Understanding Risk and Return

2001 Marshall Lectures
University of Cambridge

John Y. Campbell
Harvard University

Lecture 1: Two Puzzles of Asset Pricing

May 8, 2001

The tradeoff of risk and return is becoming ever more important for individuals, institutions, and public policy. In fact, Bernstein (1996) suggests that the rational analysis of risk is a defining characteristic of the modern age.

These lectures will explore risk and return in aggregate stock market investment. They are based on several earlier expositional and research pieces, notably Campbell (1999, 2000), Campbell and Cochrane (1999), Campbell and Shiller (2001), and Campbell and Viceira (2001).

The comparison of the stock market with the money market is startling. For example, if we look at log real returns on US stocks and Treasury bills over the period 1947.2-1998.4, we find, first, that the average stock return is 8.1%, while the average bill return is 0.9%; and second, that the volatility of the stock return is 15.6% while the volatility of the ex post real bill return is only 1.8%.[1]

These facts lead to two puzzles of asset pricing. The first was christened the *equity premium puzzle* by Mehra and Prescott (1985): Why is the *average* real stock return so high (in relation to the average short-term real interest rate)? The second might be called the *equity volatility puzzle*: Why is the *volatility* of stock returns so high (in relation to the volatility of the short-term real interest rate)? The classic reference to this second puzzle is Shiller (1981).

Financial economists have tried to resolve these puzzles by linking asset prices to aggregate consumption. This is a natural approach because consumption is the most obvious determinant of marginal utility (in simple models, the only determinant). Hence, covariance with consumption measures risk. Also, consumption can be thought of as the dividend on the portfolio of aggregate wealth. It is natural to model stocks as claims to the stream of consumption.

Unfortunately, aggregate consumption has several properties that deepen the puzzles of asset pricing. First, real consumption growth is very stable, with an annualized standard deviation of 1.1%. Second, the correlation of consumption growth and stock returns is weak (0.23 at a quarterly frequency, and 0.34 at an annual frequency).

[1] The gap between average stock and bill returns is even higher if one computes an average of simple returns (an arithmetic return average) rather than an average of log returns (a geometric return average). In this lecture I will work with log returns throughout, but I will adjust average log returns as required by the theoretical models I explore. In practice this means adding one-half the variance to the difference of average log returns, in effect converting from geometric to arithmetic average returns.

Third, stock prices have very little ability to forecast consumption growth. The R^2 statistic of a regression of consumption growth on the log dividend-price ratio is never greater than 4% at horizons from 1 to 4 years.

Financial economists also try to link stock prices to the behavior of dividends, without assuming that dividends equal consumption. Here too there are puzzles. Quarterly dividend volatility is high (28%), but this is due to strong seasonality in dividends. Annual dividend volatility is only about 6%. This volatility is much larger than consumption growth (1%), but much smaller than stock returns (16%). Stock returns are somewhat more strongly correlated with dividends than with consumption, but the maximum correlation at any horizon up to 4 years is only 0.34 at a 1-year horizon. Finally, the dividend-price ratio has little ability to forecast dividend growth. The R^2 statistic of a regression of dividend growth on the log dividend-price ratio is never greater than 8% at horizons from 1 to 4 years.

This first lecture will introduce the puzzles and explain why simple models have such difficulty with them. The second lecture, "From Puzzles to Portfolios", will discuss alternative explanations and develop implications for portfolio management.

International data

An important preliminary question is whether these features of US financial data are also apparent in other countries. If they are, this suggests that one must look for deep structural causes rather than idiosyncratic features of US experience.

Table 1, an updated version of a table in Campbell (1999), reports average stock and bill returns for Morgan Stanley Capital International data. Returns are continuously compounded (that is, measured as log returns) and deflated using consumer prices. The data start no earlier than 1970 (considerably later in some countries), and they cover 11 developed countries: Australia, Canada, France, Germany, Italy, Japan, the Netherlands, Sweden, Switzerland, the UK, and the US. The MSCI data are combined with macroeconomic data from the International Financial Statistics (IFS) of the International Monetary Fund. For some countries the IFS data are only available quarterly over a shorter sample period, so I use the longest available sample for each country. Sample start dates range from 1970.1 to 1982.2, and sample end dates range from 1997.4 to 1999.3.

For some purposes it is useful to have data over a much longer span of calendar time. The table also reports results for long-term annual data from Sweden (starting in 1920), the UK (starting in 1919), and the US (starting in 1891).

In working with these data, it is important to keep in mind that different national stock markets are of very different sizes, both absolutely and in proportion to national GDP's. At the end of 1993, for example, the Japanese MSCI index was worth only 65% of the US MSCI index, the UK MSCI index was worth only 30% of the US index, the French and German MSCI indexes were worth only 11% of the US index, and all other countries' indexes were worth less than 10% of the US index. The MSCI index accounted for about 80% of GDP in highly capitalized countries such as the UK and Switzerland, but accounted for less than 20% of GDP in Germany and Italy. The theoretical convention of treating the stock market as a claim to total consumption, or as a proxy for the aggregate wealth of an economy, makes much more sense in the highly capitalized countries.

Table 1 shows that stock markets have delivered average real returns of 5% or better in almost every country and time period. The exceptions to this occur in short-term quarterly data, and are concentrated in markets that are particularly small relative to GDP (Italy) or that predominantly represent claims on natural resources (Australia). Short-term debt, on the other hand, has rarely delivered an average real return above 3%. The exceptions to this occur in two countries, Germany and the Netherlands, whose sample periods begin in the late 1970's and thus exclude much of the surprise inflation of the oil-shock period.

Table 1 also reports annualized standard deviations and first-order autocorrelations for returns. The standard deviation of stock returns ranges from 15% to 27%. It is striking that the market with the highest volatility, Italy, is the smallest market relative to GDP and the one with the lowest average return. In the quarterly MSCI data the annualized volatility of real returns on short debt is around 3% for the UK, Italy, and Sweden, around 2.5% for Australia and Japan, and below 2% for all other countries. Volatility is higher in long-term annual data because of large swings in inflation in the interwar period, particularly in 1919-21. Much of the volatility in these real returns is probably due to unanticipated inflation and does not reflect volatility in the ex ante real interest rate.

These numbers show that the equity premium and equity volatility puzzles are not unique to the United States but characterize many other countries as well. Recently a number of authors have suggested that average excess returns in the US may be overstated by sample selection or survivorship bias. If economists study the US because it has had an unusually successful economy, then sample average US stock returns may overstate the true mean US stock return. The international data suggest that this is not a serious problem.[2]

Table 2 reports data on aggregate consumption and stock market dividends. The table illustrates the challenge of explaining asset market behavior using real data on consumption or dividends. In the postwar period the annualized standard deviation of real consumption growth is never above 3%. This is true even though the measure of consumption is total consumption, rather than nondurables and services consumption, for all countries other than the US. Even in the longer annual data, which include the turbulent interwar period, consumption volatility slightly exceeds 3% only in the US.

The volatility of dividend growth is much greater than the volatility of consumption growth, but generally less than the volatility of stock returns. The exceptions to this occur in countries with highly seasonal dividend payments; these countries have large negative first-order autocorrelations for quarterly dividend growth (reported in the last column of the table) and much smaller volatility when dividend growth is measured over a full year rather than over a quarter.

The equity premium puzzle and the stochastic discount factor

I now state the equity premium puzzle using the stochastic discount factor (SDF) paradigm. This approach to asset pricing, which has its roots in the work of Rubinstein (1976), Breeden (1979), Grossman and Shiller (1981), and Shiller (1982), has become increasingly influential since the work of Hansen and Jagannathan (1991). Cochrane (2001) provides a unified textbook treatment of asset pricing in these terms.

Consider the intertemporal choice problem of an investor, indexed by k, who can trade freely in some asset i and can obtain a gross simple rate of return $(1 + R_{i,t+1})$ on the asset held from time t to time $t + 1$. If the investor consumes C_{kt} at time t and has time-separable utility with discount factor δ and period utility $U(C_{kt})$, then his first-order condition is

$$U'(C_{kt}) = \delta \mathrm{E}_t \left[(1 + R_{i,t+1})U'(C_{k,t+1})\right] . \tag{1}$$

[2] Jorion and Goetzmann (1999) consider international stock-price data from earlier in the 20th Century and argue that the long-term average real growth rate of stock prices has been higher in the US than elsewhere. However they do not have data on dividend yields, which are an important component of total return and are likely to have been particularly important in Europe during the troubled interwar period. Dimson, Marsh, and Staunton (2000) do measure dividend yields and find that total returns in the US did not exceed returns in all other countries in the early 20th Century.

The left hand side of (1) is the marginal utility cost of consuming one real dollar less at time t; the right hand side is the expected marginal utility benefit from investing the dollar in asset i at time t, selling it at time $t+1$, and consuming the proceeds. The investor equates marginal cost and marginal benefit, so (1) must describe the optimum.

Dividing (1) by $U'(C_{kt})$ yields

$$1 = \mathrm{E}_t\left[(1+R_{i,t+1})\,\delta\frac{U'(C_{k,t+1})}{U'(C_{kt})}\right] = \mathrm{E}_t\left[(1+R_{i,t+1})\,M_{k,t+1}\right] \ , \tag{2}$$

where $M_{k,t+1} = \delta U'(C_{k,t+1})/U'(C_t)$ is the intertemporal marginal rate of substitution of the investor, also known as the stochastic discount factor or SDF. Since marginal utility must always be positive, the SDF must always be positive.

The derivation just given for equation (2) assumes the existence of an investor maximizing a time-separable utility function, but in fact the equation holds more generally. The existence of a positive stochastic discount factor is guaranteed by the absence of arbitrage in markets in which non-satiated investors can trade freely without transactions costs. In general there can be many such stochastic discount factors—for example, different investors k whose marginal utilities follow different stochastic processes will have different $M_{k,t+1}$—but each stochastic discount factor must satisfy equation (2). It is common practice to drop the subscript k from this equation and simply write

$$1 = \mathrm{E}_t\left[(1+R_{i,t+1})\,M_{t+1}\right] \ . \tag{3}$$

In complete markets the stochastic discount factor M_{t+1} is unique because investors can trade with one another to eliminate any idiosyncratic variation in their marginal utilities.

To understand the implications of (3) in a simple way, I follow Hansen and Singleton (1983) and assume that the joint conditional distribution of asset returns and the stochastic discount factor is lognormal and homoskedastic. While these assumptions are not literally realistic—stock returns in particular have fat-tailed distributions with variances that change over time—they do make it easier to discuss the main forces that should determine the equity premium.

When a random variable X is conditionally lognormally distributed, it has the convenient property that

$$\log \mathrm{E}_t X = \mathrm{E}_t \log X + \tfrac{1}{2}\mathrm{Var}_t \log X \ , \tag{4}$$

where $\mathrm{Var}_t \log X \equiv \mathrm{E}_t[(\log X - \mathrm{E}_t \log X)^2]$. If in addition X is conditionally homoskedastic, then $\mathrm{Var}_t \log X = \mathrm{E}[(\log X - \mathrm{E}_t \log X)^2] = \mathrm{Var}(\log X - \mathrm{E}_t \log X)$. Thus with joint conditional lognormality and homoskedasticity of asset returns and consumption, I can take logs of (3) and obtain

$$0 = \mathrm{E}_t r_{i,t+1} + \mathrm{E}_t m_{t+1} + \left(\tfrac{1}{2}\right)\left[\sigma_i^2 + \sigma_m^2 + 2\sigma_{im}\right]. \tag{5}$$

Here $m_t = \log(M_t)$ and $r_{it} = \log(1+R_{it})$, while σ_i^2 denotes the unconditional variance of log return innovations $\mathrm{Var}(r_{i,t+1} - \mathrm{E}_t r_{i,t+1})$, σ_m^2 denotes the unconditional variance of innovations to the stochastic discount factor $\mathrm{Var}(m_{t+1} - \mathrm{E}_t m_{t+1})$, and σ_{im} denotes the unconditional covariance of innovations $\mathrm{Cov}(r_{i,t+1} - \mathrm{E}_t r_{i,t+1},\, m_{t+1} - \mathrm{E}_t m_{t+1})$.

Equation (5) has both time-series and cross-sectional implications. Consider first an asset with a riskless real return $r_{f,t+1}$. For this asset the return innovation variance σ_f^2 and the covariance σ_{fm} are both zero, so the riskless real interest rate obeys

$$r_{f,t+1} = -\mathrm{E}_t m_{t+1} - \frac{\sigma_m^2}{2}. \tag{6}$$

The log riskless interest rate is negatively related to the conditional expectation of the log SDF. When the SDF is expected to be high, marginal utility in the future is expected to be high relative to the present; the investor has an incentive to save, and this depresses the equilibrium riskless interest rate. The log riskless interest rate also depends negatively on the conditional volatility of the log SDF. Volatility produces a precautionary savings motive, which also depresses the riskless interest rate.

Subtracting (6) from (5) yields an expression for the expected excess return on risky assets over the riskless rate:

$$\mathrm{E}_t[r_{i,t+1} - r_{f,t+1}] + \frac{\sigma_i^2}{2} = -\sigma_{im}. \tag{7}$$

The variance term on the left hand side of (7) is a Jensen's Inequality adjustment arising from the fact that we are describing expectations of log returns. This term effectively converts the return difference from a geometric average to an arithmetic average. It would disappear if we rewrote the equation in terms of the log expectation of the ratio of gross simple returns: $\log \mathrm{E}_t[(1+R_{i,t+1})/(1+R_{f,t+1})] = -\sigma_{im}$.

The right hand side of (7) says that the risk premium is the negative of the covariance of the asset with the stochastic discount factor. An asset with a high expected return must have a low covariance with the stochastic discount factor. Such an asset tends to have low returns when investors have high marginal utility. It is risky in that it fails to deliver wealth precisely when wealth is most valuable to investors. Investors therefore demand a large risk premium to hold it.

The covariance σ_{im} can be written as the product of the standard deviation of the asset return σ_i, the standard deviation of the stochastic discount factor σ_m, and the correlation between the asset return and the stochastic discount factor ρ_{im}. Since $\rho_{im} \geq -1$, $-\sigma_{im} \leq \sigma_i \sigma_m$. Substituting into (7),

$$\sigma_m \geq \frac{\mathrm{E}_t[r_{i,t+1} - r_{f,t+1}] + \sigma_i^2/2}{\sigma_i} \,. \tag{8}$$

This inequality was first derived by Shiller (1982); a multi-asset version was derived by Hansen and Jagannathan (1991). The right hand side of (8) is the excess return on an asset, adjusted for Jensen's Inequality, divided by the standard deviation of the asset's return—a logarithmic Sharpe ratio for the asset. (8) says that the standard deviation of the log stochastic discount factor must be greater than this Sharpe ratio for all assets i, that is, it must be greater than the maximum possible Sharpe ratio obtainable in asset markets.

