NOVEMBER/DECEMBER 1997

ECONOMIC PERSPECTIV

Call for Papers
1998 Conference on
Bank Structure
& Competition

A review from the
Federal Reserve Bank
of Chicago

Where is the market going?
Uncertain facts and novel theories

Index for 1997

FEDERAL RESERVE BANK
OF CHICAGO

ECONOMIC PERSPECTIVES

Goes Quarterly

To our readers...

Following a comprehensive review of Federal Reserve Bank of Chicago publications in 1997, including a readership survey in which some of you participated, the Economic Research Department has decided to expand Economic Perspectives in the coming year.

This expansion will be accompanied by a change in the publishing schedule from bimonthly to quarterly. Each quarterly issue will typically contain four articles instead of the usual two in the bimonthly format for a higher total number of articles over the course of the year.

The November/December 1997 issue of Economic Perspectives will be the last to be published according to the bimonthly schedule. In the coming months, we will begin preparing the first quarterly issue, which you will receive at the end of the first quarter of 1998.

We greatly appreciate your continuing interest in this publication and hope you will be pleased with the changes we are implementing. While we are working on the first quarter 1998 issue, you can access other Research Department publications and data via the Bank's Web site at www.frbchi.org.

On behalf of the Economic Research Department,

William C. Hunter,
Senior vice president
and director of research

Contents

ECONOMIC PERSPECTIVES

November/December 1997, Volume XXI, Issue 6

ECONOMIC PERSPECTIVES is published by the Research Department of the Federal Reserve Bank of Chicago. The views expressed are the authors' and do not necessarily reflect the views of the management of the Federal Reserve Bank.

Single-copy subscriptions are available free of charge. Please send requests for single- and multiple-copy subscriptions, back issues, and address changes to the Public Information Center, Federal Reserve Bank of Chicago, P.O. Box 834, Chicago, Illinois 60690-0834, telephone (312) 322-5111 or fax (312) 322-5515.

ECONOMIC PERSPECTIVES and other Bank publications are available on the World Wide Web at http://www.frbchi.org.

ISSN 0164-0682

Where is the market going? Uncertain facts and novel theories

John H. Cochrane

Over the last century, the stock market in the United States has yielded impressive returns to its investors. For example, in the postwar period, stock returns have averaged 8 percentage points above Treasury bills. Will stocks continue to give such impressive returns in the future? Are long-term average stock returns a fundamental feature of advanced industrial economies? Or are they the opposite of the old joke on Soviet agriculture—100 years of good luck? If not pure good luck, perhaps they result from features of the economy that will disappear as financial markets evolve.

How does the recent rise in the stock market affect our view of future returns? Do high prices now mean lower returns in the future? Or have stocks finally achieved Irving Fisher's brilliantly mistimed 1929 prediction of a "permanently high plateau?" If stocks have reached a plateau, is it a rising plateau, or is the market likely to bounce around its current level for many years, not crashing but not yielding returns much greater than those of bonds?

These questions are on all of our minds as we allocate our pension plan monies. They are also important to many public policy questions. For example, many proposals to reform social security emphasize the benefits of moving to a funded system based on stock market investments. But this is a good idea only if the stock market continues to provide the kind of returns in the future that it has in the past.

In this article, I summarize the academic, and if I dare say so, scientific, evidence on these issues. I start with the statistical analysis of past stock returns. The long-term average return is in fact rather poorly measured. The standard statistical confidence interval extends from 3 percent to 13 percent. Furthermore, average returns have been low following times of high stock prices, such as the present. Therefore, the statistical evidence suggests a period of quite low average returns, followed by slow reversion to a poorly measured long-term average, and it cautions us that statistical analysis alone leaves lots of uncertainty.

Then, I survey economic theory to see if standard models that summarize a vast amount of other information shed light on stock returns. Standard models do not predict anything like the historical equity premium. After a decade of effort, a range of drastic modifications to the standard models can account for the historical equity premium. But it remains to be seen whether the drastic modifications and a high equity premium, or the standard models and a low equity premium will triumph in the end. In sum, economic theory gives one further reason to fear that long-term average excess returns will not return to 8 percent, and it details the kind of beliefs one must have about the economy to reverse that pessimistic view.

