Journal of Economic Perspectives—Volume 11, Number 1—Winter 1997—Pages 191–200

Anomalies
The Equity Premium Puzzle

Jeremy J. Siegel and Richard H. Thaler

Economics can be distinguished from other social sciences by the belief that most (all?) behavior can be explained by assuming that rational agents with stable, well-defined preferences interact in markets that (eventually) clear. An empirical result qualifies as an anomaly if it is difficult to "rationalize" or if implausible assumptions are necessary to explain it within the paradigm. Suggestions for future topics should be sent to Richard Thaler, c/o *Journal of Economic Perspectives,* Graduate School of Business, University of Chicago, Chicago, IL 60637, or thaler@gsb.uchicago.edu.

Introduction

Suppose your great grandmother had some money lying around at the end of 1925 and, with rational expectations, anticipated your birth and decided to bequeath you $1000. Naturally, since you weren't born yet, she invested the money, and being worried about the speculative boom in stocks going on at the time, she put the money in Treasury bills, where it remained until December 31, 1995. On that date it was worth $12,720. Imagine, instead that she had invested the money in a (value-weighted) portfolio of stocks. You would now have $842,000, or 66 times

[1] The returns we quote are arithmetic returns. Geometric returns are slightly lower.

■ *Jeremy J. Siegel is Professor of Finance, Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania. Richard H. Thaler is Robert P. Gwinn Professor of Economics and Behavioral Science, Graduate School of Business, University of Chicago, Chicago, Illinois, and Research Associate, National Bureau of Economic Research, Cambridge, Massachusetts.*

as much money. This difference in returns (3.7 percent vs. 10.1 percent[1]) is strikingly large. The difference between the return on stocks and the return on a risk-free asset such as treasury bills is called the equity premium (or the equity risk premium, since it is thought to be attributable to the higher risk associated with stocks). The fact that it is too large to be explained by standard economic models is called the equity premium puzzle.[2]

The Puzzle

Of course, stocks are riskier than Treasury bills, so we should expect them to earn higher returns. How can we tell whether the equity premium is too big? Mehra and Prescott (1985) were the first to declare the equity premium an official "puzzle." They used a standard general equilibrium model in which individuals have additively separable utility functions (meaning that my utility of consumption this year does not depend on my consumption in other years) and constant relative risk aversion. In this model, the only parameter is the coefficient of relative risk aversion, A. The interpretation of A is that if consumption falls by 1 percent, then the marginal value of a dollar of income increases by A percent. The question Mehra and Prescott (1985) posed was this: what value of A is necessary to explain the historic equity premium? The value they obtained was between 30-40, which they concluded was much too high to be reasonable.

Why did they conclude that this value of A is too high? Suppose you have a gamble where you face a 50 percent chance to double your wealth (including your human capital) and a 50 percent chance to have your wealth fall by half. How much would you be willing to pay to avoid such a gamble? If your coefficient of relative risk aversion is 30, then you would be willing to pay 49 percent of your wealth to avoid the 50 percent chance of losing half your wealth. This seems absurd.[3]

There is another aspect of the data that is puzzling. A high value of A implies that individuals should want desperately to smooth consumption over time, because consumption shortfalls deliver far more pain than surpluses give pleasure.[4] Since the economy becomes richer over time, individuals should all try to borrow from their richer future in order to improve their (relatively) impoverished present. But this common desire to borrow should lead to high real interest rates. Instead, the real rate of interest has been scarcely positive over long periods of time. Thus, as

[2] Our review of this puzzle will, of necessity, be brief. For more detail see Abel (1995), Kocherlakota (1996) and Siegel (1994).

[3] Mehra and Prescott point out (p. 154) that most empirical estimates of A are in the neighborhood of 1.0–2.0. Arrow (1971, p. 98) argues on theoretical grounds that A "must hover around 1, being, if anything, somewhat less for low wealths and somewhat higher for high wealths."