Table 3 uses equation (8) to illustrate the equity premium puzzle. For each data set the first column of the table reports the average excess return on stock over short-term debt, adjusted for Jensen's Inequality by adding one-half the sample variance of the excess log return to get a sample estimate of the numerator in (8). This adjusted average excess return is multiplied by 400 to express it in annualized percentage points. The second column of the table gives the annualized standard deviation of the excess log stock return, a sample estimate of the denominator in (8). This standard deviation was reported earlier in Table 1. The third column gives the ratio of the first two columns, multiplied by 100; this is a sample estimate of the lower bound on the standard deviation of the log stochastic discount factor, expressed in annualized percentage points. In the postwar US data the estimated lower bound is a standard deviation greater than 50% a year; in the other quarterly data sets it is below 10% for Italy, between 15% and 20% for Australia and Canada, and above 30% for all the other countries. In the long-run annual data sets the lower bound on the standard deviation exceeds 30% for all three countries. These are extraordinarily high volatilities considering that the stochastic discount factor M_{t+1} is a random variable with a mean close to one that must always be positive.

The equity premium puzzle and consumption-based asset pricing

To understand why these numbers are disturbing, I now follow Rubinstein (1976), Lucas (1978), Breeden (1979), Grossman and Shiller (1979), Mehra and Prescott (1985) and other classic papers on the equity premium puzzle and assume that there is a representative agent who maximizes a time-separable power utility function defined over aggregate consumption C_t:

$$U(C_t) = \frac{C_t^{1-\gamma} - 1}{1-\gamma} \, , \tag{9}$$

where γ is the coefficient of relative risk aversion.

The assumption of power utility is not an arbitrary one. A scale-independent utility function is required to explain the fact that over the past two centuries, as wealth and consumption have grown manyfold, riskless interest rates and risk premia do not seem to have trended up or down. Power utility is one of the few utility functions that have this property.[3] Related to this, if different investors in the economy have different wealth levels but the same power utility function, then they can be aggregated into a single representative investor with the same utility function as the individual investors.

Power utility implies that marginal utility $U'(C_t) = C_t^{-\gamma}$, and the stochastic discount factor $M_{t+1} = \delta(C_{t+1}/C_t)^{-\gamma}$. The assumption made previously that the stochastic discount factor is conditionally lognormal will be implied by the assumption that aggregate consumption is conditionally lognormal (Hansen and Singleton 1983). Making this assumption for expositional convenience, the log stochastic discount factor is $m_{t+1} = \log(\delta) - \gamma \Delta c_{t+1}$, where $c_t = \log(C_t)$, and (5) becomes

$$0 = \mathrm{E}_t r_{i,t+1} + \log\delta - \gamma \mathrm{E}_t \Delta c_{t+1} + \left(\tfrac{1}{2}\right)\left[\sigma_i^2 + \gamma^2\sigma_c^2 - 2\gamma\sigma_{ic}\right] . \tag{10}$$

Here σ_c^2 denotes the unconditional variance of log consumption innovations $\mathrm{Var}(c_{t+1} - \mathrm{E}_t c_{t+1})$, and σ_{ic} denotes the unconditional covariance of innovations $\mathrm{Cov}(r_{i,t+1} - \mathrm{E}_t r_{i,t+1},\, c_{t+1} - \mathrm{E}_t c_{t+1})$.

[3] A few other utility functions also have this property. Epstein and Zin (1991) and Weil (1989) have proposed a recursive utility specification that preserves the scale-invariance of power utility but relaxes the restriction of power utility that the coefficient of relative risk aversion is the reciprocal of the elasticity of intertemporal substitution. Models of habit formation, to be discussed in the second lecture, make relative risk aversion constant in the long run but variable in the short run.

Equation (6) now becomes

$$r_{f,t+1} = -\log\delta + \gamma \mathrm{E}_t \Delta c_{t+1} - \frac{\gamma^2 \sigma_c^2}{2} \ . \tag{11}$$

This equation says that the riskless real rate is linear in expected consumption growth, with slope coefficient equal to the coefficient of relative risk aversion. The conditional variance of consumption growth has a negative effect on the riskless rate by stimulating precautionary savings.

Equation (7) becomes

$$\mathrm{E}_t[r_{i,t+1} - r_{f,t+1}] + \frac{\sigma_i^2}{2} = \gamma \sigma_{ic} \ . \tag{12}$$

The log risk premium on any asset is the coefficient of relative risk aversion times the covariance of the asset return with consumption growth. Intuitively, an asset with a high consumption covariance tends to have low returns when consumption is low, that is, when the marginal utility of consumption is high. Such an asset is risky and commands a large risk premium.

Table 3 uses (12) to illustrate the equity premium puzzle. As already discussed, the first column of the table reports a sample estimate of the left hand side of (12), multiplied by 400 to express it in annualized percentage points. The second column reports the annualized standard deviation of the excess log stock return (given earlier in Table 1), the fourth column reports the annualized standard deviation of consumption growth (given earlier in Table 2), the fifth column reports the correlation between the excess log stock return and consumption growth, and the sixth column gives the product of these three variables which is the annualized covariance σ_{ic} between the log stock return and consumption growth.

Finally, the table gives two columns with implied risk aversion coefficients. The column headed RRA(1) uses (12) directly, dividing the adjusted average excess return by the estimated covariance to get estimated risk aversion.[4] The column headed RRA(2) sets the correlation of stock returns and consumption growth equal to one before calculating risk aversion. While this is of course a counterfactual exercise, it is a valuable diagnostic because it indicates the extent to which the equity premium puzzle

[4]The calculation is done correctly, in natural units, even though the table reports average excess returns and covariances in percentage point units. Equivalently, the ratio of the quantities given in the table is multiplied by 100.

arises from the *smoothness* of consumption rather than the *low correlation* between consumption and stock returns. The correlation is hard to measure accurately because it is easily distorted by short-term measurement errors in consumption, and Campbell (1999) shows that empirically it is quite sensitive to the measurement horizon. By setting the correlation to one, the RRA(2) column indicates the extent to which the equity premium puzzle is robust to such issues. A correlation of one is also implicitly assumed in the volatility bound for the stochastic discount factor, (8), and in many calibration exercises such as Mehra and Prescott (1985) or Campbell and Cochrane (1999).

Table 4 shows that the equity premium puzzle is a robust phenomenon in international data. The coefficients of relative risk aversion in the RRA(1) column are generally extremely large. They are usually many times greater than 10, the maximum level considered plausible by Mehra and Prescott (1985). In a few cases the risk aversion coefficients are negative because the estimated covariance of stock returns with consumption growth is negative, but in these cases the covariance is extremely close to zero. Even when one ignores the low correlation between stock returns and consumption growth and gives the model its best chance by setting the correlation to one, the RRA(2) column still has risk aversion coefficients above 10 in most cases.

Could the equity premium puzzle be spurious?

The risk aversion estimates in Table 4 are point estimates and are subject to sampling error. No standard errors are reported for these estimates. However authors such as Cecchetti, Lam, and Mark (1993) and Kocherlakota (1996), studying the long-run annual US data, have found small enough standard errors that they can reject risk aversion coefficients below about 8 at conventional significance levels.

Of course, the validity of these tests depends on the characteristics of the data set in which they are used. Rietz (1988) has argued that there may be a peso problem in these data. A peso problem arises when there is a small positive probability of an important event, and investors take this probability into account when setting market prices. If the event does not occur in a particular sample period, investors will appear irrational in the sample and economists will misestimate their preferences. While it may seem unlikely that this could be an important problem in 100 years of annual data, Rietz (1988) argues that an economic catastrophe that destroys almost all stock-market value can be extremely unlikely and yet have a major depressing effect on stock prices.

One difficulty with this argument is that it requires not only a potential catastrophe, but one which affects stock market investors more seriously than investors in short-term debt instruments. Many countries that have experienced catastrophes, such as Russia or Germany, have seen very low returns on short-term government debt as well as on equity. A peso problem that affects both asset returns equally will affect estimates of the average levels of returns but not estimates of the equity premium.[5] The major example of a disaster for stockholders that did not negatively affect bondholders is the Great Depression of the early 1930's, but of course this is included in the long-run annual data for Sweden, the UK, and the US, all of which display an equity premium puzzle.

Also, the consistency of the results across countries requires investors in all countries to be concerned about catastrophes. If the potential catastrophes are uncorrelated across countries, then it becomes less likely that the data set includes no catastrophes; thus the argument seems to require a potential international catastrophe that affects all countries simultaneously.

Even if the equity premium puzzle is not entirely spurious, there are several reasons to think that stock returns exceeded their true long-run mean in the late 20th Century. Dimson, Marsh, and Staunton (2000) find that international returns were generally higher in the late 20th Century than in the early 20th Century. Siegel (1998) reports similar results for US data going back to the early 19th Century. Fama and French (2000) point out that average US stock returns in the late 20th Century were considerably higher than accountants' estimates of the return on equity for US corporations. Thus if one uses average returns as an estimate of the true cost of capital, one is forced to the implausible conclusion that corporations destroyed stockholder value by retaining and reinvesting earnings rather than paying them out.

Unusually high stock returns in the late 20th Century could have resulted from unexpectedly favorable conditions for economic growth. But they could also have resulted from a correction of historical mispricing, a one-time decline in the equity premium. Several economists have recently argued that the equity premium is now far

[5]This point is relevant for the study of Jorion and Goetzmann (1999). These authors measure average growth rates of real stock prices, as a proxy for real stock returns, but they do not look at real returns on short-term debt. They find low real stock-price growth rates in many countries in the early 20th Century; in some cases these may have been accompanied by low returns to holders of short-term debt. Note also that stock-price growth rates are a poor proxy for total stock returns in periods where investors expect low growth rates, since dividend yields will tend to be higher in such periods.

lower than it was in the early 20th Century (Jagannathan, McGrattan, and Scherbina 2001, McGrattan and Prescott 2001).[6]

Could risk aversion be higher than we thought?

It is possible that the equity premium puzzle has an extremely simple solution, namely that the coefficient of relative risk aversion γ is higher than economists traditionally thought. After all, it is hard to get evidence about risk aversion from any other source than asset markets. Experimental evidence is of very little use because it is almost impossible to design experiments involving significant stakes, and people should be almost indifferent with respect to small gambles.[7] One might think that "thought experiments", or introspection, would be sufficient to rule out very large values of γ, but Kandel and Stambaugh (1991) point out that introspection can deliver very different estimates of risk aversion depending on the size of the gamble considered. This suggests that introspection can be misleading or that some more general model of utility is needed.

The assumption of a high γ, however, leads to a second puzzle. Equation (11) implies that the unconditional mean riskless interest rate is

$$\mathrm{E}\, r_{f,t+1} = -\log\delta + \gamma g - \frac{\gamma^2 \sigma_c^2}{2}\,, \tag{13}$$

where g is the mean growth rate of consumption. Since g is positive, as shown in Table 2, high values of γ imply high values of γg. Ignoring the term $-\gamma^2\sigma_c^2/2$ for the moment, this can be reconciled with low average short-term real interest rates, shown in Table 2, only if the discount factor δ is close to or even greater than one, corresponding to a low or even negative rate of time preference. This is the riskfree rate puzzle emphasized by Weil (1989).

Intuitively, the riskfree rate puzzle is that if investors are risk-averse then with power utility they must also be extremely unwilling to substitute intertemporally. Given positive average consumption growth, a low riskless interest rate and a high rate of time preference, such investors would have a strong desire to borrow from the

[6]Glassman and Hassett (1999) take this argument to an extreme. They argue that the equity premium should be zero, and that US stock prices will rise three-fold from current levels as the transition continues.

[7]Experimental evidence is well described by the prospect theory of Kahneman and Tversky (1979), but it is not at all clear that this theory can be used to describe people's responses to the significant lifetime risks involved in financial markets.

future to reduce their average consumption growth rate. A low riskless interest rate is possible in equilibrium only if investors have a low or negative rate of time preference that reduces their desire to borrow.[8]

Of course, if the risk aversion coefficient γ is high enough then the negative quadratic term $-\gamma^2\sigma_c^2/2$ in equation (13) dominates the linear term and pushes the riskless interest rate down again. The quadratic term reflects precautionary savings; risk-averse agents with uncertain consumption streams have a precautionary desire to save, which can work against their desire to borrow. But a reasonable rate of time preference is obtained only as a knife-edge case.

Table 4 illustrates the riskfree rate puzzle in international data. The table first shows the average riskfree rate from Table 1 and the mean consumption growth rate and standard deviation of consumption growth from Table 2. These moments and the risk aversion coefficients calculated in Table 3 are substituted into equation (13), and the equation is solved for an implied time preference rate. The time preference rate is reported in percentage points per year; it can be interpreted as the riskless real interest rate that would prevail if consumption were known to be constant forever at its current level, with no growth and no volatility. Risk aversion coefficients in the RRA(2) range imply negative time preference rates in every country except Switzerland, whereas larger risk aversion coefficients in the RRA(1) range imply time preference rates that are often positive but always implausible and vary wildly across countries.

The riskfree rate puzzle can be mitigated by use of the recursive preferences suggested by Epstein and Zin (1991) and Weil (1989). These preferences allow the elasticity of intertemporal substitution to be a free parameter, independent of the coefficient of relative risk aversion, whereas power utility forces one to be the reciprocal of the other. The riskfree rate puzzle is caused by a low elasticity of intertemporal substitution rather than a high coefficient of relative risk aversion. Direct evidence on the elasticity of intertemporal substitution (Hall 1988, Campbell and Mankiw 1989) suggests that it is fairly low, certainly well below one, although possibly higher than the reciprocal of risk aversion.

[8] As Abel (1996) and Kocherlakota (1996) point out, negative time preference is consistent with finite utility in a time-separable model provided that consumption is growing, and marginal utility shrinking, sufficiently rapidly. The question is whether negative time preference is plausible.