John H. Cochrane is the Sigmund E. Edelstone Professor of Finance in the Graduate School of Business at the University of Chicago and a consultant to the Federal Reserve Bank of Chicago. His research is supported by a grant from the National Science Foundation, administered by the National Bureau of Economic Research, and by the University of Chicago. The author would like to thank George Constantinides, Andrea Eisfeldt, and David Marshall for many helpful comments.

However, I conclude with a warning that low average returns do not imply one should change one's portfolio. Someone has to hold every stock on the market. An investor should only hold less stocks than average if that investor is different from the average investor in some identifiable way, such as risk exposure, attitude, or information.

Average returns and risk

The most obvious place to start thinking about future stock returns is a statistical analysis of past stock returns.

Average real returns

Table 1 presents several measures of average real returns on stocks and bonds in the postwar period. The value weighted NYSE portfolio shows an impressive annual return of 9 percent after inflation. The S&P 500 is similar. The equally weighted NYSE portfolio weights small stocks more than the value weighted portfolio. Small stock returns have been even better than the market on average, so the equally weighted portfolio has earned more than 11 percent. Bonds by contrast seem a disaster. Long-term government bonds earned only 1.8 percent after inflation, despite a standard deviation (11 percent) more than half that of stocks (about 17 percent). Corporate bonds earned a slight premium over government bonds, but at 2.1 percent are still unappealing compared to stocks. Treasury bills earned only 0.8 percent on average after inflation.

A reward for risk

Table 1 highlights a crucially important fact. *High average returns are only earned as a compensation for risk.* High stock returns cannot be understood merely as high "productivity of the American economy" (or high marginal productivity of capital) or impatience by consumers. Such high productivity or impatience would lead to high returns on bonds as well. To understand average stock returns, and to assess whether they will continue at these levels, it is not necessary to understand why the economy gives such high returns to saving—it doesn't—but why it gives such high compensation for *bearing risk*. The risk is substantial. A 17 percent standard deviation means the market is quite likely to decline 9 – 17 = 8% or rise 9 + 17 = 26% in a year. (More precisely, there is about a 30 percent probability of a decline bigger than –8 percent or a rise bigger than 26 percent.)

Risk at short and long horizons

It is a common fallacy to dismiss this risk as "short-run price fluctuation" and to argue that stock market risk declines in the long run.

The most common way to fall into this trap is to confuse the *annualized* or average return with the actual return. For example, the two-year log or continuously compounded return is the sum of the one-year returns, $r_{0\to 2} = r_{0\to 1} + r_{1\to 2}$. Then, if returns were independent over time, like coin flips, the mean and variance would scale the same way with horizon: $E(r_{0\to 2}) = 2E(r_{0\to 1})$ and $\sigma^2(r_{0\to 2}) = 2\,\sigma^2(r_{0\to 1})$. Investors who cared about mean and variance would invest the same fraction of their wealth in stocks for any return horizon. The variance of *annualized* returns does stabilize; $\sigma(1/2\, r_{0\to 2}) = 1/2\, \sigma^2(r_{0\to 1})$. But the investor cares about the total, not annualized return. An example may clarify the distinction. Suppose you are betting $1 on a coin flip. This is a risky bet, you will either gain or lose $1. If you flip the coin 1,000 times, the average number of heads (annualized returns) will almost certainly come out quite near 50 percent.

TABLE 1

Annual real returns 1947–96

	VW	S&P500	EW	GB	CB	TB
	(- - - - - - - -	- - - - - - - - -	percent	- - - - - - - - -	- - - - - - - -	-)
Average return $E(R)$	9.1	9.5	11.0	1.8	2.1	0.8
Standard deviation $\sigma(R)$	16.7	16.8	22.2	11.1	10.7	2.6
Standard error $\sigma(R)/\sqrt{T}$	2.4	2.4	3.0	1.6	1.5	0.4

Notes: VW = value weighted NYSE, EW = equally weighted NYSE, GB = ten-year government bond, CB = corporate bond, TB = three-month Treasury bills. All less CPI inflation.

Source: All data for this and subsequent figures and tables in this article are from the Center for Research on Security Prices (CRSP) at the University of Chicago.

