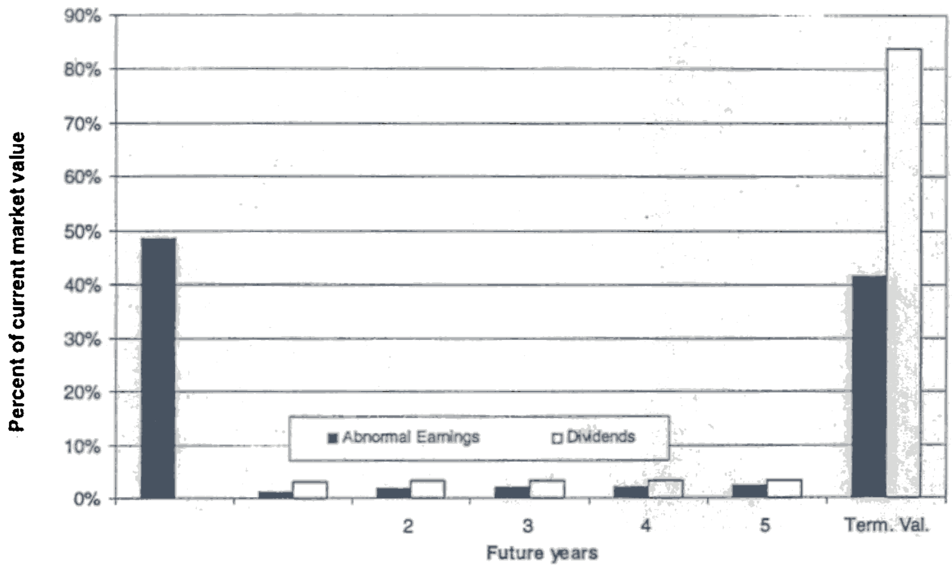


Figure 1. Comparison of value profile for abnormal earnings versus dividends, for abnormal earnings approach for U.S. stocks as of April, 1991. Based on the data in Table II, for the abnormal earnings approach described by equation (5), abnormal earnings are assumed to grow at 5.04 percent, the anticipated inflation rate, past year +5, and the resulting market discount rate (k) is 11.05 percent. For the abnormal earnings profile, the fractions represented by book value, abnormal earnings in years +1 through +5, and the terminal value are shown by the solid columns. For the dividend profile corresponding to those abnormal earnings projections, the fractions of current market capitalization that are represented by dividends in years +1 through +5 and the terminal value are shown by the hollow columns.

corresponding values of k, and the implied equity premium estimates reported in column 12 ($k^* - r_f$) are about twice those in column 11 ($k - r_f$). The mean equity premium of 7.34 percent in column 12 of Table II is approximately the same as the Ibbotson estimate. Note also the larger variation in column 12, around this mean, relative to the variation in columns 11 and 13.

The results in Table II can be used to illustrate two primary advantages of the abnormal earnings model over the dividend growth model. First, the abnormal earnings approach uses more available "hard" data (current book value and forecast abnormal earnings for years +1 to +5) to reduce the emphasis on "softer" growth assumptions (g_{ae}) used to build terminal values. Figure 1 contains a value profile for the terms in equation (5), using data for 1991. This year was selected because it represents a "median" profile: the terminal value is a smaller (larger) fraction of total value for years before (after) 1991. Recall from Table II that our estimate for k in 1991 is 11.05 percent. The terminal value is based on abnormal earnings growing at an anticipated inflation rate of 5.04 percent (g_{ae} is three percent less than the risk-free rate of 8.04 percent). The value profile for the abnormal earnings model, represented by the solid columns in Figure 1, shows that approximately 50 percent of the total value is captured by current book value, 10 percent is spread over the abnormal earnings for the next five years, and about 40 percent remains in the terminal value. This last term is the only one affected by our growth assumption. In contrast, for the dividend growth model in equation (1), the dividend growth rate (g), which is assumed to equal the five-year analyst forecast for earnings growth (g_5 = 12.12 percent), is the primary determinant of the estimated k^* (= 15.16 percent).

To offer a different perspective on why growth assumptions are more influential for projected dividends, relative to abnormal earnings, we converted the abnormal earnings profile in Figure 1 to an isomorphic value profile for dividends, represented by the hollow columns in Figure 1. (Note that these dividends refer to the flows underlying k, from the abnormal earnings model, and are different from the flows underlying k^*, the dividend growth model estimate.) The year +5 terminal value for the dividend profile in Figure 1 corresponds to a dividend growth in perpetuity of 6.8 percent.[13] Even though the abnormal earnings and dividend profiles in Figure 1 correspond to the same underlying projections, the terminal value for the dividend profile represents almost 85 percent of total value. As a result, assumed dividend growth rates have a larger impact on estimated discount rates, relative to abnormal earnings growth rate assumptions. For example, doubling the assumed value of g_{ae} to 10 percent increases the estimated discount rate by only about two percentage points. In contrast, increasing the dividend growth assumption by one percentage point raises the estimated discount rate by almost the same amount.[14]

The second major benefit of the abnormal earnings approach is that we can narrow the range of reasonable growth assumptions (g_{ae}), relative to the assumed growth rate for dividends (g). Since g is a hypothetical rate, it is not easy to determine whether 12.12 percent (the value of g underlying our 1991 estimate for k^*) is more or less reasonable than the 6.8 percent dividend growth in perpetuity (after year +5) implied by our abnormal earnings model projections. Fortunately, restating implied dividend growth rates in terms of terminal growth in abnormal earnings makes it easier to see why some dividend growth assumptions are unreasonable. The assumption that dividends grow at 12.12 percent implies that abnormal earnings past year +5 would need to grow in perpetuity at about 15 percent per year in equa-

[13] This dividend growth rate is obtained by using equation (1) on projected market value in year +5, rather than current market values (p_0) and the dividend in year six is the dividend in year +5 (= 50 percent of the earnings forecast for year +5) times the unknown growth rate. That is, solve for g in the relation $p_5 = d_5(1 + g)/(k - g)$.

[14] Note that in equation (1), changes in g increase k^* by exactly the same amount. For the dividend value profile in Figure 1, however, dividends for years +1 to +5 have been fixed by forecasted earnings and dividend payout assumptions. Therefore, increases in the dividend growth rate underlying the terminal value increase the estimated discount rate by a slightly smaller amount.

tion (5). This abnormal earnings growth rate corresponds to a real growth in rents of 10 percent (assumed long-term inflation rate is 5.04 percent), which is clearly an unreasonably optimistic assumption.

In sum, our estimates of the equity risk premium using the abnormal earnings approach are considerably lower than the Ibbotson rate, even though we believe the analyst forecasts we use, as well as the terminal growth assumptions we make, are optimistic. Adjusting for such optimism would lower our estimates further. While our estimates from the dividend growth approach are much closer to the Ibbotson rate, we believe they are biased upward because the assumed growth rate ($g = g_5$) is too high an estimate for dividend growth in perpetuity. The estimates from the abnormal earnings approach are more reliable because we use more available information to reduce the importance of assumed growth rates, and we are better able to reject growth rates as being infeasible by projecting rents rather than dividends. Additional benefits of using the abnormal earnings approach are illustrated in Section V.

IV. Equity Premium Estimates from Other Markets

Other equity markets offer a convenient opportunity to validate our domestic results. As long as the different markets are integrated with the United States and are of similar risk, those markets' estimates should proxy for the equity premium in the United States. We replicated the U.S. analysis on five other important equity markets with sufficient data to generate reasonably representative samples of those markets. Only a summary of our results is provided here; details of those analyses are in Claus and Thomas (1999b). The six markets exhibit considerable diversity in performance and underlying fundamentals over our sample period. This across-market variation increases the likelihood that the estimates we obtain from each market offer independent evidence.

As with the U.S. data, earnings forecasts, actual earnings per share, dividends per share, share prices, and the number of outstanding shares are obtained from I/B/E/S. Book values of equity as of the end of year 0 are collected from COMPUSTAT and Global Vantage for Canada and from Datastream for the remaining four countries. Unlike I/B/E/S and COMPUSTAT, Datastream drops firms that are no longer active. While such deletions are less frequent outside the United States, only surviving firms are included in our sample. Fortunately, no bias is created in this study since we equate market valuations with contemporaneous forecasts, and do not track performance.[15] Therefore, even if the surviving firms (included in our sample) performed systematically better or worse than firms that were dropped, our equity premium estimates are unbiased as long as market prices and earnings forecasts in each year are efficient and incorporate the same information.

[15] Note that there is no "backfilling" in our sample, where prior years' data for successful firms are entered subsequently.

All data are denominated in local currency. Currency risk is not an issue here, since it is present in the required rates of returns for both equities and government bonds. Thus the difference between the two rates should be comparable across countries.

We find that analysts' forecasts in these five markets exhibit an optimism bias, similar to that observed in the United States. We considered other potential sources of measurement error in the forecasts, but are confident that any biases created by these errors are unlikely to alter our equity premium estimates much. For example, in Germany, earnings could be computed in as many as four different ways: GAAP per International Accounting Standards, German GAAP, DVFA, and U.S. GAAP.[16] I/B/E/S employees indicated that they have been more successful at achieving consistency in recent years (all forecasts are on a DVFA basis), but they are not as certain about earlier years in their database. While differences in basis between forecast and actual items would affect analyst bias, they do not affect our estimates of market discount rates. Differences in basis across analysts contaminate the consensus numbers used, but the estimated market discount rates are relatively insensitive to changes in the near-term forecasts used.

To select the month of analysis for each country, we followed the same logic as that for the U.S. analysis. December was the most popular fiscal year-end for all countries except for Japan, where it was March. We then identified the period after the fiscal year-end by which annual earnings are required to be disclosed. This period differs across countries (see Table 1 in Alford et al. (1993)): it is three months for Japan and the United States, four months for France, six months for Canada and the United Kingdom, and eight months for Germany. We selected the month following the reporting deadline as the "sure to be disclosed" month to collect forecasts for any given year.

To include a country-year in our sample, we required that the total market value of all firms in our sample exceed 35 percent of the market value of "primary stock holdings" for that country, as defined by Datastream. Although we used a low hurdle to ensure that our sample contained contiguous years for all countries, a substantially greater proportion of the Datastream Market Index than our minimum hurdle is represented for most country-years.

The equity-premium estimates using the abnormal earnings and dividend growth approaches as well as the prevailing risk-free rates for different country-year combinations with sufficient data are reported in Table III. The number of years with sufficient firms to represent the overall market was highest for Canada (all 14 years between 1985 and 1998), and lowest for Japan (8 years). As with the U.S. sample, we use a 50 percent aggregate

[16] The German financial analyst society, Deutsche Vereinigung für Finanzanalyse (DVFA), has developed a system used by analysts (and often by firms) to adjust reported earnings data to provide a measure that is closer to permanent or core earnings. The adjustment process uses both reported financial information as well as firms' internal records. GAAP refers to Generally Accepted Accounting Principles or the accounting rules under which financial statements are prepared in different domiciles.