The equity volatility puzzle

So far I have asked why average stock returns are so high, given their volatility (and behavior of aggregate consumption). Now I ask where the volatility itself comes from.

In order to understand the second moments of stock returns, it is essential to have a framework relating movements in stock prices to movements in expected future dividends and discount rates. The present value model of stock prices is intractably nonlinear when expected stock returns are time-varying, and this has forced researchers to use one of several available simplifying assumptions. The most common approach is to assume a discrete-state Markov process either for dividend growth (Mehra and Prescott 1985) or, following Hamilton (1989), for conditionally expected dividend growth. The Markov structure makes it possible to solve the present value model, but the derived expressions for returns tend to be extremely complicated and so these papers usually emphasize numerical results derived under specific numerical assumptions about parameter values.

An alternative framework, which produces simpler closed-form expressions and hence is better suited for an overview of the literature, is the loglinear approximation to the exact present value model suggested by Campbell and Shiller (1988). Campbell and Shiller's loglinear relation between prices, dividends, and returns provides an accounting framework: High prices must eventually be followed by high future dividends or low future returns, and high prices must be associated with high expected future dividends or low expected future returns. Similarly, high returns must be associated with upward revisions in expected future dividends or downward revisions in expected future returns.

The loglinear approximation starts with the definition of the log return on some asset i, $r_{i,t+1} \equiv \log(P_{i,t+1}+D_{i,t+1})-\log(P_{it})$. The timing convention here is that prices are measured at the end of each period so that they represent claims to next period's dividends. The log return is a nonlinear function of log prices p_{it} and $p_{i,t+1}$ and and log dividends $d_{i,t+1}$, but it can be approximated around the mean log dividend-price ratio, $(\overline{d_{it} - p_{it}})$, using a first-order Taylor expansion. The resulting approximation is

$$r_{i,t+1} \approx k + \rho p_{i,t+1} + (1-\rho) d_{i,t+1} - p_{it} \ , \tag{14}$$

where ρ and k are parameters of linearization defined by $\rho \equiv 1/(1+\exp(\overline{d_{it} - p_{it}}))$ and $k \equiv -\log(\rho) - (1-\rho)\log(1/\rho - 1)$. When the dividend-price ratio is constant, then $\rho = P_i/(P_i + D_i)$, the ratio of the ex-dividend to the cum-dividend stock price.

In the postwar quarterly US data shown in Table 2, the average price-dividend ratio has been 28 on an annual basis, implying that ρ should be about 0.966 in annual data.[9]

The Taylor approximation (14) replaces the log of the sum of the stock price and the dividend in the exact relation with a weighted average of the log stock price and the log dividend. The log stock price gets a weight ρ close to one, while the log dividend gets a weight $1-\rho$ close to zero because the dividend is on average much smaller than the stock price, so a given percentage change in the dividend has a much smaller effect on the return than a given percentage change in the price.

Equation (14) is a linear difference equation for the log stock price. I now solve forward and impose the terminal condition that $\lim_{j\to\infty}\rho^j p_{i,t+j}=0$, effectively ruling out explosive behavior of stock prices relative to dividends (the "rational bubbles" of Blanchard and Watson (1982)).[10] Taking expectations, the implication is that the log dividend-price ratio satisfies

$$d_{it}-p_{it}=\frac{-k}{1-\rho}+\mathrm{E}_t\sum_{j=0}^{\infty}\rho^j\left[r_{i,t+1+j}-\Delta d_{i,t+1+j}\right]. \qquad (15)$$

This equation says that the log dividend-price ratio is high when dividends are expected to grow slowly, or when stock returns are expected to be high. The equation should be thought of as an accounting identity rather than a behavioral model; it has been obtained merely by approximating an identity, solving forward subject to a terminal condition, and taking expectations. Intuitively, if the stock price is high today, then from the definition of the return and the terminal condition that the stock price is non-explosive, there must either be high dividends or low stock returns in the future. Investors must then expect some combination of high dividends and low stock returns if their expectations are to be consistent with the observed price.

[9]Strictly speaking both ρ and k should have asset subscripts i, but I omit these for simplicity.

[10]There are however several reasons to rule out such bubbles. The theoretical circumstances under which bubbles can exist are quite restrictive; Tirole (1985), for example, uses an overlapping generations framework and finds that bubbles can only exist if the economy is dynamically inefficient, a condition which seems unlikely on prior grounds and which is hard to reconcile with the empirical evidence of Abel, Mankiw, Summers, and Zeckhauser (1989). Santos and Woodford (1997) also conclude that the conditions under which bubbles can exist are fragile. Empirically, bubbles imply explosive behavior of prices in relation to dividends and other measures of fundamentals; there is no evidence of this, although nonlinear bubble models are hard to reject using standard linear econometric methods.

Equation (15) describes the log dividend-price ratio rather than the log price itself. This is a useful way to write the model because in many data sets dividends appear to follow a loglinear unit root process, so that log dividends and log prices are nonstationary. In this case changes in log dividends are stationary, so from (15) the log price-dividend ratio is stationary provided that the expected stock return is stationary. Thus log stock prices and dividends are cointegrated, and the stationary linear combination of these variables involves no unknown parameters since it is just the log ratio.

Equation (15) can also be understood as a dynamic generalization of the famous formula, usually attributed to Myron Gordon (1962) but probably due originally to John Burr Williams (1938), that applies when the discount rate is a constant R and the expected dividend growth rate is a constant G:

$$\frac{D}{P} = R - G. \tag{16}$$

So far I have written asset prices as linear combinations of expected future dividends and returns. Following Campbell (1991), I can also write asset returns as linear combinations of *revisions* in expected future dividends and returns. Substituting (15) into (14), I obtain

$$r_{i,t+1} - \mathrm{E}_t\, r_{i,t+1} = (\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{i,t+1+j} - (\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=1}^{\infty} \rho^j r_{i,t+1+j}\ . \tag{17}$$

This equation says that unexpected stock returns must be associated with changes in expectations of future dividends or real returns. An increase in expected future dividends is associated with a capital gain today, while an increase in expected future returns is associated with a capital loss today. The reason is that with a given dividend stream, higher future returns can only be generated by future price appreciation from a lower current price.

I now use this accounting framework to illustrate the stock market volatility puzzle. The intertemporal budget constraint for a representative agent, $W_{t+1} = (1 + R_{p,t+1})(W_t - C_t)$, implies that aggregate consumption is the dividend on the portfolio of all invested wealth, denoted by subscript w:

$$d_{wt} = c_t. \tag{18}$$

Many authors, including Grossman and Shiller (1981), Lucas (1978), and Mehra and Prescott (1985), have assumed that the aggregate stock market, denoted by subscript

e for equity, is equivalent to the wealth portfolio and thus pays consumption as its dividend. Here I follow Campbell (1986) and Abel (1999) and make the slightly more general assumption that the dividend on equity equals aggregate consumption raised to a power λ. In logs, we have

$$d_{et} = \lambda c_t \ . \tag{19}$$

The coefficient λ can be interpreted as a measure of leverage. When $\lambda > 1$, dividends and stock returns are more volatile than the returns on the aggregate wealth portfolio. This framework has the additional advantage that a riskless real bond with infinite maturity—an inflation-indexed consol, denoted by subscript b—can be priced merely by setting $\lambda = 0$. The relative volatility of dividends and consumption suggests that $\lambda = 5$ or 6 might be a reasonable assumption.

The representative-agent asset pricing model with power utility, conditional lognormality, and homoskedasticity implies that

$$\mathrm{E}_t r_{e,t+1} = \mu_e + \gamma \mathrm{E}_t \Delta c_{t+1}, \tag{20}$$

where μ_e is an asset-specific constant term. The expected log return on equity, like the expected log return on any other asset, is just a constant plus relative risk aversion times expected consumption growth.[11]

Substituting equations (19) and (20) into equations (15) and (17), I find that

$$d_{et} - p_{et} = \frac{k_e}{1-\rho} + (\gamma - \lambda)\mathrm{E}_t \sum_{j=0}^{\infty} \rho^j \Delta c_{t+1+j} \ , \tag{21}$$

and

$$r_{e,t+1} - \mathrm{E}_t\, r_{e,t+1} = \lambda(\Delta c_{t+1} - \mathrm{E}_t \Delta c_{t+1}) - (\gamma - \lambda)(\mathrm{E}_{t+1} - \mathrm{E}_t) \sum_{j=1}^{\infty} \rho^j \Delta c_{t+1+j} \ . \tag{22}$$

Expected future consumption growth has offsetting effects on stock prices. It has a direct positive effect by increasing expected future dividends λ-for-one, but it has an indirect negative effect by increasing expected future real interest rates γ-for-one.

These offsetting effects make it almost impossible for the standard power utility model to explain the volatility of stock returns and their positive correlation with

[11] Campbell (1999) analyzes the more general Epstein-Zin-Weil model, where relative risk aversion need not equal the reciprocal of the elasticity of intertemporal substitution ψ. In that model the coefficient on expected consumption growth is actually the reciprocal $1/\psi$.

consumption growth. We already know that the coefficient of relative risk aversion γ must be large to explain the equity premium puzzle. If $\lambda < \gamma$, then good news about future consumption drives down stock prices because the interest-rate effect overwhelms the dividend effect. In this case positively autocorrelated consumption growth implies that stock returns are *negatively* correlated with consumption. If $\lambda = \gamma$, then the dividend-price ratio is constant and the volatility of stock returns is just λ times the volatility of consumption growth. Only if $\lambda > \gamma$ can we get stock returns to be positively correlated with consumption growth, and an implausibly large λ is required to match the observed volatility of stock returns.

Do stock prices forecast dividend or earnings growth?

Of course, all these calculations are dependent on the assumption made at the beginning of this subsection, that the log dividend on stocks is a multiple λ of log aggregate consumption. More general models, allowing separate variation in dividends and consumption, can in principle generate volatile stock returns from predictable variation in dividend growth without creating offsetting variation in real interest rates. But this explanation for stock market volatility requires that the stock market forecasts dividend growth.

Campbell and Shiller (2001) present a simple graphical analysis that makes it clear that stock prices have very little forecasting power for future dividend growth. They point out that if a valuation ratio, such as the dividend-price ratio, is stationary, then when the ratio is at an extreme level either the numerator or the denominator of the ratio must move in a direction that restores the ratio to a more normal level. *Something* must be forecastable based on the ratio, either the numerator or the denominator. In the case of the dividend-price ratio, a high ratio must forecast either slow dividend growth or rapid price growth.[12]

Does the dividend-price ratio forecast future dividend movements or future price movements? To answer this question, Campbell and Shiller use annual US data from 1872 to 2000, and present a pair of scatter plots shown in Figure 1. Each scatterplot has the dividend-price ratio, measured as the previous year's dividend divided by the

[12] A similar point can be understood by looking at equation (21). If the dividend-price ratio varies, then either the expected rate of dividend growth or the expected rate of return must vary. Note however that this is a slightly different point. The total rate of return includes both the dividend yield and the rate of price appreciation. This is why the argument based on equation (21) does not rely on stationarity of the dividend-price ratio. Earlier work on the ability of stock prices to predict dividends includes Shiller (1981) and Campbell and Shiller (1988).

January stock price, on the horizontal axis. (The horizontal axis scale is logarithmic but the axis is labeled in levels for ease of reference.) Over this period the historical mean value for the dividend-price ratio was 4.65%.

In the top part of the figure the vertical axis is the growth rate of real dividends (measured logarithmically as the change in the natural log of real dividends) over a time interval sufficient to bring the dividend-price ratio back to its historical mean of 4.65%. More precisely, the dividend growth rate is measured from the year preceding the year shown until the year before the dividend-price ratio again crossed 4.65%. Because dividends enter the dividend-price ratio with a one-year lag, this is the appropriate way to measure growth in dividends from the base level embodied in a given year's dividend-price ratio to the level that prevailed when the dividend-price ratio next crossed its historical mean.

Since 1872, the dividend-price ratio has crossed its mean value 29 times, with intervals between crossings ranging from one year to twenty years (the twenty-year interval being between 1955 and 1975). The different syears are indicated on the scatter diagram by two-digit numbers; a * after a number denotes a 19th Century date. The last year shown is 1983, since this is the last year that was followed by the dividend-price ratio crossing its mean. (The ratio has been below its mean ever since.) A regression line is fit through these data points, and a vertical line is drawn to indicate the dividend-price ratio at the start of the year 2000. The implied forecast for dividend growth, starting in the year 2000, is the horizontal dashed line marked where the vertical line intersects the regression line.

It is obvious from the top part of Figure 1 that the dividend-price ratio has done a poor job as a forecaster of future dividend growth to the date when the ratio is again borne back to its mean value. The regression line is nearly horizontal, implying that the forecast for future dividend growth is almost the same regardless of the dividend-price ratio. The R^2 statistic for the regression is 0.25%, indicating that only one-quarter of one percent of the variation of dividend growth is explained by the initial dividend-price ratio.

It must follow, therefore, that the dividend-price ratio forecasts movements in its denominator, the stock price, and that it is the stock price that has moved to restore the ratio to its mean value. In the lower part of Figure 1 the vertical axis shows the growth rate of real stock prices (measured logarithmically as the change in log real stock prices) between the year shown and the next year when the dividend-price ratio crossed its mean value. The scatterplot shows a strong tendency for the dividend-price ratio to predict future price changes. The regression line has a strongly positive slope, and the R^2 statistic for the regression is 63%. This answers the question: It is the denominator of the dividend-price ratio that brings the ratio back to its mean, not the numerator.

There are several reasons to be cautious in interpreting the results of Figure 1. First, the behavior of the dividend-price ratio can be altered by shifts in corporate financial policy. A permanent shift towards the use of share repurchases, for example, can reduce current dividends but permanently increase the growth rate of dividends per share by creating a steady decline in the number of shares outstanding. This may have happened in recent years, in which case the low current dividend-price ratio does not necessarily forecast low returns. To address this concern, Campbell and Shiller (2001) look at earnings as well as dividends. To eliminate the effects of short-run cyclical variation in earnings, they average earnings over 10 years as recommended in the classic investment text of Graham and Dodd (1934). They find that the ratio of prices to smoothed earnings predicts price variation rather than earnings variation, consistent with the results just reported for dividends.