[4] This follows because in the Mehra-Prescott model the coefficient of relative risk aversion, A, is also the inverse of the elasticity of intertemporal substitution.

pointed out by Weil (1989), the equity premium puzzle could as easily be called the (low) risk-free rate puzzle.[5]

Empirical Questions

There are two broad approaches to explaining the equity premium puzzle. One is to find factors that require adjustment to the empirical side of the puzzle: for example, to uncover data that would make the equity premium smaller or equity returns riskier. The other option is to explore different theoretical frameworks. We consider the empirical questions first and then discuss the theoretical variations below.

Longer Time Period

The question about whether the time period studied by Mehra and Prescott (1985) was special has been investigated by Siegel (1992a,b), who extended the U.S. data on real stock and bond returns back to 1802. He finds that early stock returns did not exceed fixed income returns by nearly the same magnitude they did in more recent data. Siegel divides the whole period into three subperiods: 1802–1871, the early period of U.S. development; 1872–1925, the middle period where data on stock and fixed income returns are of far higher quality; and 1926 to the present.

Siegel finds that real returns in the short-term fixed income market have fallen dramatically over time: from 5.4 percent in the first period, to 3.3 percent in the second, and 0.7 percent since 1926. The real return on equity, in contrast, has remained remarkably constant. As a result, the excess returns on the stock market over the risk-free rate have risen from 2.9 percent to 4.7 percent and finally 8.1 percent over the most recent period. Over the 193-year period, the excess return on equity has been 5.3 percent per year, 1.3 percentage points less than that reported by Mehra and Prescott (1985).

The reason for the fall in the real return on short-term government bonds over the last two centuries is not well understood. In the earliest period, there may have been a greater default risk perceived for a young country. As for the low return to bonds in the last 70 years, it is likely that bondholders in the early post–World War II period did not expect the high inflation of the 1970s, which diminished their real returns. However, short-term investors should presumably have captured the inflation premium in their yield, and yet real returns on short-dated government bills were persistently negative over the high-inflation 1970s.

Since 1982 (after the end of the inflationary 1970s) the real return on short-term government securities has averaged about 3 percent. This higher return on

[5] Kandel and Stambaugh (1991) note that in a world with uncertainty, the real rate of interest is also negatively related to the degree of risk aversion, which can partially explain the low observed interest rates.

bonds would have reduced the equity premium had not the return on stocks been exceptionally high.

Survivorship Bias

An alternate explanation for the equity premium puzzle is that investors are rationally worried about a small chance of an economic catastrophe of some kind, which, though it had not happened, might have (Reitz, 1988). Of the 36 stock exchanges that operated at the turn of the twentieth century, more than one-half of them had significant interruptions or were abolished outright (Brown, Goetzmann and Ross, 1995). Hence, the equity risk premium, estimated from U.S. data alone, is necessarily distorted by the fact that it is calculated for a survivor. On this argument, the riskiness of equities is understated by estimates relying on U.S. existing data, because the data do not show the catastrophe that might have occurred here also, but didn't.

This kind of explanation is not easily testable. However, several objections can be raised. First, the time period studied by Mehra and Prescott (1985) does contain an economic catastrophe, namely the 1929 stock market crash and the subsequent Great Depression. Between 1929 and 1933, stocks lost about 80 percent of their value and did not recover their lost value again until late in World War II. Moreover, careful evaluation of the stock markets that closed for extended periods shows that most did reopen and eventually rewarded equity holders. Gregor Gielen (1994) and Hirose and Tso (1995) have calculated stock returns for both Germany and Japan through World War II. Despite the defeat of the Axis powers, and the precipitous decline in their stock prices at the close of the war, the average real compound annual return on German stocks from 1926 through 1995 was 5.9 percent in Germany and 4.0 percent in Japan. In contrast, in Germany, the hyperinflation of the 1920s wiped out bondholders altogether, and in Japan, the post–World War II hyperinflation did the same. In both these hyperinflations, equities managed to regain most of their real value. Hence if the equity premium is measured as the difference between real returns on stocks and fixed income securities, it was actually *greater* for Japan and Germany during this century than for the United States. More generally, most financial holocausts that destroy stock values have been associated with hyperinflation or financial wealth confiscation where investors are often worse off in bonds than in stocks.[6]