Table III

Implied Equity Premium Using Abnormal Earnings and Dividend Growth Approaches ($k - r_f$ and $k^* - r_f$) for International Stocks (1985 to 1998)

The market is an aggregate of firms on the I/B/E/S Summary files with forecasts for years +1, +2, and a five-year earnings growth estimate (g_5) as of April each year, and actual earnings, dividends, number of shares outstanding, and prices as of the end of the prior full fiscal year (year 0). Book values of equity for year 0 (bv_0) are obtained from COMPUSTAT, Global Vantage, and Datastream. Forecasted earnings for years +3, +4, and +5 are determined by applying g_5, the forecasted 5-year growth rate, to year +2 forecasted earnings. All amounts are measured in local currencies. r_f is the 10-year government bond yield. The implied discount rate that satisfies the valuation relation in equation (5) below is k. Abnormal earnings (ae_t) equal reported earnings less a charge for the cost of equity (= beginning book value of equity * k). Assuming that 50% of earnings are retained allows the estimation of future book values from current book values and forecast earnings. The terminal value represents all abnormal earnings beyond year +5. Those abnormal earnings are assumed to grow at a constant rate, g_{ae}, which is assumed to equal the expected inflation rate, and is set equal to r_f less 3 percent. The expected rate of return on the market is also estimated using equation (1), and is labeled k^*. Equation (1) is derived from the dividend growth model, and dividend growth in perpetuity, g, is assumed to equal the five-year earnings growth rate, g_5.

$$p_0 = bv_0 + \frac{ae_1}{(1+k)} + \frac{ae_2}{(1+k)^2} + \frac{ae_3}{(1+k)^3} + \frac{ae_4}{(1+k)^4} + \frac{ae_5}{(1+k)^5} + \left[\frac{ae_5(1+g_{ae})}{(k-g_{ae})(1+k)^5}\right] \tag{5}$$

$$k^* = \frac{d_1}{p_0} + g \tag{1}$$

	Canada			France			Germany			Japan					
			k^*						k^*			k^*			k^*
	10.50%	4.41%	7.45%												
1986	8.82%	2.93%													
	9.16%	1.56%	4.53%			5.06%									
	9.66%	2.83%	4.67%			3.90%	6.78%	3.43%	4.59%						
	9.29%	3.08%		8.76%	3.64%	**5.11%**	**6.83%**	**3.87%**	5.48%				10.16%		7.24%
		1.51%	2.97%	9.66%	3.04%	4.23%	8.99%	1.10%	3.23%						
			3.71%	8.81%	2.94%	4.41%	8.42%		4.72%	6.72%	0.95%	0.38%			7.27%
	8.18%	0.42%	6.36%			**5.81%**	**7.89%**	**2.16%**	**5.03%**	**5.38%**		0.34%	9.12%	2.77%	8.69%
	7.32%					10.57%	6.14%	0.70%	4.19%	4.45%	1.05%		7.64%	3.29%	
			7.67%					1.30%	3.77%	4.24%	1.04%	4.56%			5.50%
			6.77%			**10.04%**	**6.70%**	**2.22%**	3.84%	2.80%		9.50%			
			6.89%			**12.26%**	**6.41%**	**2.14%**	3.40%	3.17%	0.79%	7.82%	7.92%	3.34%	8.43%
						9.69%	5.68%	2.28%	11.56%	2.47%	1.65%		7.02%	2.53%	
			7.44%	5.02%	2.53%						1.99%	10.89%	5.84%	2.09%	6.77%
			5.89%			7.90%	11%	2.02%	6.58%	3.86%	0.21%	5.83%			7.91%
	1.55%	1.04%	1.62%			3.27%	1.04%	1.03%		1.67%	1.31%				

dividend payout ratio to generate future dividends and book values, and assume that abnormal earnings grow at the expected inflation rate, which is assumed to be three percent less than the prevailing risk-free rate. For the few years when r_f in Japan is below three percent, we set $g_{ae} = 0$.

The equity premium values based on the abnormal earnings approach $(k - r_f)$ generally lie between two and three percent, except for Japan, where the estimates are considerably lower (and even negative in the early 1990s). Finding that none of the almost 70 estimates of $k - r_f$ reported in Tables II and III are close to the Ibbotson estimate suggests strongly that that historical estimate is too high. In contrast, the equity premium estimates based on the dividend growth approach with dividends growing in perpetuity at the five-year earnings growth forecast (g_5) are considerably higher, similar to the pattern observed in the United States. The dividend growth estimates are very close to those reported in Khorana et al. (1997), which uses a similar approach and a similar sample.

Repeating the sensitivity analyses conducted on the United States (described in Section V) on these five markets produced similar conclusions. The abnormal earnings estimates generate projections that are consistent with experience, but the dividend growth estimates are biased upward and generate projections that are too optimistic because the five-year earnings growth forecast (g_5) is too high an estimate for dividend growth in perpetuity. The values of g_5 suggest mean real dividend growth rates in perpetuity that range between 6.09 percent for Canada and 8.25 percent for Japan. These real rates exceed historic real earnings growth rates, and are at least twice as high as the real GDP growth rates forecast for these countries.

The results observed for Japan are unusual and invite speculation. While our results suggest that the equity premium in Japan increased during the sample period, from about -1 percent in the early 1990s to 2 percent in the late 1990s, these results are also consistent with a stock market bubble that has gradually burst. That is, early in our sample period, prices were systematically higher than the fundamentals (represented by analysts' forecasts) would suggest, and have gradually declined to a level that is supported by analysts' forecasts. Note that our sample excludes the peak valuations in the late 1980s before the crash. Perhaps the implied equity premium in that period would be even more negative than the numbers we estimate for the early 1990s. Regardless of whether the poor performance of Japanese equities in the 1990s is due to correction of an earlier mispricing, it is useful to contrast the inferences from a historic approach with those from a forward-looking approach such as ours: the former would conclude that equity premia have fallen in Japan during the 1990s, whereas our approach suggests the opposite.

V. Sensitivity Analyses

This section summarizes our analysis of U.S. equity data designed to gauge the robustness of our conclusion that the equity premium is much lower than historic estimates. We begin by considering two relations for P/B and P/E ratios that allow us to check whether our projections under the dividend growth and abnormal earnings models are reasonable. Next, we document the extent of analyst optimism in our data. Finally, we consider the sensitivity of our risk premium estimates to the assumed abnormal earnings growth rate (g_{ae}).[17]

A. P/B Ratios and the Level of Future Profitability

The first relation we examine is that between the P/B ratio and future levels of profitability (e.g., Penman (1999)), where future profitability is the excess of the forecast market accounting rate of return (roe_t) over the required rate of return, k.

$$\frac{p_0}{bv_0} = 1 + \frac{roe_1 - k}{(1+k)} + \frac{roe_2 - k}{(1+k)^2}\left(\frac{bv_1}{bv_0}\right) + \frac{roe_3 - k}{(1+k)^3}\left(\frac{bv_2}{bv_0}\right) + \dots, \tag{6}$$

where $roe_t = e_t/bv_{t-1}$ is the accounting return on equity in year t.

This relation indicates that the P/B ratio is explained by expected future profitability ($roe_t - k$).[18] Firms expected to earn an accounting rate of return on equity equal to the cost of capital should trade currently at book values ($p_0/bv_0 = 1$). Similarly, the P/B ratio expected in year +5 (p_5/bv_5), which is determined by the assumed growth in abnormal earnings after year +5 (g_{ae}), should be related to profitability beyond year +5. To investigate the validity of our assumed growth rates, we examine the profiles of future P/B ratios and profitability levels to check if they are reasonable and related to each other as predicted by equation (6). Future book values are generated by adding projected earnings and subtracting projected dividends (assuming a 50 percent payout) to the prior year's book value. Similarly, projected market values are obtained by growing the prior year's market value at the discount rate (k) less projected dividends.

Table IV provides data on current and projected values of P/B ratios and profitability. Current market and book values are reported in columns 1 and 2, and projected market and book values in year +5 are reported in columns

[17] We also examined Value Line data for the DOW 30 firms for two years: 1985 and 1995 (details in Claus and Thomas (1999a)). Value Line provides both dividend forecasts (over a four- or five-year horizon) and a projected price. This price is, in effect, a terminal value estimate, which obviates the need to assume dividend growth in perpetuity. Unfortunately, those risk premium estimates appear to be unreliable: The estimated discount rate is 20 percent (8.5 percent) for 1985 (1995). These results are consistent with Value Line believing that the DOW 30 firms are undervalued (overvalued) in 1985 (1995); that is, current price does not equal the present value of forecast dividends and projected prices. This view is supported by their recommendations for the proportion to be invested in equity: it was 100 percent through the 1980s, and declined through the 1990s (it is currently at 40 percent).

[18] The growth in book value terms in equation (6), bv_t/bv_0, which add a multiplicative effect, have been ignored in the discussion because of the built-in correlation with $roe_t - k$. Higher roe_t results in higher e_t, which in turn causes higher growth in bv_t because dividend payouts are held constant at 50 percent for all years.

Table IV

Price-to-Book Ratios (p_t/bv_t), Forecast Accounting Return on Equity (roe_t) and Expected Rates of Return (k) for U.S. Stocks (1985 to 1998)

To examine the validity of assumptions underlying k, which is the implied discount rate that satisfies the valuation relation in equation (5), current price-to-book ratios are compared with estimated future returns on equity (roe_t) to examine fit with equation (6) below. The market is an aggregate of firms on the I/B/E/S Summary files with forecasts for years +1, +2, and a five-year earnings growth estimate (g_5) as of April each year, and actual earnings, dividends, number of shares outstanding, and prices as of the end of the prior full fiscal year (year 0). Book values of equity for year 0 (bv_0) are obtained from COMPUSTAT. When missing, forecasted earnings for years +3, +4, and +5 are determined by applying g_5 to year +2 forecasted earnings. Assuming that 50 percent of earnings are retained allows the estimation of future book values from current book values and forecast earnings. Return on equity (roe_t) equals forecast earnings scaled by beginning book value of equity (bv_{t-1}). Market and book value amounts are in millions of dollars.

$$p_0 = bv_0 + \frac{ae_1}{(1+k)} + \frac{ae_2}{(1+k)^2} + \frac{ae_3}{(1+k)^3} + \frac{ae_4}{(1+k)^4} + \frac{ae_5}{(1+k)^5} + \left[\frac{ae_5(1+g_{ae})}{(k-g_{ae})(1+k)^5}\right] \tag{5}$$

$$\frac{p_0}{bv_0} = 1 + \frac{roe_1 - k}{(1+k)} + \frac{roe_2 - k}{(1+k)^2}\left(\frac{bv_1}{bv_0}\right) + \tag{6}$$

Forecasts as of April	Year 0 Equity Values: Market Value (p_0)	Year 0 Equity Values: Book Value (bv_0)	Year +5 Equity Values: Market Value (p_5)	Year +5 Equity Values: Book Value (bv_5)	Price/Book Ratio: In Year 0 (p_0/bv_0)	Price/Book Ratio: In Year 5 (p_5/bv_5)	Forecast Accounting Return on Equity: In Year 1 (roe_1)	In Year 2 (roe_2)	In Year 3 (roe_3)	In Year 4 (roe_4)	In Year 5 (roe_5)	In Year 6 (roe_6)	k from Eq. (5)
	1	2	3	4	5	6	7	8	9	10	11	12	13
1985	1,747,133	1,191,869	2,676,683	1,768,036	1.5	1.5	15%	16%	16%	17%	17%	17%	14.38%
1986	2,284,245	1,214,454	3,197,490	1,783,987	1.9	1.8	15%	16%	16%	17%	17%	17%	11.28%
1987	2,640,743	1,323,899	3,727,459	1,936,215	2.0	1.9	14%	16%	16%	16%	17%	17%	11.12%
1988	2,615,857	1,430,672	3,779,033	2,122,648	1.8	1.8	16%	16%	16%	17%	17%	17%	12.15%
1989	2,858,585	1,541,231	4,200,867	2,341,029	1.9	1.8	17%	17%	17%	18%	18%	18%	12.75%
1990	3,143,879	1,636,069	4,589,685	2,465,373	1.9	1.9	16%	17%	17%	18%	18%	18%	12.33%
1991	3,660,296	1,775,199	5,181,184	2,597,264	2.1	2.0	14%	16%	16%	17%	17%	17%	11.05%
1992	4,001,756	1,911,383	5,574,848	2,773,918	2.1	2.0	13%	15%	16%	16%	17%	17%	10.57%
1993	4,918,359	2,140,668	6,595,210	3,139,088	2.3	2.1	14%	16%	16%	17%	17%	17%	9.62%
1994	5,282,046	2,168,446	7,336,322	3,301,664	2.4	2.2	16%	17%	18%	18%	19%	18%	10.47%
1995	6,289,760	2,670,725	8,837,148	4,132,682	2.4	2.1	17%	18%	18%	19%	19%	19%	11.03%
1996	8,207,274	3,182,952	11,206,787	4,853,189	2.6	2.3	16%	17%	18%	18%	19%	18%	9.96%
1997	10,198,036	3,679,110	14,103,523	5,708,609	2.8	2.5	17%	18%	18%	19%	20%	19%	10.12%
1998	12,908,495	3,412,303	16,838,377	5,378,478	3.8	3.1	17%	18%	19%	20%	21%	20%	8.15%
Mean					2.2	2.1	15%	17%	17%	18%	18%	18%	11.04%

3 and 4. These values are used to generate current and year +5 P/B ratios, reported in columns 5 and 6. Columns 7 through 12 contain the forecasted accounting rate of return on equity for years 1 to 6, which can be compared with the estimated market discount rate, k, reported in column 13, to obtain forecasted profitability.