Second, the different points in the scatter diagram are not independent of one another. There are not 120 independent observations over 120 years; rather, there are only 29 independent observations corresponding to the 29 occasions on which the dividend-price ratio crossed its mean. If one uses a fixed horizon of 10 years, as Campbell and Shiller do elsewhere in their study, there are only 12 independent observations. Even allowing for this fact, however, the results appear statistically significant in a Monte Carlo study reported by Campbell and Shiller.

Third, the runup in stock prices in the late 1990's diminished the statistical evidence that valuation ratios predict stock returns. For several years in the late 1990's. the stock market delivered high returns despite record low dividend-price ratios. This evidence is not reflected in Figure 1 because the dividend-price ratio has not yet returned to its mean. On the other hand, it is extremely hard to rationalize the runup in prices using a model with a fixed discount rate, because the implied dividend growth forecasts appear wildly optimistic (Heaton and Lucas 1999); also the predictability of dividend growth from the dividend-price ratio does not seem to have increased.[13] For

[13]A more technical point is made by Lewellen (2001). He argues that the dividend-price ratio will tend to spuriously forecast returns when it appears to be less persistent than it truly is. He proposes a correction for this bias under the worst-case assumption that the dividend-price ratio truly has a unit root. Since the dividend-price ratio now appears more persistent than it used to do, Lewellen's

these reasons I believe that the experience of the late 1990's is either an extreme version of previous swings in the stock market, or possibly a one-time structural change to a permanently lower equity premium; in either case it does not alter the overall message of Figure 1.

Conclusion

In this lecture I have presented two puzzles of asset pricing, the equity premium puzzle and the equity volatility puzzle. The equity premium puzzle may in the end be explained by a combination of factors, including both high risk aversion of investors and unexpectedly high returns in the late 20th Century (possibly caused by a one-time correction of historical equity mispricing). The equity volatility puzzle is more fundamental. The data suggest that historical variations in stock prices have been driven primarily by changes in expected stock returns, rather than changes in expected future dividends. Tomorrow I will discuss alternative explanations of equity volatility, and their implications for portfolio management.

bias correction is smaller and the corrected coefficient has not changed much.

INTERNATIONAL STOCK AND BILL RETURNS

Country	Sample Period	$\overline{r_e}$	$\sigma(r_e)$	$\rho(r_e)$	$\overline{r_f}$	$\sigma(r_f)$	$\rho(r_f)$
USA	1947.2 - 1998.4	8.085	15.645	0.083	0.896	1.748	0.508
AUL	1970.1 - 1999.1	3.540	22.699	0.005	2.054	2.528	0.645
CAN	1970.1 - 1999.2	5.431	17.279	0.072	2.713	1.855	0.667
FR	1973.2 - 1998.4	9.023	23.425	0.048	2.715	1.837	0.710
GER	1978.4 - 1997.4	9.838	20.097	0.090	3.219	1.152	0.348
ITA	1971.2 - 1998.2	3.168	27.039	0.079	2.371	2.847	0.691
JAP	1970.2 - 1999.1	4.715	21.909	0.021	1.388	2.298	0.480
NTH	1977.2 - 1998.4	14.070	17.228	-0.030	3.377	1.591	-0.085
SWD	1970.1 - 1999.3	10.648	23.839	0.022	1.995	2.835	0.260
SWT	1982.2 - 1999.1	13.744	21.828	-0.128	1.393	1.498	0.243
UK	1970.1 - 1999.2	8.155	21.190	0.084	1.301	2.957	0.478
USA	1970.1 - 1998.4	6.929	17.556	0.051	1.494	1.687	0.571
SWD	1920 - 1998	7.084	18.641	0.096	2.209	5.800	0.710
UK	1919 - 1998	7.713	22.170	-0.023	1.255	5.319	0.589
USA	1891 - 1998	7.169	18.599	0.047	2.020	8.811	0.338

INTERNATIONAL CONSUMPTION AND DIVIDENDS

Country	Sample Period	$\overline{\Delta c}$	$\sigma(\Delta c)$	$\rho(\Delta c)$	$\overline{\Delta d}$	$\sigma(\Delta d)$	$\rho(\Delta d)$
USA	1947.2 - 1998.4	1.964	1.073	0.216	2.159	28.291	-0.544
AUL	1970.1 - 1999.1	2.099	2.056	-0.324	0.656	34.584	-0.450
CAN	1970.1 - 1999.2	2.082	1.971	0.105	-0.488	5.604	0.522
FR	1973.2 - 1998.4	1.233	2.909	0.029	-0.255	13.108	-0.133
GER	1978.4 - 1997.4	1.681	2.431	-0.327	1.189	8.932	0.078
ITA	1971.2 - 1998.2	2.200	1.700	0.283	-3.100	19.092	0.298
JAP	1970.2 - 1999.1	3.205	2.554	-0.275	-2.350	4.351	0.354
NTH	1977.2 - 1998.4	1.841	2.619	-0.257	4.679	4.973	0.294
SWD	1970.1 - 1999.3	0.962	1.856	-0.266	4.977	14.050	0.386
SWT	1982.2 - 1999.1	0.524	2.112	-0.399	6.052	7.698	0.271
UK	1970.1 - 1999.2	2.203	2.507	-0.006	0.591	7.047	0.313
USA	1970.1 - 1998.4	1.812	0.907	0.374	0.612	16.803	-0.578
SWD	1920 - 1998	1.770	2.816	0.150	1.551	12.894	0.315
UK	1919 - 1998	1.551	2.886	0.294	1.990	7.824	0.233
USA	1891 - 1998	1.789	3.218	-0.116	1.516	14.019	-0.087

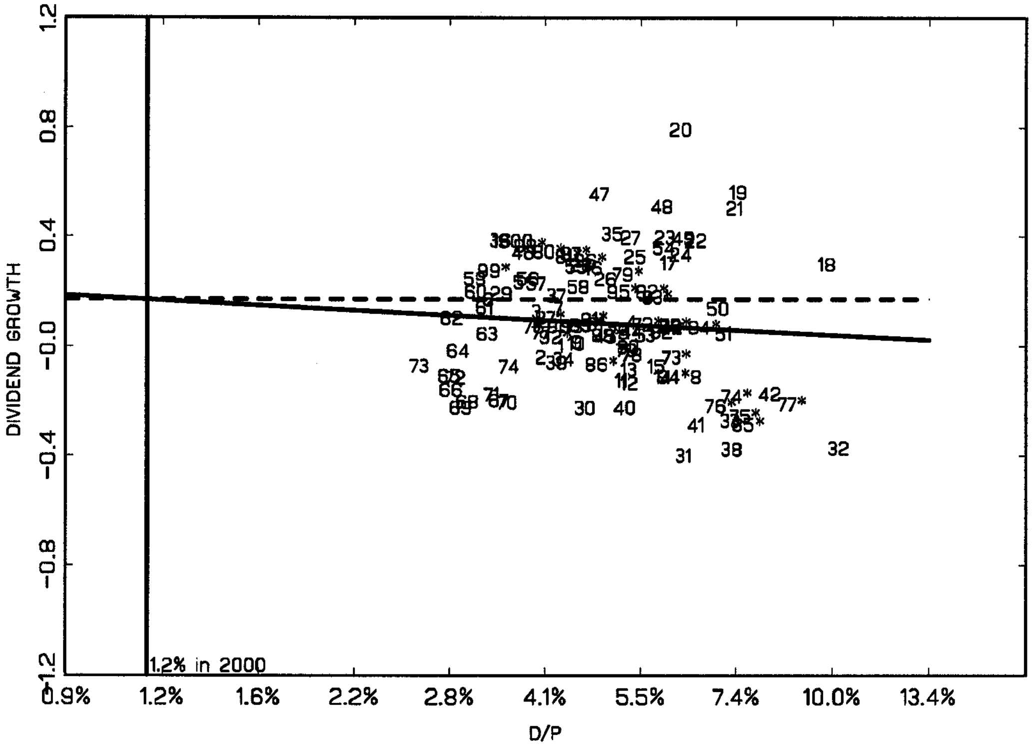
THE EQUITY PREMIUM PUZZLE

Country	Sample Period	$\overline{aer_e}$	$\sigma(er_e)$	$\sigma(m)$	$\sigma(\Delta c)$	$\rho(er_e, \Delta c)$	$cov(er_e, \Delta c)$	RRA(1)	RRA(2)
USA	1947.2 - 1998.3	8.071	15.271	52.853	1.071	0.205	3.354	240.647	49.326
AUL	1970.1 - 1998.4	3.885	22.403	17.342	2.059	0.144	6.640	58.511	8.421
CAN	1970.1 - 1999.1	3.968	17.266	22.979	1.920	0.202	6.694	59.266	11.966
FR	1973.2 - 1998.3	8.308	23.175	35.848	2.922	-0.093	-6.315	$<$ 0	12.270
GER	1978.4 - 1997.3	8.669	20.196	42.922	2.447	0.029	1.446	599.468	17.542
ITA	1971.2 - 1998.1	4.687	27.068	17.314	1.665	-0.006	-0.252	$<$ 0	10.400
JAP	1970.200 - 1998.400	5.098	21.498	23.715	2.561	0.112	6.171	82.620	9.260
NTH	1977.2 - 1998.3	11.421	16.901	67.576	2.510	0.032	1.344	849.991	26.918
SWD	1970.1 - 1999.2	11.539	23.518	49.066	1.851	0.015	0.674	1713.197	26.501
SWT	1982.2 - 1998.4	14.898	21.878	68.098	2.123	-0.112	-5.181	$<$ 0	32.076
UK	1970.1 - 1999.1	9.169	21.198	43.253	2.511	0.093	4.930	185.977	17.222
USA	1970.1 - 1998.3	6.353	16.976	37.425	0.909	0.274	4.233	150.100	41.178
SWD	1920 - 1997	6.540	18.763	34.855	5.622	0.167	8.830	74.062	12.400
UK	1919 - 1997	8.674	21.277	40.767	5.630	0.351	21.042	41.223	14.483
USA	1891 - 1997	6.723	18.496	36.345	6.437	0.495	29.450	22.827	11.293

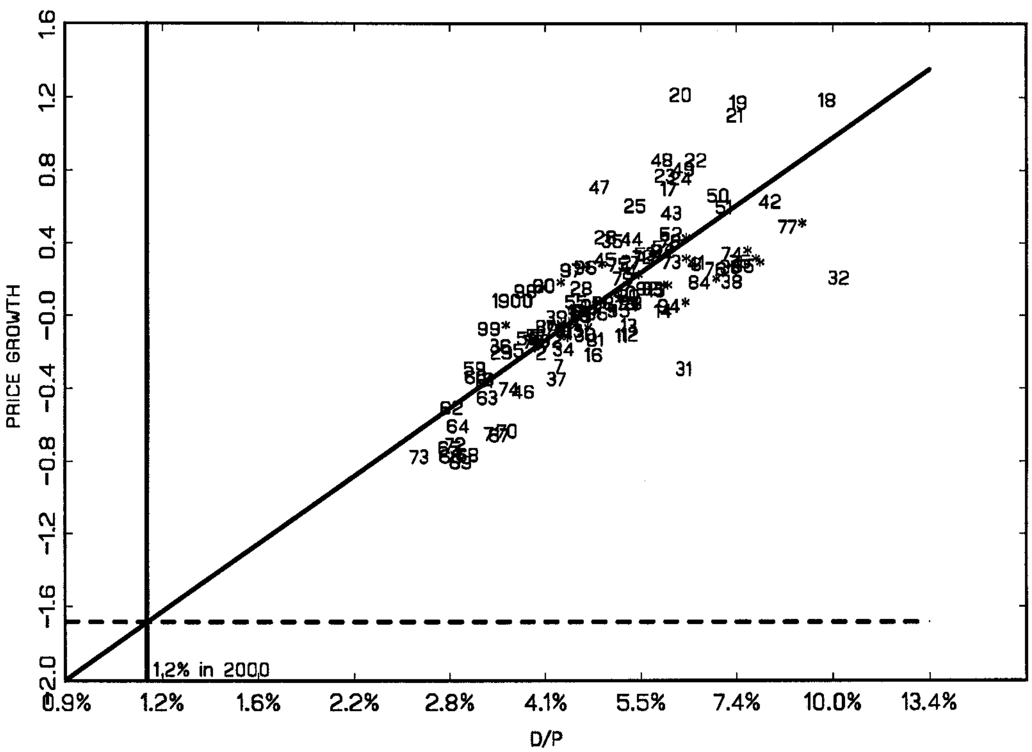
THE RISKFREE RATE PUZZLE

Country	Sample Period	$\overline{r_f}$	$\overline{\Delta c}$	$\sigma(\Delta c)$	RRA(1)	TPR(1)	RRA(2)	TPR(2)
USA	1947.2 - 1998.3	0.896	1.951	1.071	240.647	-136.270	49.326	-81.393
AUL	1970.1 - 1998.4	2.054	2.071	2.059	58.511	-46.512	8.421	-13.880
CAN	1970.1 - 1999.1	2.713	2.170	1.920	59.266	-61.154	11.966	-20.618
FR	1973.2 - 1998.3	2.715	1.212	2.922	$<$ 0	N/A	12.270	-5.735
GER	1978.4 - 1997.3	3.219	1.673	2.447	599.468	9757.265	17.542	-16.910
ITA	1971.2 - 1998.1	2.371	2.273	1.665	$<$ 0	N/A	10.400	-19.765
JAP	1970.2 - 1998.4	1.388	3.233	2.561	82.620	-41.841	9.260	-25.735
NTH	1977.2 - 1998.3	3.377	1.671	2.510	849.991	21349.249	26.918	-18.769
SWD	1970.1 - 1999.2	1.995	1.001	1.851	1713.197	48590.956	26.501	-12.506
SWT	1982.2 - 1998.4	1.393	0.559	2.123	$<$ 0	N/A	32.076	6.636
UK	1970.1 - 1999.1	1.301	2.235	2.511	185.977	676.439	17.222	-27.838
USA	1970.1 - 1998.3	1.494	1.802	0.909	150.100	-175.916	41.178	-65.701
SWD	1920 - 1997	2.209	1.730	2.811	74.062	90.793	12.400	-13.165
UK	1919 - 1997	1.255	1.472	2.815	41.223	7.913	14.483	-11.749
USA	1891 - 1997	2.020	1.760	3.218	22.827	-11.162	11.293	-11.247

Figure 1. DIVIDEND GROWTH till the next time D/P crosses its mean



PRICE GROWTH till the next time D/P crosses its mean



Understanding Risk and Return

2001 Marshall Lectures
University of Cambridge

John Y. Campbell
Harvard University

Lecture 2: From Puzzles to Portfolios

May 9, 2001

Yesterday I presented two puzzles of asset pricing, the equity premium puzzle and the equity volatility puzzle. Several solutions to the equity premium are potentially available. For example, investors may have higher risk aversion than economists used to think; returns may have been unusually high in the late 20th Century; and these high returns may have been caused in part by a one-time correction of historical equity mispricing. In this case future returns will tend to be lower than historical returns, and the equity premium will diminish as a focus of academic attention.