Mean Reversion and Aversion

The equity premium is a puzzle because the measured risk associated with equity returns is not high enough to justify the observed high returns. However,

[6] The equity premium puzzle is also observed in other smaller markets. In one recent analysis, John Campbell (1996) estimates the values of the coefficient of relative risk aversion, A, implied by the post-1970 times series of asset prices and consumption in several other countries, including Australia, Canada, France, Italy, the Netherlands, Spain and Sweden (as well as the larger countries we have already discussed). Except for three countries where the implied value of A is negative (because stock prices and consumption are negatively correlated), the obtained values for A are all very high—from 31 to over 5000. He also reports an estimate of 62 for Sweden for the period 1919-1992.

the usual measurements of risk, such as standard deviation of annual rates of returns, may mischaracterize long-term risk if year-to-year returns do not follow a random walk. To investigate the implications of this assumption, Siegel (1992b) has examined the variability of real returns of equity and fixed income assets over the period 1802–1992 and extended here through 1995. He finds that the observed deviations from a random walk in stock and bill real returns actually deepen the puzzle.

If returns are independent from one year to the next, then the standard deviation of annual average decreases with the square root of the horizon. Although the standard deviation of one-year returns is 18.15 percent, the theoretical (assuming a random walk) standard deviation of annual rates of return over 20-year periods would be just 4.06.[7] However, Poterba and Summers (1988) show that the standard deviation of stock returns actually decreases more quickly than it would if returns were a random walk because stock returns display mean reversion. Several bad years are more likely to be followed by good ones, and vice versa. Thus, Siegel finds that the actual standard deviation of 20-year rates of return is only 2.76 percent. This means that for long-horizon investors, the risk of holding stocks is less than one would expect by just looking at the annual standard deviation of returns.

Mean reversion is not, however, a characteristic of the real returns on fixed income assets. In contrast to stocks, the standard deviation of average annual real returns to bonds and treasury bills decreases *less* than the square root of the horizon. This behavior is called *mean aversion.* While the annual standard deviation of real T-bill rates of return is about a third of that of stocks (6.14 percent), the standard deviation of annual rates of returns for 20-year horizons is actually greater than that of stocks: 2.86 percent.

This analysis suggests that the equity premium is even a bigger puzzle than has previously been thought. It is not that the risk of equities is not great enough to explain their high rate of return; rather, for long-term investors, fixed income securities have been riskier in real terms. By this reasoning, the equity premium should be negative!

Theoretical Explanations

The combination of the equity premium puzzle and the real rate puzzle has captured the attention of many economic theorists who have taken on the challenge of modifying the theory of the representative agent to accommodate the historical facts. None has been completely successful in resolving the puzzle.

One approach, pioneered by Epstein and Zin (1989), uses a utility function that breaks the rigid link between the coefficient of relative risk aversion and

[7] Of course, the standard deviation of *wealth* is increasing as the horizon increases. It is the standard deviation of the annual rate of return that declines as the horizon lengthens.

elasticity of intertemporal substitution. These utility functions (which are not consistent with the axioms of expected utility theory) allow for the possibility of explaining both a high equity premium and a low real interest rate. Still, they find that this approach can only explain about a third of the equity premium.

Mankiw and Zeldes (1991) suggest that the equity premium puzzle may result from the aggregation of the consumption of stockholders and nonstockholders. Not including pension accounts, they found that almost three-quarters of individuals do not hold stocks. Of those that do, their consumption is three times more sensitive to stock market fluctuations than that found in aggregate data. But even after making this adjustment, the level of risk aversion needed to explain the equity premium puzzle is still in the neighborhood of 10.

Other economists have modified the utility function by making the utility of consumption depend on a comparison between current consumption and some benchmark. If the benchmark is taken to be prior levels of consumption, then the behavior can be described as "habit formation," as first suggested by Duesenberry (1952). In the context of the equity premium, habit formation has the effect of making an investor more sensitive to short-run reductions in consumption. This implies a high short-run risk aversion but a lower long-run risk aversion (Constantinides, 1990). However, habit formation cannot explain the difference in returns between stocks and bills (Ferson and Constantinides, 1991).