The current P/B ratio has been greater than 1 in every year in the sample period, and has increased steadily over time, from 1.5 in 1985 to 3.8 in 1998. Consistent with equation (6), all forecasted *roe* values for years 1 through 6 in Table IV exceed the corresponding values of k. Increases in the P/B ratio over the sample period are mirrored by corresponding increases in forecast profitability ($roe_t - k$) in years +1 through +5 as well as forecast profitability in the posthorizon period (after year +5), as measured by the implied price-to-book ratio in year +5. Finally, the tendency for P/B ratios to revert gradually over the horizon toward one (indicated by the year +5 values in column 6 being smaller than the year 0 values in column 5) is consistent with intuition (e.g., Nissim and Penman (1999)).

We also extended our investigation to years beyond year +5 for the assumptions underlying the abnormal earnings estimates, and find that the pattern of projections for P/B and *roe* remain reasonable. In contrast, those projections for the assumptions underlying the dividend growth model estimates suggest that the underlying growth rates are unreasonably high. To provide an illustrative example of those results, we contrast in Figure 2 the patterns for future *roe* and P/B that are projected for the dividend growth and abnormal earnings approaches for 1991. The *roe* levels are marked off on the left scale, and P/B ratios are shown on the right scale. Recall that the market discount rates estimated for the abnormal earnings and dividend growth approaches are 11.05 percent (k) and 15.16 percent (k^*) and the corresponding terminal growth rates for abnormal earnings and dividends are 5.04 percent and 12.12 percent.

The projections for the abnormal earnings method (indicated by bold lines) continue to remain reasonable. The P/B ratio always exceeds one, but it trends down over time. Consistent with P/B exceeding one, the *roe* is always above the 11.05 percent cost of capital, and trends toward it after year +5. Note that the optimistic analyst forecasts cause *roe* projections to climb for years +1 through +5, but the subsequent decline in *roe* is because the profitability growth implied by g_{ae} (our assumed growth in abnormal earnings past year +5) is lower than that implied by g_5.

The results for the dividend growth approach illustrate the benefits of using projected accounting ratios to validate assumed growth rates. The profitability (*roe*) is actually below the cost of equity of 15.16 percent (k^*), for the first three years, even though the P/B ratio is greater than one. Thereafter, the profitability keeps increasing, to a level above 20 percent by year +15. Both the high level of profitability and its increasing trend are not easily justified, especially when they are observed repeatedly for every year in our sample. Similarly, the increasing pattern for P/B, which is projected to increase from about two to about three by year +15, is hard to justify.

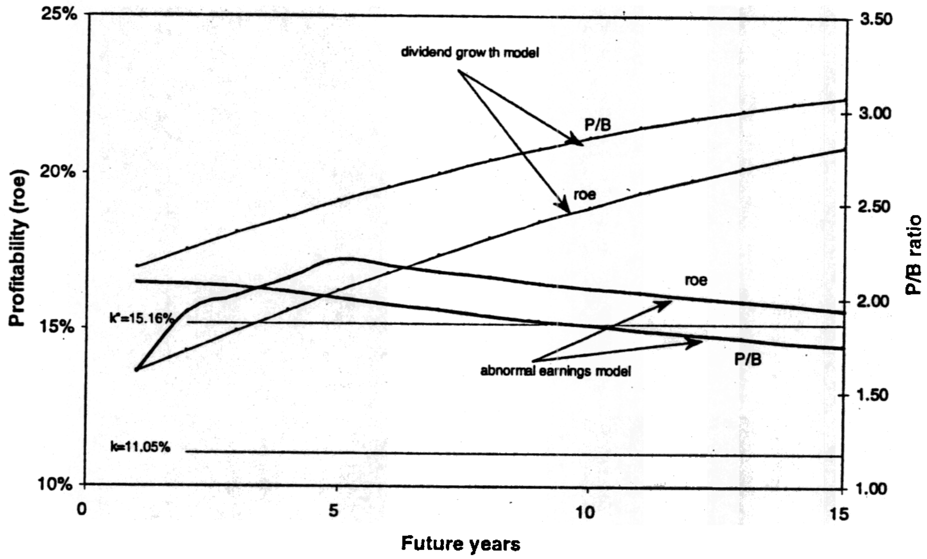


Figure 2. Pattern of future price-to-book (P/B) ratios and profitability, measured as excess of accounting return on equity (*roe*) over estimated discount rates (k^* and k), for dividend growth and abnormal earnings approaches for U.S. stocks as of April, 1991. For the dividend growth model described by equation (1) in Table II, dividends are assumed to grow at the consensus five-year earnings growth rate of 12.12 percent, and future *roe* is compared with the estimated market discount rate of 15.16 percent (k^*). For the abnormal earnings model described by equation (5) in Table II, abnormal earnings are assumed to grow at an anticipated inflation rate of 5.04 percent, and *roe* is compared with the estimated market discount rate of 11.05 percent (k). Projected P/B ratios are shown for both models.

These projections are, however, consistent with an estimated discount rate that is too high. Since near-term analysts' forecasts of profitability are below this discount rate, future levels of profitability have to be unreasonably high to compensate.

B. *P/E Ratios and Forecast Growth in Profitability*

The second relation we use to check the validity of our assumptions regarding g_{ae} is the price–earnings ratio, described by equation (7) (see derivation in Claus and Thomas, 1999a). Price–earnings ratios are a function of the present value of future changes in abnormal earnings, multiplied by a capitalization factor (= $1/k$).

$$\frac{p_0}{e_1} = \frac{1}{k}\left[1 + \frac{\Delta ae_2}{e_1(1+k)} + \frac{\Delta ae_3}{e_1(1+k)} + \dots\right], \tag{7}$$

where $\Delta ae_t = ae_t - ae_{t-1}$ is the change in expected abnormal earnings over the prior year.

The price–earnings ratio on the left-hand side deviates slightly from the traditional representation in the sense that it is a "forward" price–earnings ratio, based on expected earnings for the upcoming year, rather than a "trailing" price–earnings ratio (p_0/e_0), which is based on earnings over the year just concluded. The relation between future earnings growth and forward price–earnings ratios is simpler than that for trailing price–earnings ratios.[19] Therefore, we use only the forward price–earnings ratio here and refer to it simply as the P/E ratio.

The results reported in Table V describe P/E ratios and growth in abnormal earnings derived from analysts' forecasts for the market. The first four columns provide market values and the corresponding upcoming expected earnings for year 0 and year +5. These numbers are used to generate the current and year +5 P/E ratios reported in columns 5 and 6, which can be compared to the values of $1/k$ reported in column 18.[20] According to equation (7), absent growth in abnormal earnings, the P/E ratio should be equal to $1/k$, and the P/E ratio should be greater (less) than $1/k$ for positive (negative) expected growth in abnormal earnings. Forecast growth rates in abnormal earnings for years +2 through +6 are reported in columns 7 through 11. To maintain equivalence with the terms in equation (7), growth in abnormal earnings is scaled by earnings expected for year +1 (e_1) and then discounted.

To understand the relations among the numbers in the different columns, consider the row corresponding to 1991. The market P/E ratio of 15.1 is higher than the inverse of the discount rate ($1/k$ = 9.0). That difference of 6.1 is represented by the sum of the present value of the abnormal earnings growth terms in future years, scaled by e_1 (this sum needs to be multiplied by $1/k$ as shown in equation (7)). These growth terms decline from 13 percent in year 2 to 2 percent in year 6, and continue to decline thereafter. By year +5, the market P/E is expected to fall (to 11.7), since some of the growth in abnormal earnings (represented by the amounts in columns 7 through 11) is expected to have already occurred by then. Turning to the other sample years, the P/E ratios in year 0 (column 5) have generally increased through the sample period, and so have the values of $1/k$. Consistent with P/E ratios exceeding $1/k$ in every year, abnormal earnings are forecast to exhibit positive growth for all cells in columns 7 to 11. Also, the P/E ratios in year +5 are forecast to decline, relative to the corresponding year 0 P/E values, because of the value represented by the amounts in columns 7 to 11.

[19] Since the numerator of the P/E ratio is an ex-dividend price (p_0), the payment of a large dividend (d_0) would reduce p_0 without affecting trailing earnings (e_0), thereby destroying the relation between p_0 and e_0. This complication does not arise when expected earnings for the upcoming period (e_1) is used instead of e_0.

[20] If the numbers in Table V appear to be not as high as the trailing P/E ratios commonly reported in the popular press, note that forward P/E ratios are generally smaller than trailing P/E ratios for the following reasons. First, next year's earnings are greater than current earnings because of earnings growth. Second, current earnings contain one-time or transitory components that are on average negative, whereas forecast earnings focus on core or continuing earnings.

For purposes of comparison with other work, we also report in columns 12 through 17 of Table V the growth in forecast earnings (as opposed to growth in abnormal earnings) for years +1 through +6. Forecasted growth in earnings declines over the horizon, similar to the pattern exhibited by growth in abnormal earnings. Note the similarity in the pattern of earnings growth for all years in the sample period: the magnitudes of earnings growth estimates appear to settle at around 12 percent by year +5, before dropping sharply to values around 7 percent in the posthorizon period (year +6). Again, this decline occurs because the earnings growth implied by g_{ae} (our assumed growth in abnormal earnings past year +5) is lower than g_5.

The results in Table V confirm the predictions derived from equation (7) as well as the intuitive links drawn in the literature. As with the results for P/B ratios, the trends for P/E ratios and growth in abnormal earnings exhibit no apparent discrepancies that might suggest that the assumptions underlying our abnormal earnings model are unreasonable.

C. Bias in Analyst Forecasts

We considered a variety of biases that may exist in the I/B/E/S forecasts, but found only the well-known optimism bias to be noteworthy (details provided in Claus and Thomas (1999a)).[21] We compute the forecast error for each firm in our sample, representing the median consensus forecast as of April less actual earnings, for different forecast horizons (year +1, +2, . . . +5) for each year between 1985 and 1997. Table VI contains the median forecast errors (across all firms in the sample for each year), scaled by share price. In general, forecasted earnings exceed actual earnings, and the extent of optimism increases with the horizon.[22] There is, however, a gradual reduction in optimism toward the end of the sample period.

Since the forecast errors in Table VI are scaled by price, comparing the magnitudes of the median forecast errors with the inverse of the trailing P/E ratios (or E/P ratios) is similar to a comparison of forecast errors with earnings levels. While the trailing E/P ratios for our sample vary between 5 and 9 percent, the forecast errors in Table VI vary between values that are in the neighborhood of 0.5 percent for year +1 to around 3 percent in year +5. Comparing the magnitudes of year +5 forecast errors with the implied E/P ratios indicates that forecasted earnings exceed actual earnings by as

[21] I/B/E/S removes one-time items (typically negative) from reported earnings. That is, the level of optimism would have been even higher if we had used reported numbers instead of actual earnings according to I/B/E/S.

[22] In addition to increasing with forecast horizon, the optimism bias is greater for certain years where earnings were depressed temporarily. The higher than average dividend payouts observed in Table I for 1987 and 1992 indicate temporarily depressed earnings in those years, and the forecast errors are also higher than average for those years. For example, the two largest median year +2 forecast errors are 1.86 and 1.81 percent, and they correspond to two-year out forecasts made in 1985 and 1990.