The situation is not so favorable with respect to the equity volatility puzzle. This puzzle raises fundamental questions about the relationship between aggregate consumption and aggregate wealth. Since consumption is ultimately financed by wealth (broadly defined to include human wealth), any model with stationary asset returns implies that the ratio of consumption to wealth must be stationary. Since consumption and wealth appear individually to have unit roots, this implies that consumption and wealth are cointegrated. In the very long run, then, the annualized growth rates of consumption and wealth must be almost identical; in particular, they must have identical volatilities. The difficulty is that in the short run, the volatility of consumption growth is far smaller than the volatility of wealth growth.[14] Consumption is very smooth, while wealth is very volatile.

How can we reconcile the observed short-run properties of consumption and wealth with the properties we know they must have in the long run? There are only two possibilities. First, it may be that the annualized volatility of consumption growth increases with the horizon over which it is measured, so that ultimately it reaches the high volatility of wealth growth. This would require that consumption is not a random walk, but has positive serial correlation in growth rates. Second, it may be that the annualized volatility of wealth growth decreases with the horizon over which it is measured, so that ultimately it reaches the low volatility of consumption growth. This would require that wealth is not a random walk, but has negative serial correlation in growth rates. These two possibilities represent fundamentally different views of the world. Is the world safe as suggested by consumption, or risky as suggested by the stock market?

Recent work of Lettau and Ludvigson (2001) suggests that consumption, not wealth, accurately represents long-term risk. Lettau and Ludvigson use US Flow of

[14] Yesterday we saw that consumption is far less volatile than stock returns. While equities are not the only component of wealth, other components are not smooth enough to compensate for the volatility of stock returns (Campbell 1996, Lettau and Ludvigson 2001).

Funds data to construct a proxy for total asset wealth, including not only equities but also other assets such as real estate. They use labor income to proxy for human wealth, arguing that labor income and human wealth should be cointegrated. They analyze the three aggregate time series for consumption, labor income, and asset wealth, and find that the three are cointegrated (even though no two of them are cointegrated). The stationary linear combination of these variables forecasts wealth, not consumption or labor income. In their data consumption is extremely close to a random walk. Thus Lettau and Ludvigson find that wealth is mean-reverting and adjusts over long horizons to match the smoothness of consumption. A satisfactory model of equity volatility must be consistent with this finding.

What might explain equity volatility?

The loglinear asset pricing framework of Campbell and Shiller (1988) and Campbell (1991), discussed yesterday as equation (17), allows us to divide explanations for equity volatility into several categories. First, equity volatility might be caused by predictable variation in dividend growth (equivalent to predictable variation in consumption growth if equities are modelled as consumption claims, that is, as proxies for aggregate wealth). The empirical difficulty with this explanation, discussed yesterday, is that stock prices are not good forecasters of consumption or dividend growth. There is also a theoretical difficulty that predictable variation in consumption growth should cause offsetting movements in real interest rates that dampen the effect on stock prices.

If equity volatility is not caused by predictable variation in cash flows, then it must be caused by variation in discount rates. The first and most obvious component of the equity discount rate is the riskless real interest rate. There is some variation in the real interest rate; unfortunately it is not large enough to cause big swings in the stock market as pointed out by Campbell (1991). Also, the timing of real interest rate movements seems to be quite different from the timing of stock market movements. The 1970's, for example, saw low real rates and a depressed stock market, whereas the 1980's saw much higher real rates and a buoyant stock market. For this reason stock prices are not good forecasters of real interest rates (Campbell 1999).

The remaining component of the equity discount rate is the equity premium, the expected excess return on stocks over short-term debt. As we saw yesterday, valuation ratios in the stock market have historically predicted stock returns over long horizons, consistent with the view that stock market movements are driven by movements in the equity premium itself.

The equity premium can be thought of as volatility times the reward for bearing volatility, or the quantity of risk times the price of risk. Equity volatility does move over time, and does correlate positively with return forecasts, rising during recessions and stock market declines. However these movements of volatility are not proportional to movements in returns as pointed out by Campbell (1987) and Harvey (1989). Thus we are forced inexorably to the conclusion that the price of risk itself must be moving over time. Since stock prices tend to increase when the economy is strong and consumption is growing rapidly, the price of risk must be countercyclical, moving opposite to consumption growth. Most of this lecture will be devoted to alternative structural models that can generate countercyclical time-variation in the price of risk.

One class of models works within a representative-investor framework and asks what preferences might generate countercyclical risk aversion. Models of habit formation, such as Constantinides (1990) and Campbell and Cochrane (1999), have this property, and I discuss these models in detail in the next section. Countercylical risk aversion also arises naturally in behavioral finance models that combine the prospect theory of Kahneman and Tversky (1979) with the "house money effect" of Thaler and Johnson (1990), that is, the tendency of investors to worry less about losses that offset prior gains (Barberis, Huang, and Santos 2001).

A second class of models emphasizes the aggregation of heterogeneous agents. Each individual agent might have constant risk aversion, yet they might interact in such a way that the representative agent has time-varying risk aversion. Different models emphasize different types of heterogeneity. There might be heterogeneous constraints, so that some investors are constrained from stock market participation or are prevented from diversifying their stock portfolios (Constantinides, Donaldson, and Mehra 1998, Heaton and Lucas 1999, Vissing-Jorgensen 1999). A relaxation of such constraints allows equity risk to be shared more broadly, driving down the equilibrium price of risk. This story might explain a one-time decline in the equity premium in the late 20th Century, but is less suitable for explaining recurring cyclical variation in the price of risk.[15]

Investors might also have heterogeneous uninsurable labor income (Constantinides and Duffie 1996). Variation in the degree of idiosyncratic risk can cause a high and possibly time-varying equity premium. Heterogeneous risk aversion may also be

[15]Even on a one-time basis, it is hard to get large effects of expanding participation because new participants tend to be much poorer than old participants, so the wealth-weighted expansion in participation is relatively small.

important (Wang 1996, Chan and Kogan 2000). In this case high stock returns would tend to increase the wealth of risk-tolerant investors, increasing their weight in the aggregate and driving down the risk-aversion of the representative investor.

A third class of models emphasizes irrational expectations on the part of at least some investors. Hansen, Sargent, and Tallarini (1997) have emphasized that pessimism about long-run growth prospects can explain both the equity premium and riskfree rate puzzles.[16] The extrapolation of shocks to growth rates ("irrational exuberance" if the shocks are positive and irrational gloom if they are negative) can generate a time-varying and countercyclical price of risk (Barsky and De Long 1993, Barberis, Shleifer, and Vishny 1998, Cecchetti, Lam, and Mark 2000, Shiller 2000).

Habit formation

Sundaresan (1989) and Constantinides (1990) have argued for the importance of habit formation, a positive effect of today's consumption on tomorrow's marginal utility of consumption.

Two modeling issues arise at the outset. Writing the period utility function as $U(C_t, X_t)$, where X_t is the time-varying habit or subsistence level, the first issue is the functional form for $U(\cdot)$. Abel (1990) has proposed that $U(\cdot)$ should be a power function of the ratio C_t/X_t, while most other researchers have used a power function of the difference $C_t - X_t$. The second issue is the effect of an agent's own decisions on future levels of habit. In standard "internal habit" models such as those in Constantinides (1990) and Sundaresan (1989), habit depends on an agent's own consumption and the agent takes account of this when choosing how much to consume. In "external habit" models such as those in Abel (1990) and Campbell and Cochrane (1999), habit depends on aggregate consumption which is unaffected by any one agent's decisions. Abel calls this "catching up with the Joneses". Similar results can be obtained in either class of model, but external habit models are generally easier to work with.

The choice between ratio models and difference models of habit is important because ratio models have constant risk aversion whereas difference models have time-varying risk aversion. In Abel's (1990) ratio model, external habit adds a term to the equation describing the riskless interest rate, but does not change the equation

[16]This is similar to the peso problem story of Rietz (1988) except that investor fears are no longer required to be rational.

that describes the excess return of risky assets over the riskless interest rate. The effect on the riskless interest rate has to do with intertemporal substitution. Holding consumption today and expected consumption tomorrow constant, an increase in consumption yesterday increases the marginal utility of consumption today. This makes the representative agent want to borrow from the future, driving up the real interest rate.

This instability of the riskless real interest rate is a fundamental problem for habit formation models. Time-nonseparable preferences make marginal utility volatile even when consumption is smooth, because consumers derive utility from consumption relative to its recent history rather than from the absolute level of consumption. But unless the consumption and habit processes take particular forms, time-nonseparability also creates large swings in expected marginal utility at successive dates, and this implies large movements in the real interest rate. I now present an alternative specification in which it is possible to solve this problem, and in which risk aversion varies over time.

Campbell and Cochrane (1999) build a model with external habit formation in which a representative agent derives utility from the difference between consumption and a time-varying subsistence or habit level. They assume that log consumption follows a random walk with mean g and innovation ϵ_{t+1}. This is a fairly good approximation for U.S. data. The utility function of the representative agent is a time-separable power utility function, with curvature γ, of the difference between consumption C_t and habit X_t. Utility is only defined when consumption exceeds habit.

It is convenient to capture the relation between consumption and habit by the surplus consumption ratio S_t, defined by

$$S_t \equiv \frac{C_t - X_t}{C_t}. \tag{23}$$

The surplus consumption ratio is the fraction of consumption that exceeds habit and is therefore available to generate utility. The SDF in this model is given by

$$M_{t+1} = \delta \left(\frac{S_{t+1}}{S_t}\right)^{-\gamma} \left(\frac{C_{t+1}}{C_t}\right)^{-\gamma}. \tag{24}$$

The SDF is driven by proportional innovations in the surplus consumption ratio, as well as by proportional innovations in consumption. If the surplus consumption ratio is only a small fraction of consumption, then small shocks to consumption can be large shocks to the surplus consumption ratio; thus the SDF can be highly volatile even when consumption is smooth.

Even more important, the volatility of the SDF is itself time-varying since it depends on the level of the surplus consumption ratio. Shocks to consumption have a larger proportional effect on S_t when S_t is small than when it is large:

$$\frac{C}{S}\frac{dS}{dC} = \frac{1-S}{S}. \tag{25}$$

Hence investors are more averse to consumption risk when S_t is small. If habit X_t is held fixed as consumption C_t varies, the local coefficient of relative risk aversion is

$$\frac{-Cu_{CC}}{u_C} = \frac{\gamma}{S_t}, \tag{26}$$

where u_C and u_{CC} are the first and second derivatives of utility with respect to consumption. Risk aversion rises as the surplus consumption ratio S_t declines, that is, as consumption approaches the habit level. Note that γ, the curvature parameter in utility, is no longer the coefficient of relative risk aversion in this model.

To complete the description of preferences, one must specify how the habit X_t evolves over time in response to aggregate consumption. Campbell and Cochrane suggest an AR(1) model for the log surplus consumption ratio, $s_t \equiv \log(S_t)$:

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda\left(s_t\right)\epsilon_{t+1}\ . \tag{27}$$

The parameter ϕ governs the persistence of the log surplus consumption ratio, while the "sensitivity function" $\lambda(s_t)$ controls the sensitivity of s_{t+1} and thus of log habit x_{t+1} to innovations in consumption growth ϵ_{t+1}. This modelling strategy ensures that the habit process implied by a process for s_{t+1} always lies below consumption.

The logic of Hansen and Jagannathan (1991) implies that the largest possible Sharpe ratio is given by the conditional standard deviation of the log SDF. This is $\gamma\sigma(1+\lambda(s_t))$, so a sensitivity function that varies inversely with s_t delivers a time-varying, countercyclical Sharpe ratio.

The same mechanism helps to stabilize the riskless real interest rate. When the surplus consumption ratio falls, investors have an intertemporal-substitution motive to borrow from the future, but this is offset by an increased precautionary savings motive created by the volatility of the SDF. Campbell and Cochrane parameterize the model so that these two effects exactly cancel. This makes the riskless real interest rate constant, a knife-edge case that helps to reveal the pure effects of time-varying risk aversion on asset prices. With a constant riskless rate, real bonds of all maturities are also riskless and there are no real term premia. Thus the equity premium is also a premium of stocks over long-term bonds.

When this model is calibrated to fit the first two moments of consumption growth, the average riskless interest rate, and the Sharpe ratio on the stock market, it also roughly fits the volatility, predictability, and cyclicality of stock returns. The model does not resolve the equity premium puzzle, since it relies on high average risk aversion, but it does resolve the stock market volatility puzzle.

The Campbell-Cochrane model assumes random walk consumption and implies negative autocorrelation of stock returns. The Constantinides (1990) model of habit formation, by contrast, assumes IID asset returns and implies positive autocorrelation of consumption growth. Thus these two models take different stands on the question of whether wealth or consumption accurately represents long-run risk. The Constantinides model fits the equity premium with low risk-aversion, but it achieves this success at the cost of a positively serially correlated consumption process that contradicts the empirical findings of Lettau and Ludvigson.

Heterogeneous labor income

The heterogeneity of utility-maximizing stock market investors may have important effects. For example, if investors are subject to large idiosyncratic risks in their labor income and can share these risks only indirectly by trading a few assets such as stocks and Treasury bills, their individual consumption paths may be much more volatile than aggregate consumption. Even if individual investors have the same power utility function, so that any individual's consumption growth rate raised to the power $-\gamma$ would be a valid stochastic discount factor, the aggregate consumption growth rate raised to the power $-\gamma$ may not be a valid stochastic discount factor.

This problem is an example of Jensen's Inequality. Since marginal utility is nonlinear, the average of investors' marginal utilities of consumption is not generally the same as the marginal utility of average consumption. The problem disappears when investors' individual consumption streams are perfectly correlated with one another as they will be in a complete markets setting. Grossman and Shiller (1982) point out that it also disappears in a continuous-time model when the processes for individual consumption streams and asset prices are diffusions.