Another possible benchmark with which current consumption can be compared is the consumption levels of others. I may get utility not just from my own consumption but from knowing that I am consuming more than you are. Conversely, if you become better off and I do not, I am miserable. Abel (1990) examines asset pricing when agents have this type of utility functions, which he terms "catching up with the Joneses." A similar approach has been taken by Campbell and Cochrane (1995). Compared to the standard constant-elasticity-of-substitution utility function, this has the effect of increasing an individual's marginal utility of consumption in the future, since others at that time will also become better off. Catching up with the Joneses reduces an individual's desire to borrow against higher future consumption and hence lowers the real rate, but leaves an investor just as risk averse to contemporaneous shocks. This model can explain the equity premium with a value of A of only six: still high, but more plausible than 30.

One final "solution" to the puzzle is to deny that it is a puzzle. This is the tack taken by Kandel and Stambaugh (1991), who argue that perhaps investors really do have very high values of A. They point out that while high levels of risk aversion may lead to unreasonable behavior with respect to large changes in consumption, it does not imply implausible behavior for small changes in wealth. For example, to avoid a 50-50 chance of your consumption rising or falling by 1 percent if the coefficient of risk aversion is 10, one would pay 5 percent of the gamble. Even if the coefficient is as high as 29 (which is their estimate of A) an investor would pay only 14.3 percent of the gamble to avoid the risk of a 1 percent rise or fall in wealth. Neither of these actions seem completely unreasonable.

In evaluating this argument, however, it is important to remember that in the

domain of retirement savings the stakes are, in fact, large relative to wealth. Similarly, university endowments are a substantial portion of the wealth of private universities. This seems to put us back into the high-stakes category where values of A greater than 10 lead to absurd results.

Myopic Loss Aversion

The models discussed above use reasonable assumptions that the utility of consumption depends on the past levels of consumption or on the consumption of peers. Another approach, in a similar vein, is offered by Benartzi and Thaler (1995) using Tversky and Kahneman's (1992) prospect theory. In Benartzi and Thaler's model, all investors (individuals, pension plan sponsors, endowment fund managers, and so on) are assumed to get utility from the *changes* in the value of their portfolios; that is, utility comes from returns, not from the overall level of assets. Furthermore, investors display "loss aversion": losses are assumed to hurt significantly more than gains yield pleasure.[8]

When investors have loss averse preferences, their attitude toward risk depends crucially on the time horizon over which returns are evaluated. For example, an investor with these preferences who computed the value of her portfolio every day would find investing in stocks very unattractive, since stock prices fall almost as often as they rise on a daily basis, and losses are psychologically doubled. Consider, on the other hand, a modern version of Rip Van Winkle, who, knowing he is about to go to sleep for 20 years, makes one final phone call to his broker. Rip should sleep soundly in the knowledge that over a 20-year period, stocks have never declined in real value.

Using this interplay between the time period used to evaluate investments and loss aversion, Benartzi and Thaler (1995) estimate what evaluation period would make investors indifferent between stocks and bonds (or bills). They do this by simulating distributions of returns for stocks, bonds and bills over various horizons (from one month and up) by selecting months at random from history. They find that the evaluation period that makes stocks and bonds equally attractive is about 13 months, or just over one year.

How can this result be evaluated? One method is to use the same "plausibility" test that Mehra and Prescott (1985) apply. They declared the equity premium a puzzle because they judged a value of A much greater than 10 to be implausible. In contrast, a one-year evaluation period seems consistent with the observation that tax returns and many other activities take place once a year, making annual evaluations particularly salient.

[8] Specifically, the value function used is

$$v(x) = x^{\alpha} \quad \text{if } x \geq 0$$
$$-\lambda(-x)^{\alpha} \quad \text{if } x < 0$$

where x is returns. The parameters a and λ (the coefficient of loss aversion) have been estimated by Tversky and Kahneman to be .88 and 2.25, respectively.