Table V

Forward Price-to-Earnings Ratios (p_t/e_{t+1}) and Growth in Forecast Abnormal Earnings and Earnings for U.S. Stocks (1985 to 1998)

To examine the validity of assumptions underlying k, which is the implied discount rate that satisfies the valuation relation in equation (5), current and forecast forward price-to-earnings ratios are compared with growth in forecast abnormal earnings to examine fit with equation (7) below. The market is an aggregate of firms on the I/B/E/S Summary files with forecasts for years +1, +2, and a five-year earnings growth estimate (g_5) as of April each year, and actual earnings, dividends, number of shares outstanding, and prices as of the end of the prior full fiscal year (year 0). Book values of equity for year 0 (bv_0) are obtained from COMPUSTAT. Abnormal earnings (ae_t) equal reported earnings less a charge for the cost of equity (= beginning book value of equity * k). Future market values are projected for each year by multiplying beginning market values by (1 + k) and subtracting dividends. When missing, forecasted earnings for years +3, +4, and +5 are determined by applying g_5 to year +2 forecasted earnings. Assuming that 50 percent of earnings are retained allows the estimation of future book values from current book values and forecast earnings. Market equity values and earnings amounts are in millions of dollars.

$$p_0 = bv_0 + \frac{ae_1}{(1+k)} + \frac{ae_2}{(1+k)^2} + \frac{ae_3}{(1+k)^3} + \frac{ae_4}{(1+k)^4} + \frac{ae_5}{(1+k)^5} + \left[\frac{ae_5(1+g_{ae})}{(k-g_{ae})(1+k)^5}\right] \tag{5}$$

$$\frac{p_0}{e_1} = \frac{1}{k}\left[1 + \frac{\Delta ae_2}{e_1(1+k)} + \frac{\Delta ae_3}{e_1(1+k)^2} + \ldots\right. \tag{7}$$

Forecasts as of April	Year 0 Values: Market Value (p_0)	Year 0 Values: Earnings (e_1)	Year +5 Values: Market Value (p_5)	Year +5 Values: Earnings (e_6)	Forward P/E Ratio: In Year 0 (p_0/e_1)	Forward P/E Ratio: In Year 5 (p_5/e_6)	PV of *ae* Growth (Δae_t), Scaled by e_1: +2	+3	+4	+5	+6	Growth in Forecast Earnings: +1	+2	+3	+4	+5	+6	$1/k$ from Eq. (5)
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
1985	1,747,133	180,945			9.7	8.7	5%	3%	3%	3%	1%	17%	13%	11%	11%	12%	9%	
1986	2,284,245	178,024			12.8	10.7	7%	4%	5%	5%	1%	15%	14%	11%	11%	11%	7%	
1987	2,640,743	186,319			14.2	11.5	10%	5%	5%	5%	1%	27%	18%	11%	11%	11%	7%	
1988	2,615,857	222,497			11.8	10.4	4%	4%	4%	4%	1%	33%	11%	11%	11%	11%	8%	
1989	2,858,585	261,278			10.9	10.0	2%	3%	3%	4%	1%	14%	9%	11%	11%	11%	8%	
1990	3,143,879	257,667			12.2	10.4	7%	4%	4%	4%	1%	13%	15%	11%	12%	12%	8%	
1991	3,660,296	241,760			15.1	11.7	13%	5%	6%	6%	2%	11%	22%	12%	12%	12%	7%	
1992	4,001,756	252,109			15.9	12.0	14%	6%	6%	6%	2%	25%	22%	12%	12%	12%	7%	
1993	4,918,359	295,862			16.6	12.4	13%	6%	7%	7%	1%	19%	20%	12%	12%	12%	6%	
1994	5,282,046	339,694			15.5	11.9	11%	6%	6%	7%	1%	17%	19%	12%	12%	12%	6%	
1995	6,289,760	444,593			14.1	11.3	9%	5%	6%	6%	2%	22%	17%	12%	12%	12%	7%	
1996	8,207,274	512,921			16.0	12.5	8%	6%	7%	7%	2%	15%	15%	12%	13%	13%	7%	
1997	10,198,036	614,932			16.6	12.8	8%	7%	7%	8%	2%	19%	16%	11%	12%	12%	7%	
1998	12,908,495	577,297			22.4	15.7	12%	9%	10%	11%	2%	19%	16%	11%	12%	12%	7%	
Mean					14.6	11.6	9%	5%	6%	6%	1%	19%	16%	11%	12%	12%	7%	

Table VI

Optimism Bias in I/B/E/S Forecasts for U.S. Stocks: Median Forecast Errors for Forecasts Made Between 1985 and 1997

The following table represents the median of all forecast errors scaled by share price for each year examined. The forecast error is calculated for each firm as of April each year, and equals the median consensus forecasted earnings per share minus the actual earnings per share, scaled by price. The year when the forecasts were made is listed in the first row, while the first column lists the horizon of that forecast. For each year and horizon combination, we report the median forecast error and the number of firms in the sample. To interpret the Table, consider the values of 0.78 percent and 1,680 reported for the +1/1985 combination., in the top left-hand corner of the table. This means that the median value of the difference between the forecasted and actual earnings for 1986 was 0.78 percent of price, and that sample consisted of 1,680 firms with available forecast errors. The results confirm that analyst forecasts are systematically positively biased and that this bias increases with the forecast horizon; however, the extent of any such bias has been declining steadily over time.

		Year Forecast Was Made													
		1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	Mean
Forecast	Median	0.78%	0.65%	0.37%	0.07%	0.44%	0.58%	0.39%	0.17%	0.15%	0.03%	0.04%	0.00%	0.00%	0.28%
Year +1	Obs.	1,680	1,707	1,878	1,815	1,868	1,932	1,959	2,176	2,492	2,710	2,895	3,261	3,462	
Forecast	Median	2.05%	1.40%	0.79%	0.99%	1.74%	1.88%	1.21%	0.87%	0.58%	0.34%	0.32%	0.27%	—	1.04%
Year +2	Obs.	1,545	1,572	1,732	1,701	1,757	1,815	1,896	2,084	2,287	2,594	2,694	2,852	—	
Forecast	Median	2.84%	0.99%	1.44%	2.22%	2.78%	2.39%	1.50%	0.95%	0.63%	0.54%	0.45%	—	—	1.52%
Year +3	Obs.	1,406	1,449	1,596	1,576	1,634	1,744	1,826	1,936	2,159	2,396	2,346	—	—	
Forecast	Median	2.63%	2.04%	2.80%	3.19%	3.17%	2.83%	1.54%	0.91%	0.77%	0.60%	—	—	—	2.05%
Year +4	Obs.	1,285	1,344	1,492	1,474	1,586	1,696	1,724	1,825	2,024	2,132	—	—	—	
Forecast	Median	3.54%	3.44%	3.86%	3.59%	3.43%	2.91%	1.36%	0.94%	0.74%	—	—	—	—	2.65%
Year +5	Obs.	1,201	1,260	1,411	1,432	1,528	1,621	1,618	1,704	1,815	—	—	—	—	

much as 50 percent at that horizon. These results suggest that our equity premium estimates are biased upward because we do not adjust for the considerable optimism in earnings forecasts for years +1 to +5. They also suggest that we are justified in dropping assumed growth rates for earnings past year +5 (column 17 versus column 16 in Table V).

D. Impact of Variation in the Assumed Growth Rate in Abnormal Earnings Beyond Year +5 (g_{ae})

We begin by considering two alternative cases for g_{ae}: three percent less and three percent more than our base case, where g_{ae} is assumed to equal the expected inflation rate. As mentioned in the Appendix, our base growth rate of $g_{ae} = r_f - 3\%$ is higher than any rate assumed in the prior abnormal earnings literature. Adding another three percent to the growth rate, which would require rents to grow at a three percent real rate in perpetuity, raises the level of optimism further. Dropping three percent from the base case, in the lower growth scenario, would be equivalent to assuming a very low nominal growth rate in abnormal earnings, and would be only slightly more optimistic than the assumptions in much of the prior abnormal earnings literature.

For the higher (lower) growth rate scenario, corresponding to $g_{ae} = r_f$ ($g_{ae} = r_f - 6\%$), the average risk premium over the 14-year sample period increases (decreases) to a mean of 4.66 (2.18), from a mean of 3.40 percent for the base case. Even for the high growth rate in abnormal earnings, the increase in the estimated risk premium is modest, and leaves it substantially below the traditional estimates of the risk premium. While increasing (decreasing) the growth rate increases (decreases) the terminal value, it also reduces (increases) the present value of that terminal value because of the higher (lower) discount rate it engenders.

We also considered a synthetic market portfolio each year constructed to have no expected future abnormal earnings, to avoid the need for an assumed abnormal earnings growth rate beyond year +5. As described in equation (6), portfolios with P/B = 1 should exhibit no abnormal earnings; that is, the roe_t should on average equal k for this synthetic market. The last term in equation (5), representing the terminal value of abnormal earnings beyond year +5, is set to zero and the estimates for k obtained iteratively each year. The mean estimate for $k - r_f$ from this synthetic market is 2.20 percent, which is slightly lower than the mean risk premium of 3.40 percent in Table II. Note that a lower discount rate is not expected for the synthetic market, since it has a beta close to one each year and has a lower P/B than the market. (Low P/B firms are expected to generate higher returns (e.g., Gebhardt, Lee, and Swaminathan (forthcoming).) The higher discount rates observed for the assumptions underlying our abnormal earnings model support our view that the analyst forecasts we use and our assumption that the terminal growth in abnormal earnings equals expected inflation ($g_{ae} = r_f - 3\%$) are both optimistic.

VI. Conclusion

Barring some notable exceptions (e.g, Siegel (1992 and 1998), Blanchard (1993), Malkiel (1996), and Cornell (1999)), academic financial economists generally accept that the equity premium is around eight percent, based on the performance of the U.S. market since 1926. We claim that these estimates are too high for the post-1985 period that we examine, and the equity premium is probably no more than three percent. Our claim is based on estimates of the equity premium obtained for the six largest equity markets, derived by subtracting the 10-year risk-free rate from the discount rate that equates current prices to forecasted future flows (derived from I/B/E/S earnings forecasts). Growth rates in perpetuity for dividends and abnormal earnings need to be much higher than is plausible to justify equity premium estimates of about eight percent. Not only are such growth rates substantially in excess of any reasonable forecasts of aggregate growth (e.g., GDP), the projected streams for various indicators, such as price-to-book and price-to-earnings ratios, are also internally contradictory and inconsistent with intuition and past experience.

We agree that the weight of the evidence provided by the historical performance of U.S. stock markets since 1926 is considerable. Yet there are reasons to believe that this performance exceeded expectations, because of potential declines in the equity premium, good luck, and survivor bias. While projecting dividends to grow at earnings growth rates forecast by analysts provides equity premium estimates as high as eight percent, we show that those growth forecasts exhibit substantial optimism bias and need to be adjusted downward. In addition to our results, theory-based work, historical evidence from other periods and other markets, and surveys of institutional investors all suggest that the equity premium is much lower than eight percent. Overall, we believe that an eight percent equity premium is not supported by an analysis that compares current market prices with reasonable expectations of future flows for the markets and years that we examine.

Appendix: Assumed Growth Rates in Perpetuity for Dividends (g) and Abnormal Earnings (g_{ae})

While the conceptual definition of g is clear—it is the dividend growth rate that can be sustained in perpetuity, given current capital and future earnings[23]—determining this rate from fundamentals is not easy. To illustrate, take two firms that are similar in every way, except that they have announced different dividend policies in the current period, which results in a higher expected forward dividend yield (d_1/p_0) for one firm than the other, say 7 percent and 1 percent. What can be said about g for the two firms?

[23] Assuming too high a rate would cause the capital to be depleted in some future period, and assuming too low a rate would cause the capital to grow "too fast."

Examination of equation (1) indicates that g for the low dividend yield firm must be 6 percent higher than g for the higher dividend yield firm, assuming they both have the same discount rate (k^*). If k^* equals 10 percent, for example, the value of g for the two firms must be 3 percent and 9 percent. These two values of g are substantially different from each other, even though the two firms are not.

In addition to being a hypothetical rate, g need not be related to historic or forecasted near-term growth rates for earnings or dividends. Dividend payout ratios can change over time because of changes in the investment opportunity set available and the relative attractiveness of cash dividends versus stock buybacks. Since changes in dividend payout affect the dividend yield, which in turn affects g, historic growth rates may not be relevant for g. Also, if dividend policies are likely to change over time, g need not be related to g_5 (the growth rate forecast for earnings over the next five years), a rate that is frequently used to proxy for g. Various scenarios can be constructed for the two firms in the example above to obtain similar historic and/or near-term forecast growth rates and yet have substantially different values for g.

Despite the difficulties noted above, both historic and forecast rates for aggregate dividends, earnings, and other macroeconomic measures (such as GDP) have been used as proxies for g. We note that these proxies create additional error. First, it is important to hold the unit of investment constant through the period where growth is measured. In particular, any growth created at the aggregate level by the issuance/retirement of equity since the beginning of the period should be ignored. Second, profits from all activities conducted outside the publicly traded corporate sector that are included in the macroeconomic measures should be deleted, and all overseas profits relating to this sector that are excluded from some macroeconomic measures should be included.