However, the risk of the bet (total return) is much larger: It only takes an average number of heads equal to 0.499 (that is, 499/1,000) to lose a dollar; if the average number of heads is 0.490, still very close to 0.5, you lose $10. Just as we care about dollars, not the fraction of heads, we care about total returns, not annualized rates.

To address the short-run price fluctuation fallacy directly, table 2 shows that mean returns and standard deviations scale with horizon just about as this independence argument suggests, out to five years.

(In fact, returns are not exactly independent over time. Estimates in Fama and French [1988a] and Poterba and Summers [1988] suggest that the variance grows a bit less slowly than the horizon for the first five to ten years, and then grows with horizon as before, so stocks are in fact a bit safer for long horizons than the independence assumption suggests. However, this qualification does not rescue the annualized return fallacy. Also bear in mind that long-horizon statistics are measured even less well than annual statistics; there are only five nonoverlapping ten-year samples in the postwar period.)

The stock market is like a coin flip, but it is a biased coin flip. Thus, even though mean and variance may grow at the same rate with horizon, the probability that one loses money in the stock market does decline over time. (For example, for the normal distribution, tail probabilities are governed by $E(r)/\sigma(r)$, which grows at the square root of horizon.) However, portfolio advice is not based on pure probabilities of making or losing money; but on measures such as the mean and variance of return. Based on such measures, there is not much presumption that stocks are dramatically safer for long-run investments.

I cannot stress enough that the high average returns come only as compensation for risk. Our task below is to understand this risk and people's aversion to it. Many discussions, including those surrounding the move to a funded social security system, implicitly assume that one gets the high returns without taking on substantial risk. What happens to a funded social security system if the market goes down?

TABLE 2

How risk and return vary with investment horizon

Horizon h (years)	$\frac{E(R^e)}{h}$	$\frac{\sigma(R^e)}{\sqrt{h}}$	$\frac{1}{\sqrt{h}}\frac{E(R^e)}{\sigma(R^e)}$
1	8.6	17.1	0.50
2	9.1	17.9	0.51
3	9.2	16.8	0.55
5	10.5	21.9	0.48

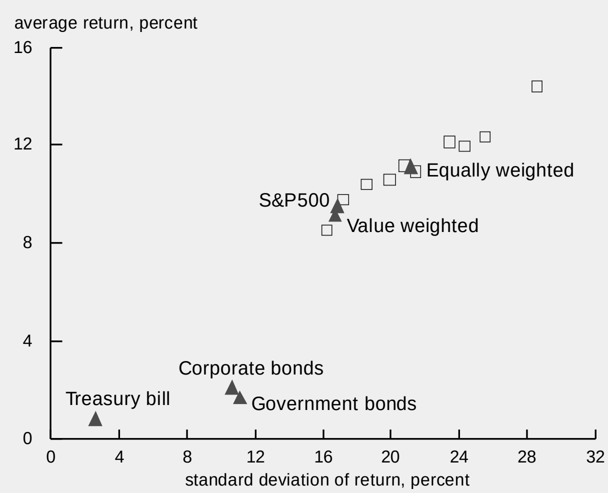
Notes: R^e = value weighted return less T-bill rate. Column one shows average excess return divided by horizon. Column two shows standard deviation of excess return divided by square root of horizon. Column three shows Sharpe ratio divided by square root of horizon. All statistics in percent.

Means versus standard deviations—Sharpe ratio

Figure 1 presents mean returns versus their standard deviations. In addition to the portfolios listed in table 1, I include ten portfolios of NYSE stocks sorted by size. This picture shows that average returns alone are not a particularly useful measure. By taking on more risk, one can achieve very high average returns. In the picture, the small stock portfolio

FIGURE 1

Mean vs. standard deviation of real returns, 1947–96



Notes: Triangles are equally weighted and value weighted NYSE; S&P 500; three-month Treasury bill; ten-year government bond; and corporate bond returns. Unmarked squares are NYSE size portfolios.

earns over 15 percent per year average real return, though at the cost of a huge standard deviation. Furthermore, one can form portfolios with even higher average returns by leveraging—borrowing money to buy stocks—or investing in securities such as options that are very sensitive to stock returns. Since standard deviation (and beta or other risk measures) grow exactly as fast as mean return, the extra mean return gained in this way exactly corresponds to the extra risk of such portfolios. When considering economic models, it is easy to get them to produce higher mean returns (along with higher standard deviations) by considering claims to leveraged capital.