An implication of myopia as an explanation is that if people did concentrate on long-term results, they would invest more in stocks. In a follow-up paper, Benartzi and Thaler (1996) report an experiment in which groups of university employees were shown distributions of returns for two hypothetical retirement funds, *A* and *B*, where the distributions were derived from the actual distributions of stocks and bonds since 1926. One group was shown a distribution of annual returns; this group invested 40 percent of their money in stocks. Another group was shown a simulated distribution of 30-year returns derived from the annual return data by drawing years at random. Although this group is given essentially the same information, they chose to invest 90 percent of their money in stocks, presumably because they found the long-run return distribution for stocks more attractive than for bonds.

Commentary

The equity premium puzzle is a rare bird among economics anomalies. Because economic theory rarely makes quantitative predictions, many tests of the theory come down to whether the sign of some magnitude is the same as the theory predicts. Such tests are hard to fail. Indeed, since stocks are riskier than bonds or bills (at least on an annual basis) according to a crude sign test, asset pricing conforms to economic theory. The ingenious contribution of Mehra and Prescott (1985) was to come up with a quantitative test of the theory. They established a value of 10 as a reasonable upper bound for *A* (we feel an even lower upper bound could be justified), and their results were a resounding rejection of the theory.

What should we make of these results? One view is that history has just been kind to stock markets, especially those in the larger markets. According to this view, we have just experienced 200 years of good luck. Yet we have shown that the equity returns in such countries as Germany and Japan, which have experienced much bad luck, still greatly outperform short-term fixed income securities. And the equity premium has been high over extremely long periods of time. This raises the question of how long it should take for investors in an economy to learn about the true risk and return on financial assets.

Another view is that investors really are extremely risk averse, as argued by Kandel and Stambaugh (1991). Eugene Fama (1991, p. 1596) offers a different interpretation:

> . . . a large equity premium is not necessarily a puzzle; high risk aversion (or low intertemporal elasticity of substitution for consumption) may be a fact. Roughly speaking, a large equity premium says that consumers are extremely averse to small negative consumption shocks. This is in line with the perception that consumers live in morbid fear of recessions (and economists devote enormous energy to studying them) even though, at

least in the post war period, recessions are associated with small changes in consumption.

Note that Fama seems to suggest that while the risk aversion displayed is real, that it is some kind of mistake by consumers. A similar view is offered by Thomas MaCurdy and John Shoven (1992). They find it so difficult to understand why investors put any of their retirement funds in bonds that they conclude (p. 12) that people must be "confused about the relative safety of different investments over long horizons." In the myopic loss aversion explanation investors are also making a mistake: they fail to aggregate over time periods.[9]

What are the practical implications of this? If you believe that the equity premium is simply a fair return for bearing the risks associated with buying stocks, then you can base your asset allocation decision in part on whether you think you are more or less risk averse than the marginal investor. However, if you think that the equity premium is partially derived from other peoples' mistakes and fears, then you might find equities very attractive. Indeed, most economists we know have a very high proportion of their retirement wealth invested in equities (as we do). Are economists just risk lovers, or do they think that the equity premium is big enough to be attractive?

We must stress that our analysis has all been on historical data, which suggest that the equity premium has been too large *in the past.* Of course, as mutual fund companies typically say in their advertisements, past returns are no guarantee of future returns. What is the equity premium now? There is reason to believe that it is lower than it has been in the past. As we noted above, current estimates of expected inflation imply real rates of return of 3-4 percent on long-term bonds, just 3 percent below the historic real return on equities. In a recent detailed investigation of this question, Blanchard (1993) concluded that the equity premium was about 3 percent (or half what it has been over the last 70 years), which he attributes in part to the fading memories of the Great Depression and to the disappearance of significant inflation. Still, even a 3 percent equity premium is substantial when compounded over long periods, and for long-horizon investors such as the young saving for retirement, pension plans and endowments, we find the case for equities compelling. However, if after reading this piece you decide to put more of your retirement savings in stocks, remember we are stressing long-term results and will not accept complaints for 20 years. Feel free to call us in 2017.