To control for the unit of investment problem, we use forecasted growth in per-share earnings rather than aggregate earnings, and to mitigate the problems associated with identifying g, we focus on growth in rents (abnormal earnings), g_{ae}, rather than dividends. To understand the benefits of switching to g_{ae}, it is important to describe some features of abnormal earnings. Expected abnormal earnings would equal zero if book values of equity reflected market values.[24] If book values measure input costs fairly, but do not include the portion of market values that represent economic rents (not yet earned), abnormal earnings would reflect those rents. However, the magnitude of such rents at the aggregate market level is likely to be small, and any rents that emerge are likely to be dissipated over time for the usual reasons (antitrust actions, global competition, etc.). As a result, much of the

[24] That is, if market prices are efficient and book values are marked to market values each period, market (book) values are expected to adjust each period so that no future abnormal returns (abnormal earnings) are expected.

earlier literature using the abnormal earnings approach has assumed zero growth in abnormal earnings past the "horizon" date.[25]

Returning to the two-firm example, shifting the focus from growth in dividends to growth in rents removes much of the confusion caused by transitory changes in dividend payouts and dividend yields: these factors should have no impact on growth in rents, since the level of and growth in rents are determined by economic factors such as monopoly power. That is, even though the two firms have different forecasted earnings and dividends, the forecasted abnormal earnings and growth in abnormal earnings should be identical.

We believe, however, that the popular assumption of zero growth in abnormal earnings may be too pessimistic because accounting statements are conservative and understate input costs: assets (liabilities) tend to be understated (overstated) on average. For example, many investments (such as research and development, advertising, and purchased intangibles) are written off too rapidly in many domiciles. As a result, abnormal earnings tend to be positive, even in the absence of economic rents. Growth in abnormal earnings under conservative accounting is best understood by examining the behavior of the excess of *roe* (the accounting rate of return on the book value of equity) over k (the discount rate). Simulations and theoretical analyses (e.g., Zhang (2000)) of the steady-state behavior of the accounting rate of return under conservative accounting suggest two important determinants: the long-term growth in investment and the degree of accounting conservatism. These analyses also suggest that *roe* approaches k, but remains above it in the long-term.

Even though a decline in the excess of *roe* over k should cause the magnitude of abnormal earnings to fall over time, a countervailing factor is the growth in investment, which increases the base on which abnormal earnings are generated. We assume as a first approximation that the latter effect is greater than the former, and that abnormal earnings increase in perpetuity at the expected inflation rate. Since we recognize that this assumption is an approximation, we elected to err on the side of choosing too high a growth rate to ensure that our equity premium estimates are not biased downward. Also, we conduct sensitivity analyses to identify the impact on our equity premium estimates of varying the assumed growth rate within a reasonable range.

REFERENCES

Abarbanell, Jeffrey and Victor Bernard, 1999, Is the U.S. stock market myopic? *Journal of Accounting Research* 38, 221–242.

Abel, Andrew B., 1999, Risk premia and term premia in general equilibrium, *Journal of Monetary Economics* 43, 3–33.

Alford, Andrew, Jennifer Jones, Richard Leftwich, and Mark Zmijewski, 1993, The relative informativeness of accounting disclosures in different countries, *Journal of Accounting Research* 31, 183–223.

[25] That is, abnormal earnings persist, but show no growth. Some papers are even more conservative, and have assumed that abnormal earnings drop to zero past the horizon date.

Ang, Andrew and Jun Liu, 2000. A generalized affine earnings model, Working paper, Columbia University.

Bagnoli, Mark, Messod D. Beneish, and Susan G. Watts, 1999, Whisper forecasts of quarterly earnings per share, *Journal of Accounting and Economics* 28, 27–50.

Benore, Charles, 1983, A survey of investor attitudes toward the electrical power industry, Paine Webber Mitchell Hutchins, Inc., New York.

Blanchard, Oliver J., 1993, Movements in the equity premium, *Brookings Papers on Economic Activity* 2, 75–138.

Bogle, John C., 1999, *Common Sense on Mutual Funds* (John Wiley & Sons, New York).

Botosan, Christine A., 1997, Disclosure level and the cost of equity capital, *Accounting Review* 72, 323–350.

Brigham, Eugene F., Dilip K. Shome, and Steve R. Vinson, 1985, Risk premium approach to measuring a utility's cost of equity, *Financial Management* 14, 33–45.

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

Burr, Barry B., 1998, 20%+ again? Keep dreaming, *Pensions and Investments* 26, 33–34.

Campbell, John Y., 1991, A variance decomposition for stock returns, *The Economic Journal* 101 (405), 157–179.

Campbell, John Y. and Robert J. Shiller, 1988, The dividend-price ratio and expectations of future dividends and discount factors, *Review of Financial Studies* 1, 195–228.

Claus, James J. and Jacob K. Thomas, 1999a, The equity risk premium is lower than you think it is: Empirical estimates from a new approach, Working paper, Columbia University.

Claus, James J. and Jacob K. Thomas, 1999b, Measuring the equity premium using earnings forecasts: An international analysis, Working paper, Columbia University.

Cochrane, John H., 1997, Where is the market going? Uncertain facts and novel theories, *Economic Perspectives* 21, 3–37.

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

Edwards, Edgar O. and Philip W. Bell, 1961, *The Theory and Measurement of Business Income* (University of California Press, Berkeley, CA).

Fama, Eugene F. and Kenneth R. French, 2000, The equity premium, Working paper, University of Chicago.

Frankel, Richard and Charles M. C. Lee, 1998a, Accounting valuation, market expectation, and cross-sectional stock returns, *Journal of Accounting and Economics* 25, 283–319.

Frankel, Richard and Charles M. C. Lee, 1998b, Accounting diversity and international valuation, Working paper, University of Michigan.

Gebhardt, William, Charles M. C. Lee and Bhaskaran Swaminathan, Toward an ex ante cost of capital, *Journal of Accounting Research* forthcoming.

Glassman, James K. and Kevin A. Hassett, 1998, Are stocks overvalued? Not a chance, *Wall Street Journal*, March 30.

Gordon, Myron, 1962, *The investment, financing and valuation of the corporation* (Irwin, Homewood, IL).

Ibbotson Associates, Inc., 1999, Stocks, Bonds, Bills and Inflation, *Yearbook*, Chicago, IL.

Jorion, Philipe and William N. Goetzmann, 1999, Global stock markets in the twentieth century, *Journal of Finance* 54, 953–980.

Khorana, Ajay, R. Charles Moyer, and Ajay Patel, 1997, An examination of equity risk premium forecasts in G-6 countries, Working paper, Babcock Graduate School of Management, Wake Forest University.

Kocherlakota, Naryana, 1996, The equity premium: It's still a puzzle, *Journal of Economic Literature* 34, 42–71.

Liu, Jing and Jacob K. Thomas, 2000, Stock returns and accounting earnings, *Journal of Accounting Research* 38, 71–101.

Malkiel, Burton G., 1996, The equity premium puzzle, Keynote speech at 1996 Ibbotson Associates equity risk premium conference, Chicago, IL.

Matsumoto, Dawn A., 1999, Management's Incentives to Guide Analysts' Forecasts, Working paper, Harvard Business School.

Mayfield, E. Scott, 1999, Estimating the market risk premium, Working paper, Harvard Business School.

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

Moyer, R. Charles, and Ajay Patel, 1997, The equity market risk premium: a critical look at alternative ex ante estimates, Working paper, Babcock Graduate School of Management, Wake Forest University.

Myers, James N., 1999, Conservative accounting and finite firm life: Why residual income valuation estimates understate stock price, Working paper, University of Washington.

Nissim, Doron and Stephen H. Penman, 1999, Ratio analysis and equity valuation, Working paper, Columbia Business School.

Penman, Stephen H., 1999, *Financial Statement Analysis and Security Valuation* (Irwin/McGraw Hill, New York).

Preinreich, Gabriel A. D., 1938, Annual survey of economic theory: the theory of depreciation, *Econometrica* 6, 219–241.

Samuelson, Paul, 1965, Proof that properly anticipated prices fluctuate randomly, *Industrial Management Review* 6, 41–49.

Siegel, Jeremy J., 1992, The equity premium: Stock and bond returns since 1802, *Financial Analysts Journal* January/February, 28–38.

Siegel, Jeremy J., 1998, *Stocks for the long run,* 2nd edition (McGraw-Hill, New York).

Siegel, Jeremy J., and Richard H. Thaler, 1997, The equity risk premium puzzle, *Journal of Economic Perspectives* 11, 191–200.

Stulz, René M., 1999, Globalization, corporate finance, and the cost of capital, *Journal of Applied Corporate Finance* 12(3), 8–25.

Vuolteenaho, Tuomo, 1999, Understanding the aggregate book-to-market ratio, Working paper, University of Chicago.

Welch, Ivo, 1999, Views of financial economists on the equity premium and professional controversies, Working paper, UCLA.

Wien, Byron R., 1998, U.S. strategy: Risk premium—R.I.P. Morgan Stanley Dean Witter.

Williams, John. B., 1938, *The Theory of Investment Value* (Harvard University Press, Cambridge, MA).

Zhang, Xiao-jun, 2000, Conservative accounting and equity valuation, *Journal of Accounting and Economics* 29, 125–149.

Long-Run Stock Returns: Participating in the Real Economy

Roger G. Ibbotson and Peng Chen

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

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

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

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

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

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

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

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

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

Six Methods for Decomposing Returns

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

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

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

and

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

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

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

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

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

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

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

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

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

where *Rinv* is reinvestment return.

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

$$\bar{R} = [(1 + \overline{CPI})(1 + \overline{Rcg}) - 1] + \overline{Inc} + \overline{Rinv}$$
$$10.70\% = [(1 + 3.08\%) \times (1 + 3.02\%) - 1] + 4.28\% + 0.20\%.$$

The second column in Figure 1 shows the decomposition of historical equity returns for 1926–2000 according to the capital gain and income method.

Method 3. Earnings. The real-capital-gain portion of the return in the capital gain and income method can be broken into growth in real EPS, g_{REPS}, and growth in P/E, $g_{P/E}$:

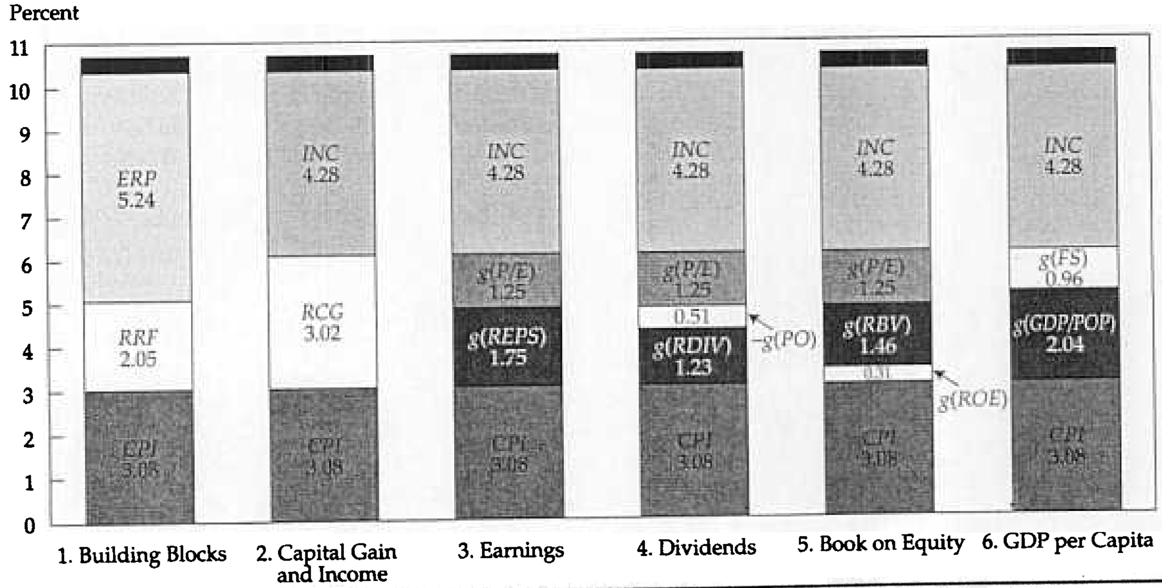
$$\begin{aligned} Rcg_t &= \frac{P_t}{P_{t-1}} - 1 \\ &= \frac{P_t/E_t}{P_{t-1}/E_{t-1}}\left(\frac{E_t}{E_{t-1}}\right) - \\ &= (1 + g_{P/E,t})(1 + g_{REPS,t}) - 1. \end{aligned} \tag{5}$$

