

THE EQUITY PREMIUM PUZZLE

Rajnish Mehra

University of California, Santa Barbara

and

University of Chicago

April 2001

To appear as a chapter in *Mastering Investments* published by Prentice Hall 2002

Historical data provides us with a wealth of evidence that for over a century, stock returns have been considerably higher than those for Treasury bills. The average annual real return (that is to say, the inflation-adjusted return) on the US stock market over the last 110 years has been about 7.9 per cent. Over the same period, the return on a relatively riskless security was a paltry 1 per cent. The difference between these two returns - 6.9 per cent - is termed the equity premium. This statistical difference is even more pronounced over the post-war period, with the premium of stock returns over bills being almost 8 per cent.

Further, this pattern of excess returns to equity holdings is not unique to US capital markets. Equity returns compared to the return to debt holdings in other countries also exhibit this historical regularity. The annual return on the British stock market was 5.7 per cent over the post war period, an impressive 4.6 per cent premium over the average bond return of 1.1 per cent. Similar statistical differentials are documented for France, Germany, Italy and Spain.

The finance expert Jeremy Siegel has analysed data on US stock and bond returns going back to 1802 and found a similar though somewhat smaller premium in place for the past 200 years. Table 1 summarises the foregoing data.

	% real return on a market index	% real return on a relatively riskless security	% risk premium
Time period	Mean	Mean	Mean
1802-1998	7.0	2.9	4.1
1889-1978	6.98	0.80	6.18
1889-2000	7.9	1.0	6.9
1926-2000	8.7	0.7	8.0
1947-2000	8.4	0.6	7.8
Source: 1802-1998 from Siegel (1998), 1889-1978 from Mehra & Prescott (1985) and the rest are the author's estimates.			

The dramatic investment implications of these differential rate of return can be seen in Table 2, which maps the capital appreciation of \$1 invested in different assets from 1802-1997 and from 1925-2000.

Investment period	Terminal value of \$1 invested in			
	Stocks		Bills	
	Real	Nominal	Real	Nominal
1802-1997	\$558,945	\$7,470,000	\$276	\$3,679
1925-2000	\$266.47	\$2,586.52	\$1.71	\$16.56
Source: Ibbotson (2001) and Siegel (1998)				

As the table illustrates, \$1 invested in a diversified stock index yields an ending wealth of \$558,945 versus a value of \$266, *in real terms*, for \$1 invested in a portfolio of Treasury bills for the period 1802-1997. While for the 75-year period 1925-2000 the corresponding values are \$266.47 and \$1.71. (This table assumes that all payments to the underlying asset, such as dividend payments to stock and interest payments to bonds, are reinvested and no taxes are paid.)

This long-term perspective underscores the remarkable wealth-building potential of the equity premium. It should come as no surprise that the equity premium is of central importance in portfolio allocation decisions and estimates of the cost of capital, and is center in the current debate about the advantages of investing Social Security funds in the stock market. In putting together an investment portfolio investors choose among different assets such as stock, real estate or corporate bonds. Typically an investor chooses the portfolio that is right for him by assessing the relative risk characteristics and expected return of various alternatives. The risk premium spells out the differential return that an investor can expect different assets to yield.

Are stocks riskier?

Why have stocks been such an attractive investment relative to bonds? Why has the rate

of return on stocks been higher than on relatively risk free assets? One intuitive answer is that since stocks are "riskier" than bonds, investors require a larger premium or reward for bearing this additional risk. Indeed, the standard deviation of the returns to stocks (about 20 per cent per annum historically) is larger than that of the returns to Treasury bills (about 4 per cent per annum), so, obviously they are considerably more risky than bills. But are they?

To deepen our understanding of the trade-off between risk and return in the pricing of financial assets, let us look at why different assets give different rates of return. In asset pricing theory, assets are priced such that the "incremental loss of well-being" incurred by sacrificing current consumption and buying an asset at a certain price is equal to the expected "incremental gain in well-being" contingent on the increase in consumption that is anticipated when the asset pays off in the future.

The operative phrase here is the "incremental loss or gain of well-being due to consumption", which is different from "incremental consumption". In other words, the *same* amount of consumption may result in different degrees of well-being at different times. (A five-course dinner after a heavy lunch yields considerably less satisfaction than a similar dinner when one is hungry.)

Hence assets that pay off when times are good and consumption levels are high - that is, when the incremental value of additional consumption is low - are less desirable than those that pay off an equivalent amount when times are bad and additional consumption is more desirable and more highly valued.