Recently Constantinides and Duffie (1996) have provided a simple framework within which the effects of heterogeneity can be understood. Constantinides and Duffie postulate an economy in which individual investors k have different consumption levels C_{kt}. The cross-sectional distribution of individual consumption is lognormal, and the change from time t to time $t+1$ in individual log consumption is cross-sectionally uncorrelated with the level of individual log consumption at time t. All investors have the same power utility function with time discount factor δ and coefficient of relative risk aversion γ.

In this economy each investor's own intertemporal marginal rate of substitution is a valid stochastic discount factor. Hence the cross-sectional average of investors' intertemporal marginal rates of substitution is a valid stochastic discount factor. I write this as

$$M^*_{t+1} \equiv \delta \mathrm{E}^*_{t+1} \left[\left(\frac{C_{k,t+1}}{C_{kt}} \right)^{-\gamma} \right], \tag{28}$$

where E^*_t denotes an expectation taken over the cross-sectional distribution at time t. That is, for any cross-sectionally random variable X_{kt}, $\mathrm{E}^*_t X_{kt} \equiv \lim_{K\to\infty}(1/K)\sum_{k=1}^{K} X_{kt}$, the limit as the number of cross-sectional units increases of the cross-sectional sample average of X_{kt}. Note that $\mathrm{E}^*_t X_{kt}$ will in general vary over time and need not be lognormally distributed conditional on past information.

The assumption of cross-sectional lognormality means that the log stochastic discount factor, $m^*_{t+1} \equiv \log(M^*_{t+1})$, can be written as a function of the cross-sectional mean and variance of the change in log consumption:

$$m^*_{t+1} = -\log(\delta) - \gamma \mathrm{E}^*_{t+1} \Delta c_{k,t+1} + \left(\frac{\gamma^2}{2} \right) \mathrm{Var}^*_{t+1} \Delta c_{k,t+1}, \tag{29}$$

where Var^*_t is defined analogously to E^*_t as $\mathrm{Var}^*_t X_{kt} = \lim_{K\to\infty}(1/K)\sum_{k=1}^{K}(X_{kt} - \mathrm{E}^*_t X_{kt})^2$, and like E^*_t will in general vary over time.

An economist who knows the underlying preference parameters of investors but does not understand the heterogeneity in this economy might attempt to construct a representative-agent stochastic discount factor, M^{RA}_{t+1}, using aggregate consumption:

$$M^{RA}_{t+1} \equiv \delta \left(\frac{\mathrm{E}^*_{t+1}[C_{k,t+1}]}{\mathrm{E}^*_t[C_{kt}]} \right)^{-\gamma}. \tag{30}$$

The log of this stochastic discount factor can also be related to the cross-sectional mean and variance of the change in log consumption:

$$\begin{aligned} m^{RA}_{t+1} &= -\log(\delta) - \gamma \mathrm{E}^*_{t+1}\Delta c_{k,t+1} - \left(\frac{\gamma}{2}\right)[\mathrm{Var}^*_{t+1}c_{k,t+1} - \mathrm{Var}^*_t c_{kt}] \\ &= -\log(\delta) - \gamma \mathrm{E}^*_{t+1}\Delta c_{k,t+1} - \left(\frac{\gamma}{2}\right)[\mathrm{Var}^*_{t+1}\Delta c_{k,t+1}], \end{aligned} \tag{31}$$

where the second equality follows from the relation $c_{k,t+1} = c_{kt} + \Delta c_{k,t+1}$ and the fact that $\Delta c_{k,t+1}$ is cross-sectionally uncorrelated with c_{kt}.

The difference between these two variables can now be written as

$$m^*_{t+1} - m^{RA}_{t+1} = \frac{\gamma(\gamma+1)}{2}\mathrm{Var}^*_{t+1}\Delta c_{k,t+1}. \tag{32}$$

The time series of this difference can have a nonzero mean, helping to explain the riskfree rate puzzle, and a nonzero variance, helping to explain the equity premium puzzle. If the cross-sectional variance of log consumption growth is negatively correlated with the level of aggregate consumption, so that idiosyncratic risk increases in economic downturns, then the true stochastic discount factor m^*_{t+1} will be more strongly countercyclical than the representative-agent stochastic discount factor constructed using the same preference parameters; this has the potential to explain the high price of risk without assuming that individual investors have high risk aversion. Mankiw (1986) makes a similar point in a two-period model. It is also possible that the correlation between idiosyncratic risk and aggregate consumption itself moves over time in such a way that the price of risk is time-varying.

An important unresolved question is whether the heterogeneity we can measure has the characteristics that are needed to help resolve the asset pricing puzzles. In the Constantinides-Duffie model the heterogeneity must be large to have important effects on the stochastic discount factor; a cross-sectional standard deviation of log consumption growth of 20 percent, for example, is a cross-sectional variance of only 0.04, and it is variation in this number over time that is needed to explain the equity premium puzzle. Interestingly, the effect of heterogeneity is strongly increasing in risk aversion since $\mathrm{Var}^*_{t+1}\Delta c_{k,t+1}$ is multiplied by $\gamma(\gamma+1)/2$ in (32). This suggests that heterogeneity may supplement high risk aversion but cannot altogether replace it as an explanation for the equity premium puzzle.

Cogley (1998) looks at consumption data and finds that measured heterogeneity has only small effects on the SDF. Lettau (1997) reaches a similar conclusion by assuming that individuals consume their income, and calculating the risk-aversion coefficients needed to put model-based stochastic discount factors inside the Hansen-Jagannathan volatility bounds. This procedure is conservative in that individuals trading in financial markets are normally able to achieve some smoothing of consumption relative to income. Nevertheless Lettau finds that high individual risk aversion is still needed to satisfy the Hansen-Jagannathan bounds.

These conclusions may not be surprising given the Grossman-Shiller (1982) result that the aggregation problem disappears in a continuous-time diffusion model. In such a model, the cross-sectional variance of consumption is locally deterministic and hence the false SDF M_{t+1}^{RA} correctly prices risky assets. In a discrete-time model the cross-sectional variance of consumption can change randomly from one period to the next, but in practice these changes are likely to be small. This limits the effects of consumption heterogeneity on asset pricing.

It is also important to note that idiosyncratic shocks are assumed to be permanent in the Constantinides-Duffie model. Heaton and Lucas (1996) calibrate individual income processes to micro data from the Panel Study of Income Dynamics (PSID). Because the PSID data show that idiosyncratic income variation is largely transitory, Heaton and Lucas find that investors can minimize its effects on their consumption by borrowing and lending. This prevents heterogeneity from having any large effects on aggregate asset prices.

To get around this problem, several recent papers have combined heterogeneity with constraints on borrowing. Heaton and Lucas (1996) and Krusell and Smith (1997) find that borrowing constraints or large costs of trading equities are needed to explain the equity premium. Constantinides, Donaldson, and Mehra (1998) focus on heterogeneity across generations. In a stylized three-period overlapping generations model young agents have the strongest desire to hold equities because they have the largest ratio of labor income to financial wealth. If these agents are prevented from borrowing to buy equities, the equilibrium equity premium is large.

Heterogeneity in preferences may also be important. Several authors have recently argued that trading between investors with different degrees of risk aversion or time preference, possibly in the presence of market frictions or portfolio insurance constraints, can lead to time-variation in the market price of risk (Dumas (1989), Grossman and Zhou (1996), Wang (1996), Chan and Kogan (1999)). Intuitively, risk-tolerant agents hold more risky assets so they control a greater share of wealth in good states than in bad states; aggregate risk aversion therefore falls in good states, producing effects similar to those of habit formation.

Irrational expectations

A number of papers have explored the consequences of relaxing the assumption that investors have rational expectations and understand the behavior of dividend and consumption growth. In the absence of arbitrage, there exist positive state prices that can rationalize the prices of traded financial assets. These state prices equal subjective state probabilities multiplied by ratios of marginal utilities in different states. Thus given any model of utility, there exist subjective probabilities that produce the necessary state prices and in this sense explain the observed prices of traded financial assets. The interesting question is whether these subjective probabilities are sufficiently close to objective probabilities, and sufficiently related to known psychological biases in behavior, to be plausible.

Many of the papers in this area work in partial equilibrium and assume that stocks are priced by discounting expected future dividends at a constant rate. This assumption makes it easy to derive any desired behavior of stock prices directly from assumptions on dividend expectations. Barsky and De Long (1993), for example, assume that investors believe dividends to be generated by a doubly integrated process, so that the dividend growth rate has a unit root. These expectations imply that rapid dividend growth increases stock prices more than proportionally, so that the price-dividend ratio rises when dividends are growing strongly. If dividend growth is in fact stationary, then the high price-dividend ratio is typically followed by dividend disappointments, low stock returns, and reversion to the long-run mean price-dividend ratio. Under this assumption of stationary dividend growth, Barsky and DeLong's model produces overreaction of stock prices to dividend news, and this accounts for the equity volatility puzzle and the predictability of stock returns.[17]

Another potentially important form of irrationality is a failure to understand the difference between real and nominal magnitudes. Modigliani and Cohn (1979) argued that investors suffer from inflation illusion, in effect discounting real cash flows at nominal interest rates. Ritter and Warr (1999) and Sharpe (1999) argue that inflation illusion may have led investors to bid up stock prices as inflation has declined since the early 1980s. An interesting issue raised by this literature is whether misvaluation is

[17]Shiller (2000) discusses psychological factors that contribute to the formation of extrapolative expectations, with special reference to the runup in stock prices during the 1990's. Barberis, Shleifer, and Vishny (1998) present a related model.

caused by a high *level* of inflation (in which case it is unlikely to be important today) or whether it is caused by *changes* in inflation from historical benchmark levels (in which case it may contribute to high current levels of stock prices).

A limitation of these models is that they do not consider general equilibrium issues, in particular the implication of irrational beliefs for aggregate consumption. Using for simplicity the fiction that dividends equal consumption, investors' irrational expectations about dividend growth should be linked to their irrational expectations about consumption growth. Interest rates are not exogenous, but like stock prices, are determined by investors' expectations. Thus it is significantly harder to build a general equilibrium model with irrational expectations.

To see how irrationality can affect asset prices in general equilibrium, consider first a static model in which log consumption follows a random walk with drift. Investors understand that consumption is a random walk, but they underestimate its drift. Such irrational pessimism lowers the average riskfree rate, increases the equity premium, and has an ambiguous effect on the price-dividend ratio. Thus pessimism has the same effects on asset prices as a low rate of time preference and a high coefficient of risk aversion, and it can help to explain both the riskfree rate puzzle and the equity premium puzzle (Hansen, Sargent, and Tallarini 1997).

To explain the volatility puzzle, a more complicated model of irrationality is needed. Suppose now that log consumption growth follows an AR(1) process, but that investors overestimate the persistence of this process. In this model the equity premium falls when consumption growth has been rapid, and rises when consumption growth has been weak. This model, which can be seen as a general equilibrium version of Barsky and De Long (1993) or Shiller (2000) fits the apparent cyclical variation in the market price of risk. One difficulty with this story is that it has strong implications for bond market behavior. When investors become irrationally exuberant, their optimism should lead to a strong desire to borrow from the future, which should drive up the riskless interest rate even while it drives down the equity premium. Cecchetti, Lam, and Mark (2000) handle this problem by allowing the degree of investors' irrationality itself to be stochastic and time-varying.

Implications for portfolio choice

I have argued that the price of risk is time varying. It follows that a rational investor, who lives entirely off financial wealth without idiosyncratic labor income, must have time-varying risk aversion in order to buy and hold an aggregate equity index. This leads naturally to the question, what should a rational investor do if he lives off financial wealth and has constant risk aversion?

This topic of portfolio choice is the original subject of modern financial economics. Mean-variance analysis, developed almost fifty years ago by Markowitz (1952), has provided a basic paradigm for portfolio choice. This approach usefully emphasizes the ability of diversification to reduce risk, but it ignores several critically important factors. Most notably, the analysis is static; it assumes that investors care only about risks to wealth one period ahead. However many investors, both individuals and institutions such as charitable foundations or universities, seek to finance a stream of consumption over a long lifetime.

Merton (1969, 1971, 1973) showed thirty years ago that the solution to a long-term portfolio choice problem can be very different from the solution to a short-term problem. In particular, if investment opportunities are varying over time, then long-term investors care about shocks to investment opportunities—the productivity of wealth—as well as shocks to wealth itself. They may seek to hedge their exposures to wealth productivity shocks, and this gives rise to intertemporal hedging demands for financial assets. Brennan, Schwartz, and Lagnado (1997) have coined the phrase "strategic asset allocation" to describe this far-sighted response to time-varying investment opportunities.

Unfortunately Merton's intertemporal model is hard to solve. Until recently solutions to the model were only available in those trivial cases where it reduces to the static model. Therefore the Merton model has not become a usable empirical paradigm, has not displaced the Markowitz model, and has had only limited influence on investment practice. Recently this situation has begun to change as a result of advances in both analytical and numerical methods. A new empirical paradigm is emerging. Interestingly, this paradigm both supports and qualifies traditional rules of thumb used by financial planners. Campbell and Viceira (1999, 2001, 2002) present an integrated empirical approach to the recent portfolio choice literature.

Time-variation of the equity premium has two effects on optimal portfolio choice for investors with constant risk aversion. First, it implies that investors should "time the market", increasing their equity allocations at times when the equity premium is high and reducing them at times when the equity premium is low.[18]

[18] Note that these adjustments take place gradually, since the variables that predict the equity premium move relatively slowly. Thus they are nothing like the rapid moves that are sometimes

A second effect on portfolio choice arises from the fact that the equity premium tends to fall when stock prices rise, because valuation ratios such as D/P move inversely with prices. This implies mean-reversion in stock returns, that is, a tendency for the annualized volatility of returns to fall with the investment horizon. Direct evidence for reduction in volatility at long horizons is presented by Siegel (1998). Campbell and Viceira (2002) use a simple time-series model, related to the evidence presented yesterday on stock return predictability, to generate implied volatilities of returns on stocks, bonds, and Treasury bills at all horizons. Their results are summarized graphically in figures 1 and 2 for quarterly postwar and long-term annual US data. In both data sets equity volatility is in the range 16% to 18% over one year, but it falls to 9% in the quarterly data and 13% in the annual data over longer holding periods. The volatility of Treasury bill investments, by contrast, increases with the holding period because real interest rates vary over time in a persistent fashion.