Therefore, equity's total return can be broken into four components—inflation, growth in real EPS, growth in P/E, and income return:

$$\begin{aligned} R_t = {} & [(1 + CPI_t)(1 + g_{REPS,t})(1 + g_{P/E,t}) - 1] \\ & + Inc_t + Rinv_t. \end{aligned} \tag{6}$$

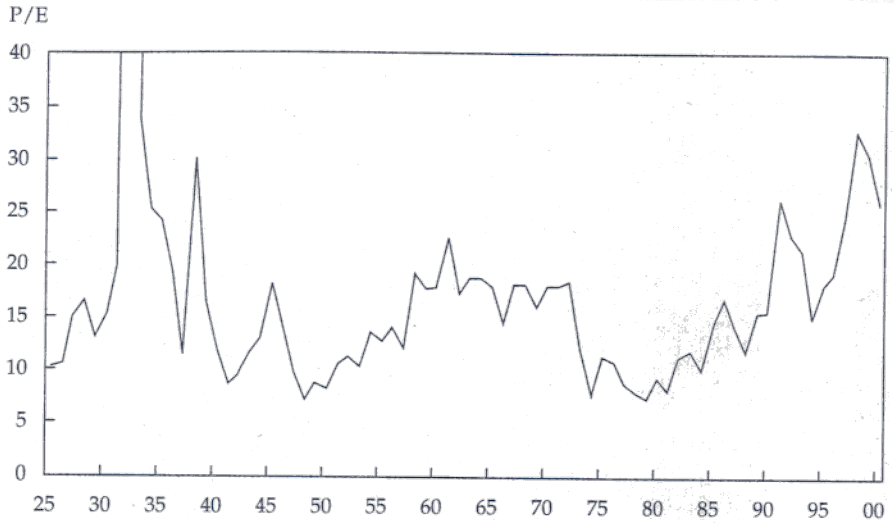
The real earnings of U.S. equity increased 1.75 percent annually between 1926 and 2000. The P/E, as **Figure 2** illustrates, was 10.22 at the beginning of 1926 and 25.96 at the end of 2000. The highest P/E (136.50 and off the chart in Figure 2) was recorded during the Great Depression, in December 1932, when earnings were near zero, and the lowest in the period (7.07) was recorded in 1948. The average year-end P/E was 13.76.[10]

Figure 1. Decomposition of Historical Equity Returns by Six Methods, 1926–2000



Notes: The block on the top of each column is the reinvestment return plus the geometric interactions among the components. Including the geometric interactions ensured that the components summed to 10.70 percent in this and subsequent figures. The table that constitutes Appendix A gives detailed information on the reinvestment and geometric interaction for all the methods.

Figure 2. P/E, 1926–2000



The U.S. equity returns from 1926 and 2000 can be computed according to the earnings method as follows:

$$\bar{R} = [(1+\overline{CPI})(1+\overline{g_{REPS}})(1+\overline{g_{P/E}})-1] + \overline{Inc} + \overline{Rinv}$$

$$10.70\% = [(1+3.08\%)\times(1+1.75\%)\times(1+1.25\%)-1] + 4.28\% + 0.20\%.$$

The third column in Figure 1 shows the decomposition of historical equity returns for 1926–2000 according to the earnings method.

Method 4. Dividends. In this method, real dividends, *RDiv*, equal the real earnings times the dividend-payout ratio, *PO*, or

$$REPS_t = \frac{RDiv_t}{PO_t}; \tag{7}$$

therefore, the growth rate of earnings can be calculated by the difference between the growth rate of real dividends, g_{RDiv}, and the growth rate of the payout ratio, g_{PO}:

$$(1+g_{REPS,t}) = \frac{(1+g_{RDiv,t})}{(1+g_{PO,t})}. \tag{8}$$

If dividend growth and payout-ratio growth are substituted for the earnings growth in Equation 6, equity total return in period t can be broken into (1) inflation, (2) the growth rate of P/E, (3) the growth rate of the dollar amount of dividends after inflation, (4) the growth rate of the payout ratio, and (5) the dividend yield:

$$R_t = \left[(1+CPI_t)(1+g_{P/E,t})\left(\frac{1+g_{RDiv,t}}{1+g_{PO,t}}\right)-1\right] + Inc_t + Rinv_t. \tag{9}$$

Figure 3 shows the annual income return (dividend yield) of U.S. equity for 1926–2000. The dividend yield dropped from 5.15 percent at the beginning of 1926 to only 1.10 percent at the end of 2000. **Figure 4** shows the year-end dividend-payout ratio for 1926–2000. On average, the dollar amount of dividends after inflation grew 1.23 percent a year, while the dividend-payout ratio decreased 0.51 percent a year. The dividend-payout ratio was 46.68 percent at the beginning of 1926. It had decreased to 31.78 percent at the end of 2000. The highest dividend-payout ratio was recorded in 1932, and the lowest was the 31.78 percent recorded in 2000.

The U.S. equity returns from 1926 through 2000 can be computed in the dividends method according to

$$\bar{R} = \left[(1+\overline{CPI})(1+\overline{g_{P/E}})\left(\frac{1+\overline{g_{RDiv}}}{1+\overline{g_{PO}}}\right)-1\right] + \overline{Inc} + \overline{Rinv}$$

$$10.70\% = \left[(1+3.08\%)\times(1+1.25\%)\times\left(\frac{1+1.23\%}{1-0.51\%}\right)-1\right] + 4.28\% + 0.20\%.$$

The decomposition of equity return according to the dividends method is given in the fourth column of Figure 1.

Method 5. Return on Book Equity. Earnings can be broken into the book value of equity, *BV*, and return on the book value of equity, *ROE*:

$$EPS_t = BV_t(ROE_t). \tag{10}$$

The growth rate of earnings can be calculated from the combined growth rates of real book value, g_{RBV}, and of *ROE*:

$$1+g_{REPS,t} = (1+g_{RBV,t})(1+g_{ROE,t}). \tag{11}$$

Figure 3. Income Return (Dividend Yield), 1926–2000

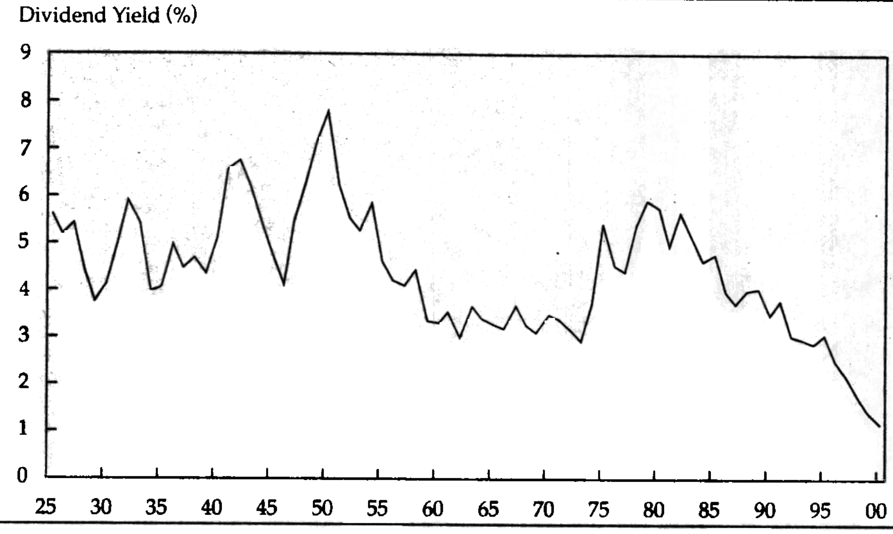
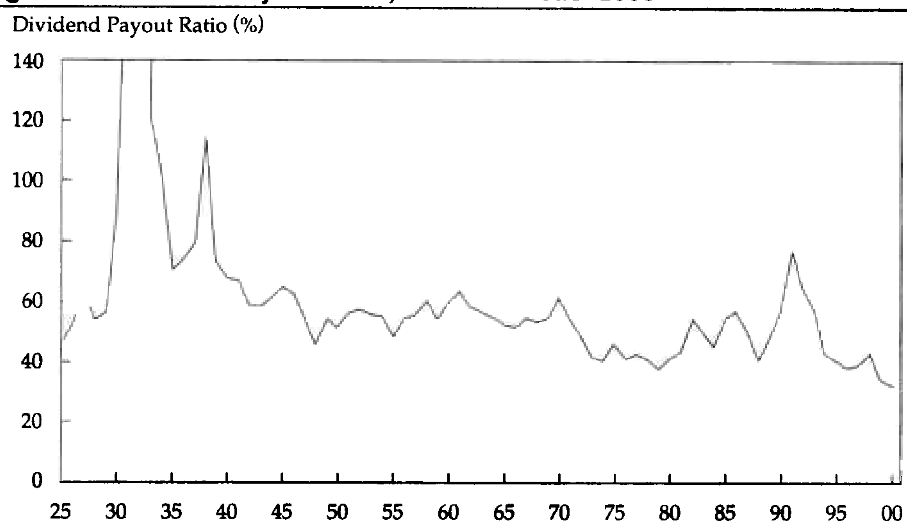


Figure 4. Dividend-Payout Ratio, Year-End 1926–2000



Note: The dividend-payout ratio was 190.52 percent in December 1931 and 929.12 percent in December 1932.

In this method, *BV* growth and *ROE* growth are substituted for earnings growth in the equity return decomposition, as shown in the fifth column of Figure 1. Then, equity's total return in period t can be computed by

$$R_t = [(1 + CPI_t)(1 + g_{P/E,t})(1 + g_{RBV,t})(1 + g_{ROE,t}) - 1] + Inc_t + Rinv_t. \quad (12)$$

We estimated that the average growth rate of the book value after inflation was 1.46 percent for 1926–2000.[11] The average *ROE* growth a year during the same time period was calculated to be 0.31 percent:

$$\bar{R} = [(1 + \overline{CPI})(1 + \overline{g_{P/E}})(1 + \overline{g_{BV}})(1 + \overline{g_{ROE}}) - 1] + \overline{Inc} + \overline{Rinv}$$

$$10.70\% = [(1 + 3.08\%)(1 + 1.25\%)(1 + 1.46\%)(1 + 0.31\%) - 1] + 4.28\% + 0.20\%.$$

Method 6. GDP per Capita. Diermeier et al. proposed a framework to analyze the aggregate supply of financial asset returns. Because we were interested only in the supply model of the equity returns in this study, we developed a slightly different supply model based on the growth of economic productivity. In this method, the market return over the long run is decomposed into (1) inflation, (2) the real growth rate of overall economic productivity (GDP per capita, $g_{GDP/POP}$), (3) the increase in the equity market relative to overall economic productivity (the increase in the factor share of equities in the overall economy, g_{FS}), and (4) dividend yields.[12] This model is expressed by the following equation:

$$R_t = [(1 + CPI_t)(1 + g_{GPD/POP,t})(1 + g_{FS,t}) - 1] + Inc_t + Rinv_t. \quad (13)$$

Figure 5 shows the growth of the U.S. stock market, GDP per capita, earnings, and dividends initialized to unity ($1.00) at the end of 1925. The level of all four factors dropped significantly in the early 1930s. For the whole period, GDP per capita slightly outgrew earnings and dividends, but all four factors grew at approximately the same rate. In other words, overall economic productivity increased slightly faster than corporate earnings or dividends over the past 75 years. Although GDP per capita outgrew earnings and dividends, the overall stock market price grew faster than GDP per capita. The primary reason is that the market P/E increased 2.54 times during the same time period.