Let us illustrate this issue by considering the standard popular model of modern finance, the Capital Asset Pricing Model (CAPM) as an example. The model postulates a linear relationship between an asset's "beta", a measure of its risk, and expected return. Thus, high beta stocks yield a high expected rate of return. In the CAPM, good times and bad times are captured by the return on the market. The performance of the market as captured by a broad-based index acts as a surrogate indicator for the relevant state of the

economy.

A high beta stock tends to pay off more when the market return is high, that is, when times are good and consumption is plentiful. As discussed earlier, such a stock provides less incremental well-being than a security that pays off when consumption is low, and hence is less valuable and consequently sells for less. To use the jargon of modern asset pricing theory, assets that pay off in states of low marginal utility will sell for a lower price than a similar asset that pays off in states of high marginal utility. Since rates of return are inversely proportional to asset prices, the latter class of assets will, on average, give a lower rate of return than the former.

Another way of looking at asset pricing is to realise that economic agents prefer to smooth out patterns of consumption over time. Assets that pay off a relatively larger amount at times when consumption is already high, “destabilise” these patterns of consumption, whereas assets that pay off when consumption levels are low “smooth” out consumption. Naturally, the latter are more valuable and thus require a lower rate of return to induce investors to hold these assets. (Insurance policies are a classic example of assets that smooth consumption. Individuals willingly purchase and hold them, in spite of their very low rates of return.)

To return to the original question: are stocks that much more riskier than bills so as to justify a 7 per cent differential in their rates of return?

The puzzle

Stocks and bonds pay off in approximately the same states of nature or economic scenarios where consumption is similar. Hence, as argued earlier, they should command approximately the same rate of return. In fact, in a 1979 paper by this author and Edward Prescott, we showed that stocks on average should command at most a 1 per cent return premium over Treasury bonds. Since, for as long as we had reliable data (about 100 years), the mean premium on stocks over bonds was considerably and consistently

higher, we realised that we had a puzzle on our hands.

It should be stressed that the equity premium puzzle is a puzzle of degree rather than quality. Standard theory is consistent with our notion of risk that, on average, stocks should return more than bonds. The puzzle arises from the fact that the returns predicted by the theory are very different from those that have been historically documented.

The puzzle cannot be dismissed lightly, since much of our economic intuition is based on the very class of models that fall short so dramatically when confronted with financial data. It shows that concepts central to financial and economic modelling fail to capture the single characteristic that appears to make stocks comparatively so risky. It also questions the viability of using this class of models for a quantitative assessment, say, to gauge the welfare implications of alternative stabilisation policies since the costs and benefits associated with these policies are now suspect.

For these reasons, finance specialists and economists have made repeated attempts to solve the puzzle over the past 15 years or so. The majority of their research efforts fall into two camps. Some have proposed modifications of the utility functions that are typically used to model investors as highly risk averse. Others have proposed explanations based on market imperfections, transactions costs, potential disaster states, selection bias, and the inability to insure against risk and disaster scenarios.

Two recent approaches to resolving the puzzle appear promising. The first, expounded by academics John Campbell and John Cochrane, incorporates the possibility of economic recession as a variable in the calculations. In this approach, the risk aversion of investors rises dramatically when the chances of a recession become larger.

The second approach, proposed by George Constantinides, John Donaldson and this author, incorporates consumer heterogeneity and departs from the representative agent model. Equity thus is no longer a homogenous asset; it has a very different investment characteristic for the middle-aged cohort who typically have stable employment and

salaries than for a young person who typically has uncertain future wage income.

In this model, an economy consists of three overlapping generations: the young, the middle-aged and the old. The consumption and investment decisions of each generation affect the demand for and thus the prices of assets in the economy. The young are restricted from participating in the stock market because they face a binding borrowing constraint.

Hence stocks are priced not by the young, for whom it is an attractive asset, but by the middle-aged, for whom stocks tend to be far less attractive. For the young, stocks and wages are not highly correlated and thus stocks are a potential hedge against wage fluctuations. The middle-aged, however, have no significant future wages, and thus the fluctuation in their consumption arises directly from fluctuations in the value of their holdings of stocks. At this stage of the life cycle, equity income is highly correlated with consumption, and it no longer helps to diversify the effects of risky human capital. Hence, for middle-aged and older persons to hold equity, it must offer a higher rate of return. That is, in equilibrium it must command a large premium over safe securities. This new approach that has yielded interesting insights into the discussion on optimal portfolio allocation incorporates a life-cycle approach to asset pricing. We elaborate on the Life-cycle perspective in a companion article in this series.

Another form of heterogeneity, proposed by George Constantinides and Darrell Duffie, captures the notion that consumers are subject to idiosyncratic income shocks that cannot be insured away. Simply put, consumers face the risk of job loss, or other major personal disasters that cannot be hedged away or insured against.