Mean-reversion in stock returns creates a horizon effect on portfolio choice: Long-term investors may invest differently from short-term investors. The reduction in long-term stock market risk is directly relevant for long-term buy-and-hold investors (Barberis 2000). These investors will increase their equity holdings, relative to the holdings of otherwise identical short-term investors, because they perceive equities as having lower risk.

Long-term investors who can rebalance their portfolios each period have intertemporal hedging demand (Merton 1973). They may wish to hedge the risk that future investment opportunities will deteriorate. If their risk aversion is greater than one, they wish to hold assets that increase in value when investment opportunities deteriorate. The most obvious example of such an asset is an inflation-indexed bond, whose value increases when real interest rates fall. But stocks also have this property, because an increase in stock prices signals a decrease in future stock returns and thus a deterioration in investment opportunities (Campbell and Viceira 1999).

Figure 3 illustrates alternative portfolio rules. The horizontal axis shows the equity premium, with its long-run average marked by a vertical dashed line. In the presence of mean-reversion, the equity premium will fall if stock prices have risen (one will move to the left in the diagram) and will rise if stock prices have fallen (one will move to the right). The vertical axis shows the portfolio allocation to stocks, assuming that the alternative is to hold cash at a constant riskless interest rate and that there are no constraints on leverage or short sales.

recommended by commercial market timing or tactical asset allocation models.

The three lines in the figure are three alternative portfolio rules. The horizontal line marked "Myopic Investor" is the traditional buy-and-hold allocation that would come out of a single-period mean-variance analysis, ignoring time-variation in the equity premium. The sloped line marked "Tactical Investor" is the allocation that would be recommended by single-period mean-variance analysis that takes account of time-variation in the equity premium, in the manner of commercial tactical asset allocation strategies. This line passes through the origin, because an equity premium of zero would imply zero allocation to stocks. The sloped line marked "Strategic Investor" is the optimal portfolio rule derived by Campbell and Viceira (1999) for long-term investors with constant relative risk aversion greater than one. It has almost the same slope as the tactical portfolio rule (if anything it is slightly steeper), but it is shifted upward by positive intertemporal hedging demand. A strategic investor should hold some equities even if the equity premium temporarily dips to zero, in order to hedge against further deterioration in investment opportunities.

It is interesting to relate these results to recent discussions of stock market risk. Equities have traditionally been regarded as risky assets. They may be attractive because of their high average returns, but these returns represent compensation for risk; thus equities should be treated with caution by all but the most aggressive investors. In recent years, however, authors such as Siegel (1998) and Glassman and Hassett (1999) have argued that equities are actually relatively safe assets for investors who are able to hold for the long term.

The revisionist view that stocks are safe assets is based on the evidence that excess stock returns are less volatile when they are measured over long holding periods. Mathematically, such a reduction in stock market risk at long horizons can only be due to mean-reversion in excess stock returns, which is equivalent to time-variation in the equity premium. Yet revisionist investment advice typically ignores the implications of a time-varying equity premium. Siegel (1998) recommends an aggressive buy-and-hold strategy, like the horizontal line in Figure 3 but shifted upwards to reflect the reduced risk of stocks for long-term investors. The optimal policy is instead the sloped line marked "Strategic Investor" in Figure 3.

The difference between the optimal strategy and the strategy recommended by Siegel is particularly dramatic at times like the present, when recent stock returns have been spectacular. At such a time, the optimal equity allocation may be no higher—it may even be lower—than the allocation implied by a traditional short-term portfolio analysis. To put it another way, investors who are attracted to the stock market by the prospect of high returns combined with low long-term risk are trying to have their cake and eat it too. If expected stock returns are constant over time, then one can hope to earn high stock returns in the future similar to the high returns of the past; but in this case stocks are much riskier than bonds in the long term, just as they are in the short term. If instead stocks mean-revert, then they are relatively safe assets for long-term investors; but in this case future returns are likely to be meagre as mean-reversion unwinds the spectacular stock market runup of the past decade.

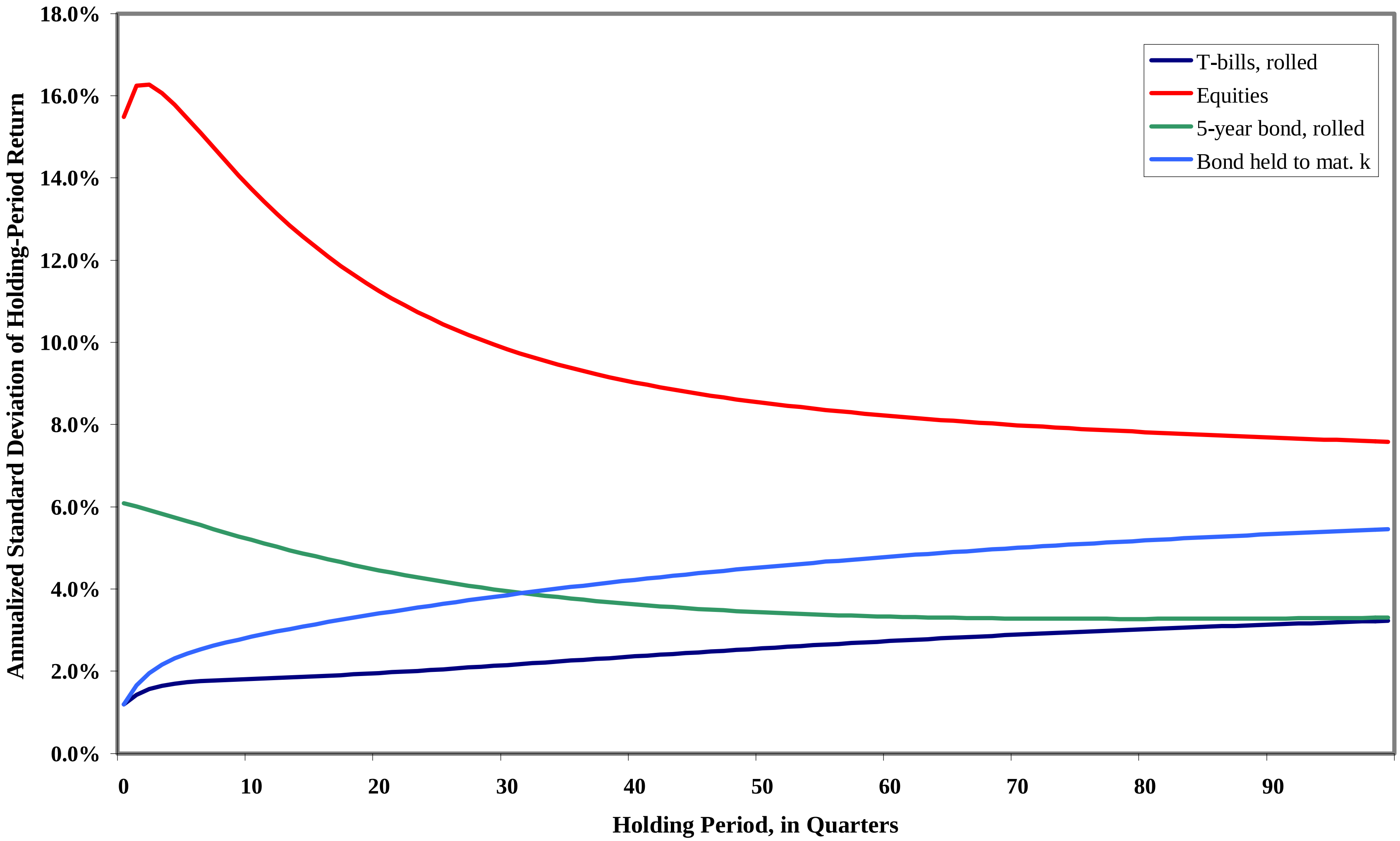
It is important to keep in mind two limitations of this portfolio analysis. First, it ignores constraints that might prevent investors from short-selling or from borrowing to invest in risky assets. The Siegel strategy of buying and holding stocks might be much closer to optimal for an aggressive investor who cannot borrow to leverage a stock market position, and who therefore normally holds the maximum 100% weight in equities.

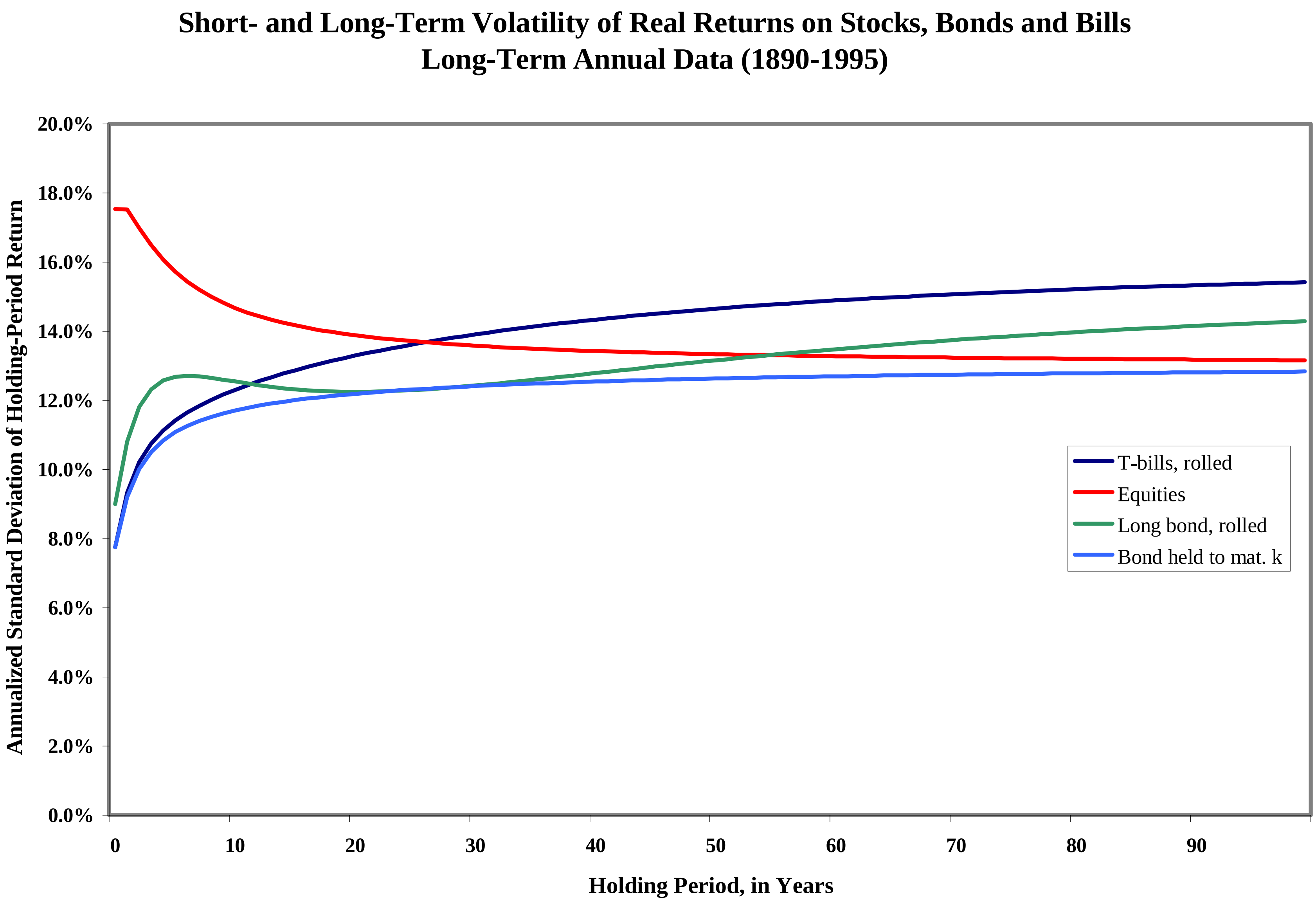
Second, I have solved the microeconomic portfolio choice problem of a rational investor with constant relative risk aversion and no human wealth, but such an investor cannot be the representative investor. As we discussed earlier on, the representative investor must have different preferences, constraints, or beliefs in order to be content to hold the aggregate wealth portfolio. Thus the portfolio advice of Figure 3 can only be used by atypical investors.

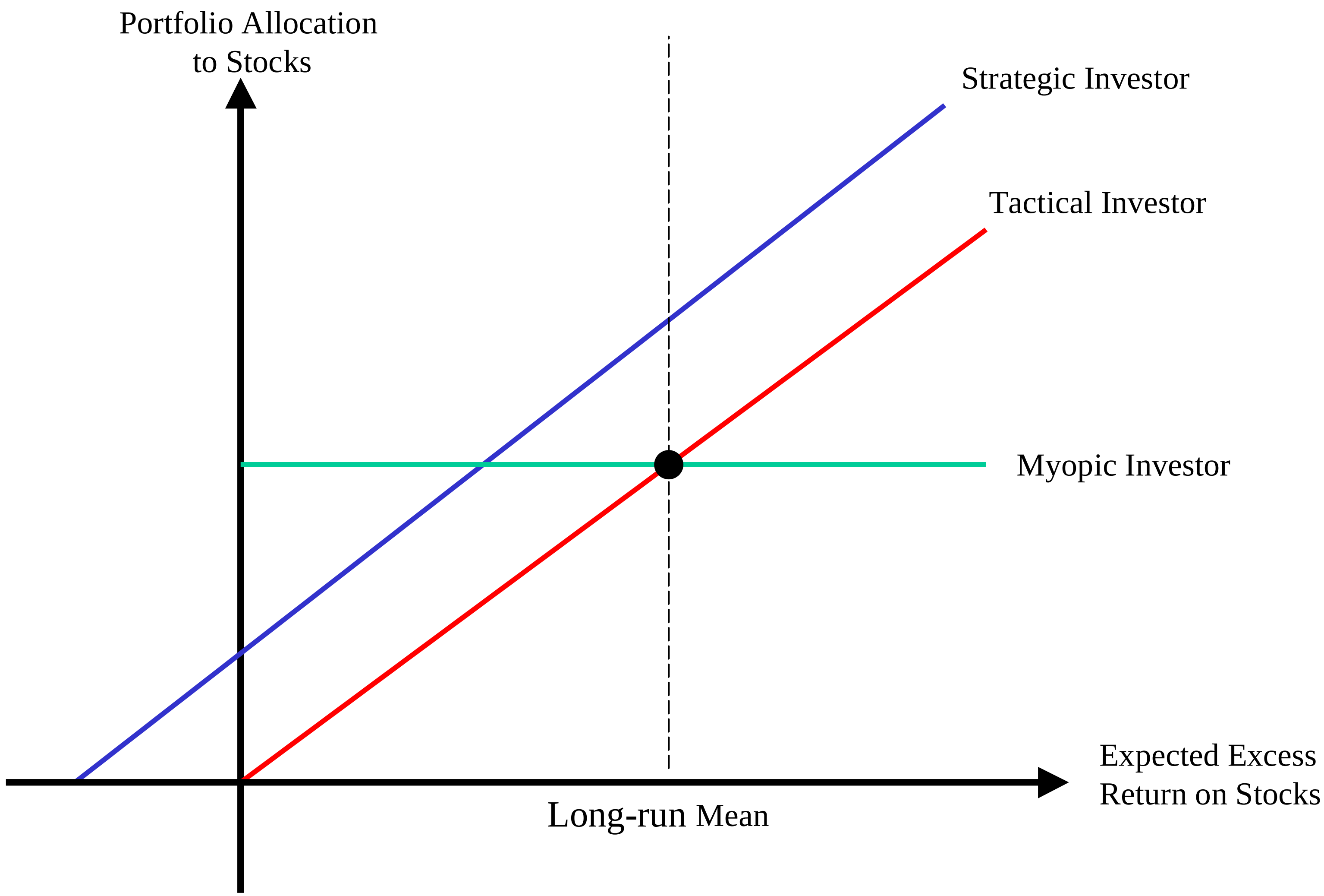
Conclusion

In these lectures I have argued that the stock market moves as if risk aversion is volatile and countercyclical. This could be due to habit formation, countercyclical idiosyncratic labor income risk, or irrational expectations. These findings have interesting implications for the optimal portfolio choice of investors with constant risk aversion and no labor income risk. Such investors should invest more aggressively when consumption and stock prices are low than when they are high. Also, long-term investors with constant risk aversion greater than one should invest more aggressively on average than short-term investors with the same risk aversion. This last result supports the view, sometimes expressed by financial planners, that investors can afford to take greater stock market risk if they have a long investment horizon.