■ *Thanks for helpful comments go to Andrew Abel, Olivier Blanchard, George Constantinides, Robert Stambaugh and the editors.*

[9] That is, even if loss aversion is real, investors should realize that they should care about retirement consumption, not returns along the way. To paraphrase the well-known country song, loss averse long-term investors must learn not to "count their money while they're sittin' at the table. . . ."

References

Abel, Andrew B., "Asset Prices Under Habit Formations and Catching up with the Joneses," *America Economic Review,* March 1990, *80,* 38–42.

Abel, Andrew B., "The Equity Premium Puzzle," *Business Review, Federal Reserve Bank of Philadelphia,* September/October 1995, 3–14.

Arrow, Kenneth J., *Essays in the Theory of Risk Bearing.* Amsterdam: North Holland, 1971.

Benartzi, Shlomo, and Richard H. Thaler, "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics,* February 1995, *110*:1, 75–92.

Benartzi, Shlomo, and Richard H. Thaler, "Risk Aversion or Myopia: The Fallacy of Small Numbers and its Implications for Retirement Savings," working paper, University of California, Los Angeles, 1996.

Blanchard, Olivier, "Movements in the Equity Premium," *Brookings Papers on Economic Activity,* 1993, *2,* 519–43.

Brown, Stephen J., William N. Goetzmann, and Stephen A. Ross, "Survival," *Journal of Finance,* July 1995, *50,* 853–74.

Campbell, John, "Consumption and the Stock Market: Interpreting International Experience." NBER Working Paper No. 5619, 1996.

Campbell, John, and Steve Cochrane, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior." NBER Working Paper No. 4995, 1995.

Constantinides, George M., "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy,* June 1990, *98*:3, 519–43.

Duesenberry, James, *Income Savings and the Theory of Consumer Behavior.* Cambridge: Harvard University Press, 1952.

Epstein, Larry G., and Stanley E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica,* 1989, *57,* 937–69.

Fama, Eugene F., "Efficient Capital Markets: II," *Journal of Finance,* December 1991, *46*:5, 1575–617.

Ferson, Wayne, and George M. Constantinides, "Habit Persistence and Durability in Aggregate Consumption: Empirical Tests," *Journal of Financial Economics,* October 1991, *39,* 199–240.

Gielen, Gregor, *Konen Aktienkurse noch Steigen?* Wiesbaden, Ger.: Gabler, 1994.

Hirose, Hiroaki, and Yi Tso, "Japanese Market Returns," unpublished paper, Wharton School, December 29, 1995.

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

Kocherlakota, Narayana R., "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature,* March 1996, *34,* 42–71.

MaCurdy, Thomas, and John Shoven, "Accumulating Pension Wealth with Stocks and Bonds," working paper, Stanford University, January 1992.

Mankiw, N. Gregory, and Stephen P. Zeldes, "The Consumption of Stockholders and Nonstockholders," *Journal of Financial Economics,* March 1991, *29,* 97–112.

Mehra, Rajnish, and Edward C. Prescott, "The Equity Premium: A Puzzle," *Journal of Monetary Economics,* March 1985, *15,* 145–62.

Poterba, James A., and Lawrence H. Summers, "Mean Reversion in Stock Prices: Evidence and Implications," *Journal of Financial Economics,* October 1988, *22,* 27–59.

Reitz, Thomas A., "The Equity Risk Premium: A Solution," *Journal of Monetary Economics,* July 1988, *22,* 117–31.

Siegel, Jeremy J., "The Equity Premium: Stock and Bond Returns Since 1802," *Financial Analyst Journal,* January/February 1992a, *48,* 28–38.

Siegel, Jeremy J., "The Real Rate of Interest from 1800-1990: A Study of the U.S. and U.K.," *Journal of Monetary Economics,* April 1992b, *29,* 227–52.

Siegel, Jeremy J., *Stocks for the Long Run: A Guide to Selecting Markets for Long-Term Growth.* Burr Ridge, Ill.: Irwin Professional Publishing, 1994.

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

Weil, Philippe, "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics,* November 1989, *24,* 401–21.