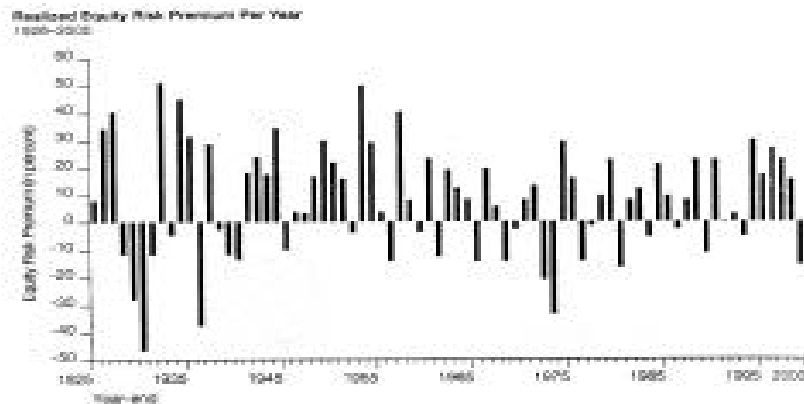
Thus, equities and cyclically related investments exhibit the undesirable feature that they drop in value when the probability of job-loss increases. In economic downturns, consumers need an extra incentive to hold equities and other similar investment instruments. The equity premium is thus rationalised as the added incentive needed to

make equities palatable to investors.

There is also an alternative point of view, held by a group of academicians and professionals who claim that at present there is no equity premium and by implication no puzzle. To address these claims we need to differentiate between two different interpretations of the term equity premium. One is the ex-post or realised equity premium. This is the actual, *historically observed* difference between the return on the market, as captured by a stock index, and the risk-free rate, as proxied by the return on government bills. This is what Edward Prescott and this author addressed in the 1979 paper.

However, there is a related concept - the ex-ante equity premium. This is a forward-looking measure of the premium, that is, the equity premium that is *expected* to prevail in the future or the conditional equity premium given the current state of the economy. To elaborate, after a bull market, when stock valuations are high relative to fundamentals, the ex-ante equity premium is likely to be low. Nonetheless, it is precisely when the market has risen sharply, that the ex-post, or the realised premium is high. Conversely, after a major downward correction, the ex-ante (expected) premium is likely to be high, while the realised premium will be low. This should not come as a surprise, since returns to stock have been documented to revert to a mean over time.

Which of these interpretations of the equity premium is relevant for an investment advisor? Clearly this depends on the planning horizon. The 1979 paper referred to very long investment horizons. It has almost nothing to say about what the premium will be over the next few years. The ex-post equity premium is the realisation of a stochastic process over a certain period and it has varied considerably over time as shown in the figure below. Further, the variation depends on the time horizon over which it is measured. There have even been periods when it has been negative.



Source: Ibbotson 2001

Market watchers and others who are interested in short-term investment planning will wish to project the conditional equity premium over their planning horizon. This is no simple task. However, even if current market conditions suggest that the conditional equity premium is small (and there seems to be general consensus that it is), this in itself does not mean that it was obvious that either the historical premium was too high or that the equity premium has diminished.

The data used to document the equity premium over the past hundred years it is probably as good as any economic data we have and, a hundred years is long series when it comes to economic data. Before we dismiss the premium, we not only need to have an understanding of the observed phenomena but also why the future is likely to be different. In the absence of this, we can make the following claim based on what we know. Over the long horizon the equity premium is likely to be similar to what it has been in the past and the returns to investment in equity will continue to substantially dominate those in bonds for investors with a long planning horizon.

Further reading

Campbell, J.Y. and Cochrane, J.H. (1999) “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behaviour,” *Journal of Political Economics* 107: 205-251.

Cochrane, J.H. (1997) “Where is the Market Going? Uncertain Facts and Novel Theories,” *Economic Perspectives*, 21: 3-37.

Constantinides, G.M., Donaldson, J.B., and Mehra, R. “Junior Can’t Borrow: A New Perspective on the Equity Premium Puzzle”. Forthcoming *Quarterly Journal of Economics*. <http://www.econ.ucsb.edu/~mehra/junior.pdf>

Kocherlakota, N.R. (1996) “The Equity Premium: It’s Still a Puzzle.” *J. Econ. Literature*, 34: 42-71.

Mehra, R. and Prescott, E.C. (1985) “The Equity Premium: A Puzzle.” *J. Monetary Econ.* 15: 145-61.

Siegel, J. (1998) *Stocks for the Long Run*, 2nd edn, New York. Irwin.