Average equity market return can be calculated according to this model as follows:

$$\begin{aligned}\bar{R} &= [(1 + \overline{CPI})(1 + \overline{g_{GDP/POP}})(1 + \overline{g_{FS}}) - 1] \\ &\quad + \overline{Inc} + \overline{Rinv} \\ 10.70\% &= [(1 + 3.08\%)(1 + 2.04\%)(1 + 0.96\%) - 1] \\ &\quad + 4.28\% + 0.20\%.\end{aligned}$$

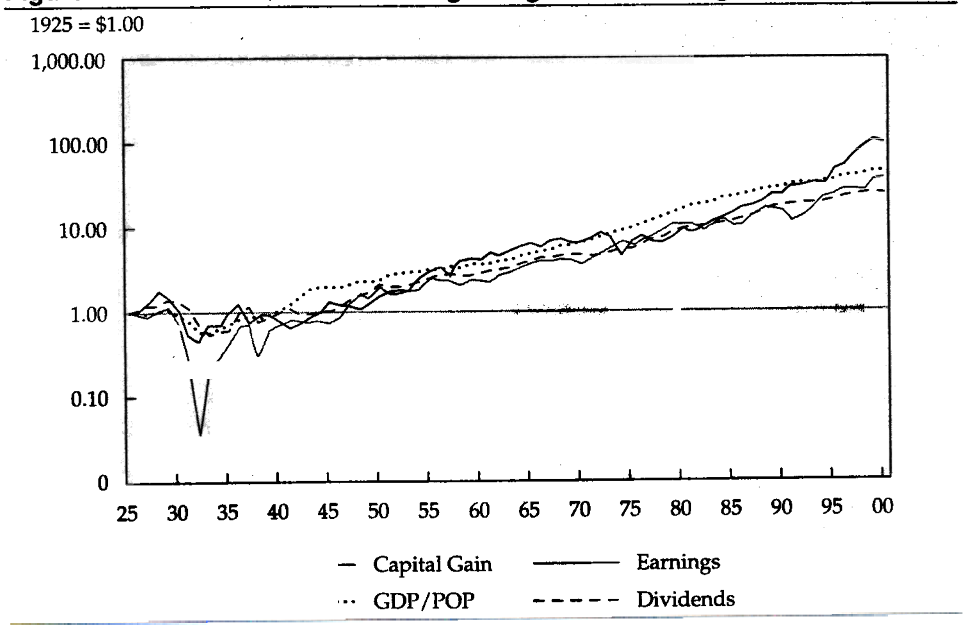
We calculated the average annual increase in the factor share of the equity market relative to the overall economy to be 0.96 percent. The increase in this factor share is less than the annual increase of the P/E (1.25 percent) over the same time period. This finding suggests that the increase in the equity market share relative to the overall economy can be fully attributed to the increase in its P/E.

The decomposition of historical equity returns by the GDP per capita model is given in the last column of Figure 1.

Summary of Equity Returns and Components. The decomposition of the six models into their components can be compared by looking at Figure 1. The differences among the five models arise from the different components that represent the capital gain portion of the equity returns.

This analysis produced several important findings. First, as Figure 5 shows, the growth in corporate earnings has been in line with the growth of overall economic productivity. Second, P/E increases accounted for only 1.25 pps of the 10.70 percent total equity return. Most of the return has been attributable to dividend payments and nominal earnings growth (including inflation and real earnings growth). Third, the increase in the relative factor share of equity can be fully attributed to the increase in P/E. Overall, economic productivity outgrew both corporate earnings and dividends from 1926 through 2000. Fourth, despite the record earnings growth in the 1990s, the dividend yield and the payout ratio declined sharply, which renders dividends alone a poor measure for corporate profitability and future earnings growth.

Figure 5. Growth of $1 from the Beginning of 1926 through 2000



Long-Term Forecast of Equity Returns

Supply-side models can be used to forecast the long-term expected equity return. The supply of stock market returns is generated by the productivity of the corporations in the real economy. Over the long run, the equity return should be close to the long-run supply estimate. In other words, investors should not expect a much higher or a much lower return than that produced by the companies in the real economy. Therefore, we believe investors' expectations for long-term equity performance should be based on the supply of equity returns produced by corporations.

The supply of equity returns consists of two main components—current returns in the form of dividends and long-term productivity growth in the form of capital gains. In this section, we focus on two of the supply-side models—the earnings model and the dividends model (Methods 3 and 4).[13] We studied the components of these two models by identifying which components are tied to the supply of equity returns and which components are not. Then, we estimated the long-term, sustainable return based on historical information about these supply components.

Model 3F. Forward-Looking Earnings. According to the earnings model (Equation 6), the historical equity return can be broken into four components—the income return, inflation, the growth in real EPS, and the growth in P/E. Only the first three of these components are historically supplied by companies. The growth in P/E reflects investors' changing predictions of future earnings growth. Although we forecasted that the past supply of corporate growth will continue, we did not forecast any change in investor predictions. Thus, the supply side of equity return, *SR*, includes only inflation, the growth in real EPS, and income return:[14]

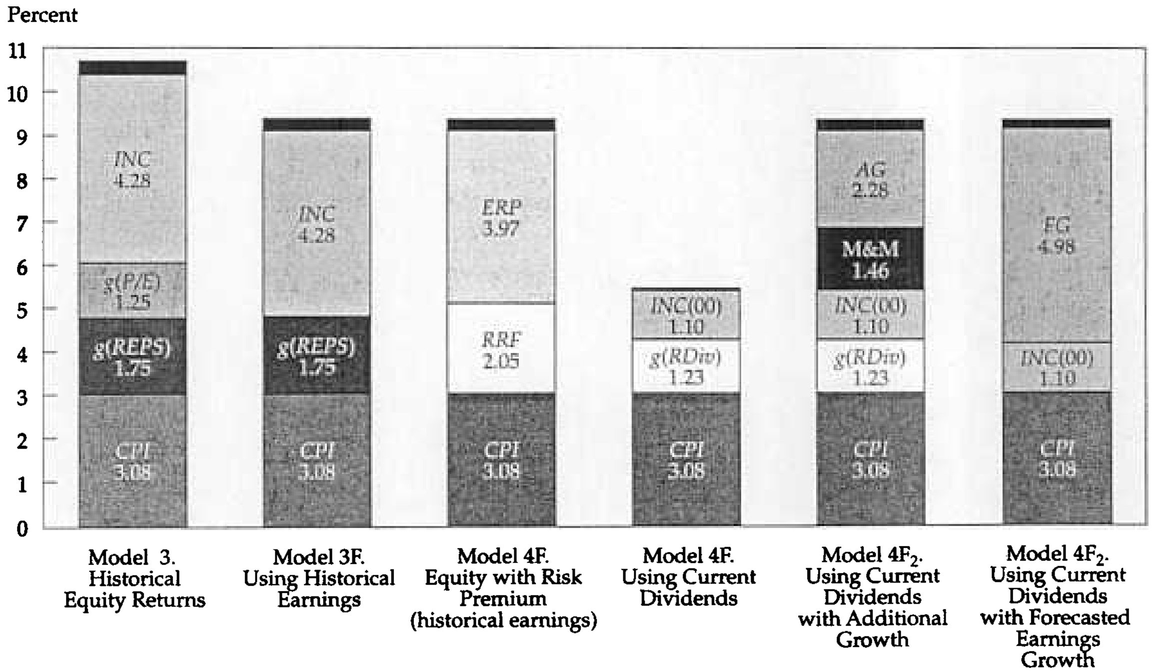
$$SR_t = [(1+CPI_t)(1+g_{REPS,t})-1]+Inc_t+Rinv_t. \quad (14)$$

The long-term supply of U.S. equity returns based on the earnings model is 9.37 percent, calculated as follows:

$$\overline{SR} = [(1+\overline{CPI})(1+\overline{g_{REPS}})-1]+\overline{Inc}+\overline{Rinv}$$
$$9.37\% = [(1+3.08\%)(1+1.75\%)-1]+4.28\%+0.20\%.$$

The decomposition according to Model 3F is compared with that of Method 3 (based on historical data plus the estimated equity risk premium) in the first two columns of **Figure 6**.

Figure 6. Historical vs. Current Dividend-Yield Forecasts Based on Earnings and Dividends Models



Notes: Inc(00) is the dividend yield in year 2000. *FG* is the real earnings growth rate, forecasted to be 4.98 percent. Model $4F_2$ corrects Model 4F as follows: add 1.46 pps for M&M consistency and add 2.24 pps for the additional growth, *AG*, implied by the high current market P/E.

The supply-side equity risk premium, *ERP*, based on the earnings model is calculated to be 3.97 pps:

$$\begin{aligned}\overline{ERP} &= \frac{(1+\overline{SR})}{(1+\overline{CPI})(1+\overline{RRf})} - 1 \\ &= \frac{1+9.37\%}{(1+3.08\%)(1+2.05\%)} - 1 \\ &= 3.97\%.\end{aligned}$$

The *ERP* is taken into account in the third column of Figure 6.

Model 4F. Forward-Looking Dividends. The forward-looking dividends model is also referred to as the constant-dividend-growth model (or the Gordon model). In it, the expected equity return equals the dividend yield plus the expected dividend growth rate. The supply of the equity return in the Gordon model includes inflation, the growth in real dividends, and dividend yield.

As is commonly done with the constant-dividend-growth model, we used the current dividend yield of 1.10 percent instead of the historical dividend yield of 4.28 percent. This decision reduced the estimate of the supply of equity returns to 5.44 percent:

$$\begin{aligned}\overline{SR} &= [(1+\overline{CPI})(1+\overline{g_{RDiv}})-1] + Inc(00) + \overline{Rinv} \\ 5.54\% &= [(1+3.08\%)(1+1.23\%)-1] + 1.10\% + 0.20\%,\end{aligned}$$

where *Inc*(00) is the dividend yield in year 2000. The equity risk premium was estimated to be 0.24 pps:

$$\begin{aligned}\overline{ERP} &= \frac{(1+\overline{SR})}{[illegible]} \\ &= \frac{[illegible]}{(1+3.08\%)+(1+2.05\%)} \\ &= 0.24\%.\end{aligned}$$

Figure 6 allows a comparison of forecasted equity returns including the equity risk premium estimates based on the earnings model and the dividends model. In the next section, we show why we disagree with the dividends model and prefer to use the earnings model to estimate the supply-side equity risk premium.

Differences between the Earnings Model and the Dividends Model. The earnings model (3F) and the dividends model (4F) differ in essentially two ways. The differences relate to the low current payout ratio and the high current P/E. These two differences are reconciled in what we will call Model $4F_2$ shown in the two right-hand columns of Figure 6. First, to reflect growth in productivity, the earnings model uses historical earnings growth whereas the dividend model uses historical dividend growth. Historical dividend growth underestimates historical earnings growth, however, because of the decrease in the payout ratio. Overall, the dividend growth underestimated the increase in earnings productivity by 0.51 pps a year for 1926–2000. Today's low dividend yield also reflects the current payout ratio, which is at a historical low of 31.8 percent (compared with the historical average of 59.2 percent). Applying such a low rate to the future would mean that even more earnings would be retained in the future than in the historical period studied. But had more earnings been retained, the historical earnings growth would have been 0.95 pps a year higher, so (assuming the historical average dividend-payout ratio) the current yield of 1.10 percent would need to be adjusted upward by 0.95 pps.

By using the current dividend-payout ratio in the dividend model, Model 4F creates two errors, both of which violate Miller and Modigliani theory. A company's dividend-payout ratio affects only the form in which shareholders receive their returns (i.e., dividends versus capital gains), not their total returns. The *current* low dividend-payout ratio should not affect our forecast. Companies today probably have such low payout ratios to reduce the tax burden on their investors. Instead of paying dividends, many companies reinvest earnings, buy back shares, or use the cash to purchase other companies.[15] Therefore, the dividend growth model has to be upwardly adjusted by 1.46 pps (0.51 pp plus 0.95 pp) so as not to violate M&M theory.

The second difference between Model 3F and Model 4F is related to the fact that the current P/E (25.96) is much higher than the historical average (13.76). The current yield (1.10 percent) is at a historic low—because of the previously mentioned low payout ratio and because of the high P/E. Even assuming the historical average payout ratio, the current dividend yield would be much lower than its historical average (2.05 percent versus 4.28 percent). This difference is geometrically estimated to be 2.28 pps a year. In Figure 6, the additional growth, *AG*, accounts for 2.28 pps of the return; in the last column, the forecasted real earnings growth rate, *FG*, accounts for 4.98 pps. The high P/E could be caused by (1) mispricing, (2) a low required rate of return, and/or (3) a high expected future earnings growth rate. Mispricing as a cause is eliminated by our assumption of market efficiency, and a low required rate of return is eliminated by our assumption of a constant equity risk premium through the past and future periods that we are trying to estimate. Thus, we interpret the high P/E as the market expectation of higher earnings growth and the following equation is the model for

Model $4F_2$, which reconciles the differences between the earnings model and the dividends model:[16]

$$\overline{SR} = [(1+\overline{CPI})(1+\overline{g_{RDiv}})(1-\overline{g_{PO}})-1] + Inc(00) + AY + AG + \overline{Rinv}$$

$$9.67\% = [(1+3.08\%)(1+1.23\%)(1+0.51\%)-1] + 1.10\% + 0.95\% + 2.28\% + 0.20\%.$$

To summarize, the earnings model and the dividends model have three differences. The first two differences relate to the dividend-payout ratio and are direct violations of M&M. The third difference results from the expectation of higher-than-average earnings growth, which is predicted by the high current P/E. Reconciling these differences reconciles the earnings and dividends models.