Short- and Long-Term Volatility of Real Returns on Stocks, Bonds and Bills
Postwar Quarterly Data




Short- and Long-Term Volatility of Real Returns on Stocks, Bonds and Bills
Long-Term Annual Data (1890-1995)
Annualized Standard Deviation of Holding-Period Return
20.0%
18.0%
16.0%
14.0%
12.0%
10.0%
8.0%
6.0%
4.0%
2.0%
0.0%
0
10
20
30
40
50
60
70
80
90
Holding Period, in Years
T-bills, rolled
Equities
Long bond, rolled
Bond held to mat. k



Portfolio Allocation
to Stocks
Strategic Investor
Tactical Investor
Myopic Investor
Expected Excess
Return on Stocks
Long-run Mean

Understanding Risk and Return

2001 Marshall Lectures
University of Cambridge

John Y. Campbell
Harvard University

Bibliography

Abel, Andrew B., 1990, "Asset Prices under Habit Formation and Catching Up with the Joneses", *American Economic Review Papers and Proceedings* 80, 38–42.

Abel, Andrew B., 1999, "Risk Premia and Term Premia in General Equilibrium", *Journal of Monetary Economics* 43, 3–33.

Abel, Andrew B., N. Gregory Mankiw, Lawrence H. Summers, and Richard J. Zeckhauser, 1989, "Assessing Dynamic Efficiency: Theory and Evidence", *Review of Economic Studies* 56, 1–20.

Barberis, Nicholas, 2000, "Investing for the Long Run when Returns are Predictable", *Journal of Finance* 55, 225–264.

Barberis, Nicholas, Ming Huang, and Tano Santos, 2001, "Prospect Theory and Asset Prices", *Quarterly Journal of Economics* 116, 1–53.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, "A Model of Investor Sentiment", *Journal of Financial Economics* 49, 307–343.

Barsky, Robert B. and J. Bradford De Long, 1993, "Why Does the Stock Market Fluctuate?", *Quarterly Journal of Economics* 107, 291–311.

Bernstein, Peter, 1996, *Against the Gods: The Remarkable Story of Risk*, Wiley.

Blanchard, Olivier J. and Mark Watson, 1982, "Bubbles, Rational Expectations, and Financial Markets", in P. Wachtel ed., *Crises in the Economic and Financial Structure: Bubbles, Bursts, and Shocks*, Lexington: Lexington, MA.

Breeden, Douglas T., 1979, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities", *Journal of Financial Economics* 7, 265–296.

Brennan, Michael J., Eduardo S. Schwartz, and Ronald Lagnado, 1997, "Strategic Asset Allocation", *Journal of Economic Dynamics and Control* 21, 1377–1403.

Campbell, John Y., 1987, "Stock Returns and the Term Structure", *Journal of Financial Economics* 18, 373–399.

Campbell, John Y., 1991, "A Variance Decomposition for Stock Returns", *Economic Journal* 101, 157–179.

Campbell, John Y., 1996, “Understanding Risk and Return”, *Journal of Political Economy* 104, 298–345.

Campbell, John Y., 1999, “Asset Prices, Consumption, and the Business Cycle”, Chapter 19 in John Taylor and Michael Woodford eds. *Handbook of Macroeconomics Vol. 1*, 1231–1303.

Campbell, John Y., 2000, “Asset Pricing at the Millennium”, *Journal of Finance* 55, 1515–1567.

Campbell, John Y. and John H. Cochrane, 1999, “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior”, *Journal of Political Economy* 107, 205–251.

Campbell, John Y. and N. Gregory Mankiw, 1989, “Consumption, Income, and Interest Rates: Reinterpreting the Time Series Evidence”, in Olivier J. Blanchard and Stanley Fischer eds., *National Bureau of Economic Research Macroeconomics Annual* 4, 185–216.

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.

Campbell, John Y. and Robert J. Shiller, 2001, “Valuation Ratios and the Long-Run Stock Market Outlook: An Update”, NBER Working Paper No. 8221.

Campbell, John Y. and Luis M. Viceira, 1999, “Consumption and Portfolio Decisions When Expected Returns Are Time Varying, *Quarterly Journal of Economics* 114, 433–495.

Campbell, John Y. and Luis M. Viceira, 2001, “Who Should Buy Long-Term Bonds?”, *American Economic Review* 91, 99–127.

Campbell, John Y. and Luis M. Viceira, 2002, *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, forthcoming Oxford University Press.

Cecchetti, Stephen G., Pok-sang Lam, and Nelson C. Mark, 1993, “The Equity Premium and the Risk-Free Rate: Matching the Moments”, *Journal of Monetary Economics 31*, 21–45.

Cecchetti, Stephen G, Pok-sang Lam, and Nelson C. Mark, 2000, "Asset Pricing with Distorted Beliefs: Are Equity Returns Too Good to Be True?" *American Economic Review* 90, 787–805.

Chan, Y. Lewis and Leonid Kogan, 2000, "Heterogeneous Preferences and the Dynamics of Asset Prices", unpublished paper, Wharton School, University of Pennsylvania.

Cochrane, John H., 2001, *Asset Pricing*, Princeton University Press, Princeton, NJ.

Cogley, Timothy, 1998, "Idiosyncratic Risk and the Equity Premium: Evidence from the Consumer Expenditure Survey, working paper, Federal Reserve Bank of San Francisco.

Constantinides, George M., 1990, "Habit Formation: A Resolution of the Equity Premium Puzzle", *Journal of Political Economy* 98, 519–543.

Constantinides, George M., John B. Donaldson, and Rajnish Mehra, 1998, "Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle", NBER Working Paper No. 6617.

Constantinides, George M. and Darrell Duffie, 1996, "Asset Pricing with Heterogeneous Consumers", *Journal of Political Economy* 104, 219–240.

Dimson, Elroy, Paul Marsh, and Mike Staunton, 2000, *The Millennium Book: A Century of Investment Returns*, ABN Amro and London Business School, London.

Dumas, Bernard, 1989, "Two-person Dynamic Equilibrium in the Capital Market, *Review of Financial Studies* 2, 157–188.

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

Epstein, Lawrence and Stanley Zin, 1991, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation", *Journal of Political Economy* 99, 263–286.

Fama, Eugene and Kenneth R. French, 2000, "The Equity Premium", unpublished paper, University of Chicago and MIT.

Glassman, James K. and Kevin A. Hassett, 1999, *Dow 36,000*, Times Books: New York, NY.

Gordon, Myron, 1962, *The Investment, Financing, and Valuation of the Corporation*, Irwin: Homewood, IL.

Graham, Benjamin and David L. Dodd, 1934, *Security Analysis*, McGraw-Hill: New York, NY.

Grossman, Sanford J. and Robert J. Shiller, 1981, "The Determinants of the Variability of Stock Market Prices", *American Economic Review* 71, 222–227.

Grossman, Sanford J. and Robert J. Shiller, 1982, "Consumption Correlatedness and Risk Measurement in Economies With Non-Traded Assets and Heterogeneous Information", *Journal of Financial Economics* 10, 195–210.

Grossman, Sanford J. and Z. Zhou, 1996, "Equilibrium Analysis of Portfolio Insurance", *Journal of Finance* 51, 1379–1403.

Hall, Robert E., 1988, "Intertemporal Substitution in Consumption", *Journal of Political Economy* 96, 221–273.

Hamilton, James D., 1989, "A New Approach to the Analysis of Nonstationary Returns and the Business Cycle", *Econometrica* 57, 357–384.

Hansen, Lars P. and Ravi Jagannathan, 1991, "Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns", *Journal of Political Economy* 99, 225-262.

Hansen, Lars P., Thomas J. Sargent, and Thomas Tallarini, 1997, "Robust Permanent Income and Pricing", Working paper, University of Chicago.

Harvey, Campbell R., 1989, "Time-Varying Conditional Covariances in Tests of Asset Pricing Models", *Journal of Financial Economics* 24, 289–317.

Heaton, John and Deborah Lucas, 1996, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing", *Journal of Political Economy* 104, 668–712.

Heaton, John and Deborah Lucas, 1999, "Stock Prices and Fundamentals", *NBER Macroeconomics Annual* 213–242, MIT Press, Cambridge, MA.

Jagannathan, Ravi, Ellen R. McGrattan, and Anna Scherbina, 2001, "The Declining US Equity Premium", NBER Working Paper No. 8172, Cambridge, MA.

Jorion, Philippe and William N. Goetzmann, 1999, "Global Stock Markets in the Twentieth Century", *Journal of Finance* 54, 953–980.

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

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

Krusell, Per and A.A. Smith, Jr., 1997, "Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns", working paper, University of Rochester and Carnegie Mellon University.

Lettau, Martin, 1997, "Idiosyncratic Risk and Volatility Bounds", working paper, CentER, Tilburg University.

Lettau, Martin and Sydney Ludvigson, 2001, "Consumption, Aggregate Wealth, and Expected Stock Returns", forthcoming *Journal of Finance.*

Lewellen, Jonathan, 2001, "Predicting Returns with Financial Ratios", unpublished paper, Sloan School, MIT.

Liang, Nellie and Steven A. Sharpe, 1999, "Share Repurchases and Employee Stock Options and their Implications for S&P500 Share Retirements and Expected Returns", unpublished paper, Board of Governors of the Federal Reserve System.

Lucas, Robert E. Jr., 1978, "Asset Prices in an Exchange Economy", *Econometrica* 46, 1429–1446.

Mankiw, N. Gregory, 1986, "The Equity Premium and the Concentration of Aggregate Shocks, *Journal of Financial Economics* 17, 211–219.

Markowitz, Harry, 1952, "Portfolio Selection", *Journal of Finance* 7, 77–91.

McGrattan, Ellen R., and Edward C. Prescott, 2001, "Is the Stock Market Overvalued?", NBER Working Paper No. 8077, Cambridge, MA.

Mehra, Rajnish and Edward C. Prescott, 1985, "The Equity Premium: A Puzzle", *Journal of Monetary Economics* 15, 145–161.

Merton, Robert C., 1969, "Lifetime Portfolio Selection Under Uncertainty: The Continuous Time Case", *Review of Economics and Statistics* 51, 247–257.

Merton, Robert C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model, *Journal of Economic Theory* 3, 373–413.

Merton, Robert C., 1973, "An Intertemporal Capital Asset Pricing Model", Econometrica 41, 867–887.

Modigliani, Franco and Richard A. Cohn, 1979, "Inflation and the Stock Market, *Financial Analysts Journal* 35, 24–44.

Rietz, Thomas, 1988, "The Equity Premium Puzzle: A Solution?", *Journal of Monetary Economics* 21, 117–132.

Ritter, Jay R. and Richard S. Warr, 1999, "The Decline of Inflation and the Bull Market of 1982 to 1997", working paper, University of Florida.

Rubinstein, Mark, 1976, "The Valuation of Uncertain Income Streams and the Pricing of Options", *Bell Journal of Economics* 7, 407–25.

Santos, Manuel S. and Michael Woodford, 1997, "Rational Asset Pricing Bubbles", *Econometrica* 65, 19–57.

Sharpe, Steven A., 1999, "Stock Prices, Expected Returns, and Inflation", working paper, Board of Governors of the Federal Reserve System.

Shiller, Robert J., 1981, "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?", *American Economic Review* 71, 421–436.

Shiller, Robert J., 1982, "Consumption, Asset Markets, and Macroeconomic Fluctuations", *Carnegie-Rochester Conference Series on Public Policy* 17, 203–238.

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

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

Sundaresan, Suresh M., 1989, "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth", *Review of Financial Studies* 2, 73–88.

Thaler, Richard H. and Eric J. Johnson, 1990, "Gambling With the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice", *Management Science* 36, 643–660.

Tirole, Jean, 1985, "Asset Bubbles and Overlapping Generations", *Econometrica* 53, 1499–1527.

Vissing-Jorgensen, Annette, 1999, "Limited Stock Market Participation and the Equity Premium Puzzle", unpublished paper, University of Chicago.

Wang, Jiang, 1996, "The Term Structure of Interest Rates in a Pure Exchange Economy with Heterogeneous Investors", *Journal of Financial Economics* 41, 75–110.

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

Williams, John Burr, 1938, *The Theory of Investment Value*, Harvard University Press: Cambridge, MA.