Geometric vs. Arithmetic. The estimated equity return (9.37 percent) and equity risk premium (3.97 pps) are geometric averages. The arithmetic average, however, is often used in portfolio optimization. One way to convert the geometric average into an arithmetic average is to assume the returns are independently lognormally distributed over time. Then, the arithmetic average, R_A, and geometric average, R_G, have roughly the following relationship:

$$R_A = R_G + \frac{\sigma^2}{2}, \tag{15}$$

where σ^2 is the variance.

The standard deviation of equity returns is 19.67 percent. Because almost all the variation in equity returns is from the equity risk premium, rather than the risk-free rate, we need to add 1.93 pps to the geometric estimate of the equity risk premium to convert the returns into arithmetic form, so $R_A = R_G + 1.93$ pps. The arithmetic average equity risk premium then becomes 5.90 pps for the earnings model.

To summarize, the long-term supply of equity return is estimated to be 9.37 percent (6.09 percent after inflation), conditional on the historical average risk-free rate. The supply-side equity risk premium is estimated to be 3.97 pps geometrically and 5.90 pps arithmetically.[17]

Conclusions

We adopted a supply-side approach to estimate the forward-looking, long-term, sustainable equity return and equity risk premium. We analyzed historical equity returns by decomposing returns into factors commonly used to describe the aggregate equity market and overall economic productivity—inflation, earnings, dividends, P/E, the dividend-payout ratio, BV, ROE, and GDP per capita. We examined each factor and its relationship to the long-term supply-side framework. We used historical information in our supply-side models to forecast the equity risk premium. A complete tabulation of all the numbers from all models and methods is presented in Appendix A.

Contrary to several recent studies on the equity risk premium declaring the forward-looking premium to be close to zero or negative, we found

Appendix A. Summary Tabulations for Forecasted Equity Return

Method/Model	Sum	Inflation	Real Risk-Free Rate	Equity Risk Premium	Real Capital Gain	g(Real *EPS*)	g(Real *Div*)	–g(Payout Ratio)
A. Historical								
Method 1	10.70	3.08	2.05	5.24				
Method 2	10.70	3.08			3.02			
Method 3	10.70	3.08				1.75		
Method 4	10.70	3.08					1.23	0.51
Method 5	10.70	3.08						
Method 6	10.70	3.08						
B. Forecast with historical dividend yield								
Model 3F	9.37	3.08				1.75		
Model 3F (*ERP*)	9.37	3.08	2.05	3.97				
C. Forecast with current dividend yield								
Model 4F	5.44	3.08					1.23	
Model 4F (*ERP*)	5.44	3.08	2.05	0.24				
Model $4F_2$	9.37	3.08					1.23	0.51
Model $4F_2$ (*FG*)	9.37	3.08						

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

the long-term supply of the equity risk premium to be only slightly lower than the straight historical estimate. We estimated the equity risk premium to be 3.97 pps in geometric terms and 5.90 pps on an arithmetic basis. These estimates are about 1.25 pps lower than the historical estimates. The differences between our estimates and the ones provided by several other recent studies result principally from the inappropriate assumptions those authors used, which violate the M&M theorem. Also, our models interpret the current high P/E as the market forecasting high future growth rather than a low discount rate or an overvaluation. Our estimate is in line with both the historical supply measures of public corporations (i.e., earnings) and overall economic productivity (GDP per capita).

The implication of an estimated equity risk premium being far closer to the historical premium than zero or negative is that stocks are expected to outperform bonds over the long run. For long-term investors, such as pension funds and individuals saving for retirement, stocks should continue to be a favored asset class in a diversified portfolio. Because our estimate of the equity risk premium is lower than historical performance, however, some investors should lower their equity allocations and/or increase their savings rate to meet future liabilities.

Notes

1. In our study, we defined the equity risk premium as the difference between the long-run expected return on stocks and the long-term risk-free (U.S. Treasury) yield. [Some other studies, including Ibbotson and Sinquefield (1976a, 1976b) used short-term U.S. T-bills as the risk-free rate.] We did all of our analysis in geometric form, then converted to arithmetic data at the end, so the estimate is expressed in both arithmetic and geometric forms.
2. See also Mehra (2003).
3. Comparing estimates from one study with another is sometimes difficult because of changing points of reference. The equity risk premium estimate can be significantly different simply because the authors used arithmetic versus geometric returns, a long-term risk-free rate versus a short-term risk-free rate, bond income return (yield) versus bond total return, or long-term strategic forecasting versus short-term market-timing estimates. We provide a detailed discussion of arithmetic versus geometric returns in the section "The Long-Term Forecast."
4. Welch's survey reported a 7 pp equity risk premium measured as the arithmetic difference between equity and T-bill returns. To make an apples-to-apples comparison, we converted the 7 pp number into a geometric equity risk premium relative to the long-term U.S. government bond income return, which produced an estimate of almost 4 pps.
5. For further discussion of approaches to estimating the equity risk premium, see the presentations and discussions at www.aimrpubs.org/ap/home.html from AIMR's *Equity Risk Premium Forum*.
6. Each per share quantity is per share of the S&P 500 portfolio. Hereafter, we will merely refer to each factor without always mentioning "per share"—for example, "dividends" instead of "dividends per share."
7. Many theoretical models suggest that the equity risk premium is dynamic over time. Recent empirical studies (e.g., Goyal and Welch 2001; Ang and Bekaert 2001) found no evidence, however, of long-horizon return predictability by using either earnings or dividend yields. Therefore, instead

g(*BV*)	*g*(*ROE*)	*g*(*P/E*)	*g*(*Real GDP/ POP*)	*g*(*FS-GDP/POP*)	Income Return	Reinvestment + Interaction	Additional Growth	Forecasted Earnings Growth
						0.33		
					4.28	0.32		
		1.25			4.28	0.34		
		1.25			4.28	0.35		
1.25	0.31	1.25			4.28	0.31		
			2.04	0.96	4.28	0.32		
					4.28	0.26		
						0.27		
					1.10[a]	0.03		
						0.07		
					2.05[b]	0.21	2.28	
					1.10[a]	0.21		4.98

of trying to build a model for a dynamic equity risk premium, we assumed that the long-term equity risk premium is constant. This assumption provided a benchmark for analysis and discussion.

8. We updated the series with data from Standard and Poor's to include the year 2000.
9. Appendix A summarizes all the tabulations we discuss.
10. The average P/E was calculated by reversing the average earnings-to-price ratio for 1926–2000.
11. Book values were calculated from the book-to-market ratios reported in Vuolenteenaho (2000). The aggregate book-to-market ratio was 2.0 in 1928 and 4.1 in 1999. We used the growth rate in book value calculated for 1928–1999 as the proxy for the growth rate for 1926–2000. The average ROE growth rate was calculated from the derived book value and the earnings data.
12. Instead of assuming a constant equity factor share, we examined the historical growth rate of the equity factor share relative to the overall growth of the economy.
13. We did not use Methods 1, 2, and 5 in forecasting because the forecasts of Methods 1 and 2 would be identical to the historical estimate reported in the previous section and because the forecast of Method 5 would require more complete BV and ROE data than we currently have available. We did use Method 6 to forecast future stock returns but found the results to be very similar to those for the earnings model; therefore, we do not report the results here.
14. This model uses historical income return as an input for reasons that are discussed in the section "Differences between the Earnings Model and the Dividends Model."
15. The current tax code provides incentives for companies to distribute cash through share repurchases rather than through dividends. Green and Hollifield (2001) found that the tax savings through repurchases are on the order of 40–50 percent of the taxes that investors would have paid if dividends were distributed.
16. Contrary to efficient market models, Shiller (2000) and Campbell and Shiller argued that the P/E appears to forecast future stock price change.
17. We could also use the GDP per capita model to estimate the long-term equity risk premium. This model implies long-run stock returns should be in line with the productivity of the overall economy. The equity risk premium estimated by using the GDP per capita model would be slightly higher than the *ERP* estimate from the earnings model because GDP per capita grew slightly faster than corporate earnings in the study period. A similar approach can be found in Diermeier et al., who proposed using the growth rate of the overall economy as a proxy for the growth rate in aggregate wealth in the long run.

References

Ang, Andrew, and Geert Bekaert. 2001. "Stock Return Predictability: Is It There?" National Bureau of Economic Research (NBER) Working Paper 8207 (April).

Arnott, Robert D., and Clifford S. Asness. 2003. "Surprise! Higher Dividends = Higher Earnings Growth." *Financial Analysts Journal*, vol. 59, no. 1 (January/February):70–87.

Arnott, Robert D., and Peter L. Bernstein. 2002. "What Risk Premium Is 'Normal'?" *Financial Analysts Journal*, vol. 58, no. 2 (March/April):64–84.

Arnott, Robert D., and Ronald Ryan. 2001. "The Death of the Risk Premium: Consequences of the 1990s." *Journal of Portfolio Management*, vol. 27, no. 3 (Spring):61–74.

Campbell, John Y., and Robert J. Shiller. 2001. "Valuation Ratios and the Long-Run Stock Market Outlook: An Update." *NBER* Working Paper No. 8221.

Diermeier, Jeffrey J., Roger G. Ibbotson, and Laurance B. Siegel. 1984. "The Supply for Capital Market Returns." *Financial Analysts Journal*, vol. 40, no. 2 (March/April):2–8.

Fama, Eugene F., and Kenneth R. French. 2001. "Disappearing Dividends: Changing Firm Characteristics or Lower Propensity to Pay?" *Journal of Financial Economics*, vol. 60, no. 1 (April):3–43.

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

Gordon, Myron. 1962. *Investment, Financing, and Valuation of the Corporation*. Homewood, IL: Irwin.

Goyal, Amit, and Ivo Welch. 2001. "Predicting the Equity Premium with Dividend Ratios." Working paper. Yale School of Management and UCLA.

Graham, John R., and Campbell R. Harvey. 2001. "Expectations of Equity Risk Premia, Volatility and Asymmetry from a Corporate Finance Perspective." Working paper, Fuqua School of Business, Duke University (August).

Green, Richard C., and Burton Hollifield. 2001. "The Personal-Tax Advantages of Equity." Working paper, Carnegie Mellon University (January).

Ibbotson Associates. 2001. *Stocks, Bonds, Bills, and Inflation: 2001 Yearbook*. Chicago, IL: Ibbotson Associates.

Ibbotson, Roger G., and Rex A. Sinquefield. 1976a. "Stocks, Bonds, Bills, and Inflation: Year-by-Year Historical Returns (1926–1974)." *Journal of Business*, vol. 49, no. 1 (January):11–47.

———. 1976b. "Stocks, Bonds, Bills, and Inflation: Simulations of the Future (1976–2000)." *Journal of Business*, vol. 49, no. 3 (July):313–338.

Ibbotson, Roger G., Jeffrey J. Diermeier, and Laurance B. Siegel. 1984. "The Demand for Capital Market Returns: A New Equilibrium Theory." *Financial Analysts Journal*, vol. 40, no. 1 (January/February):22–33.

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

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

Miller, Merton, and Franco Modigliani. 1961. "Dividend Policy, Growth and the Valuation of Shares." *Journal of Business*, vol. 34, no. 4 (October):411–433.

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

Siegel, Jeremy J. 1999. "The Shrinking Equity Risk Premium." *Journal of Portfolio Management*, vol. 26, no. 1 (Fall):10–17.

Vuolenteenaho, Tuomo. 2000. "Understanding the Aggregate Book-to-Market Ratio and Its Implications to Current Equity-Premium Expectations." Working paper, Harvard University.

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

Wilson, Jack W., and Charles P. Jones. 2002. "An Analysis of the S&P 500 Index and Cowles' Extensions: Price Indexes and Stock Returns, 1870–1999." *Journal of Business*, vol. 75, no. 3 (July):505–535.