In sum, *excess returns* of stocks over Treasury bills are more interesting than the level of returns. This is the part of return that is a compensation for risk, and it accounts for nearly all of the amazingly high average stock returns. Furthermore, the *Sharpe ratio* of mean excess return to standard deviation, or the slope of a line connecting stock returns to a risk-free interest rate in figure 1, is a better measure of the fundamental characteristic of stocks than the mean excess return itself, since it is invariant to leveraging. The stock portfolios listed in table 1 all have Sharpe ratios near 0.5.

Standard errors

The average returns and Sharpe ratios look impressive. But are these true or just chance? One meaning of chance is this: Suppose that the average excess return really is low, say 3 percent. How likely is it that a 50-year sample has an average excess return of 8 percent? Similarly, if the next 50 years are "just like" the last 50, in the sense that the structure of the economy is the same but the random shocks may be different, what is the chance that the average return in the *next* 50 years will be as good as it was in the last 50?

Since we only see one sample, these questions are really unanswerable at a deep level. Statistics provides an educated guess in the *standard error*. Assuming that each year's return is statistically independent, our best guess of the standard deviation of the *average* return is $\sigma / \sqrt{T}$, where σ is the standard deviation of annual returns and T is the data size.

This formula tells us something important: *Stock returns are so volatile that it is very hard to statistically measure average returns*. Table 1 includes standard errors of stock returns measured in the last 50 years, and table 3 shows standard errors for a variety of horizons. The confidence interval, mean +/–2 standard deviations, represents the 95 percent probability range. As the table shows, even very long-term averages leave a lot of uncertainty about mean returns. For example, with 50 years of data, an 8 percent average excess return is measured with a 2.4 percentage point standard error. Thus, the confidence interval says that the true average excess return is between $8 - 2 \times 2.4 = 3\%$ and $8 + 2 \times 2.4 = 13\%$ with 95 percent probability.[1] This is a wide band of uncertainty about the true market return, given 50 years of data. One can also see that five- or ten-year averages are nearly useless; it takes a long time to statistically discern that the average return has increased or decreased. As a cold winter need not presage an ice age, so even a decade of bad returns need not change one's view of the true underlying average return.

The standard errors are also the standard deviations of average returns over the next T years, and table 3 shows that there is quite a lot of uncertainty about those returns. For example, if the true mean excess return is and will continue to be 8 percent, the five-year standard error of 7.6 percent is almost as large as the mean. This means that there is still a good chance that the next five-year return will average less than the Treasury bill rate.

On the other hand, though the average return on stocks is not precisely known, the 2.4 percent standard error means that we can confidently reject the view that the true mean excess return was zero or even 2–3 percent. The argument that *all* the past equity premium was luck doesn't hold up well against this simple statistical argument.

TABLE 3

Standard error of average return at various horizons

Horizon T (years)	Standard error $\sigma / \sqrt{T}$ (percentage points)
5	7.6
10	5.4
25	3.8
50	2.4

Note: Returns assumed to be statistically independent with standard deviation σ = 17 percent.

Selection and crashes

Two important assumptions behind the standard error calculation, however, suggest ways in which the postwar average stock return might still have been largely due to luck. Argentina and the U.S. looked very similar at the middle of the last century. Both economies were underdeveloped relative to Britain and Germany and had about the same per capita income. If Argentina had experienced the U.S.'s growth and stock returns, and vice versa, this article would be written in Spanish from the Buenos Aires Federal Reserve Bank, with high Argentine stock returns as the subject.

The statistical danger this story points to is *selection* or *survival* bias. If you flip one coin ten times, the chance of seeing eight heads is low. But if you flip ten coins ten times, the chance that the coin with the greatest number of heads exceeds eight heads is much larger. Does this story more closely capture the 50-year return on U.S. stocks? Brown, Goetzmann, and Ross (1995) present a strong case that the uncertainty about true average stock returns is much larger than $\sigma / \sqrt{T}$ suggests. As they put it, "Looking back over the history of the London or the New York stock markets can be extraordinarily comforting to an investor—equities appear to have provided a substantial premium over bonds, and markets appear to have recovered nicely after huge crashes. . . . Less comforting is the past history of other major markets: Russia, China, Germany, and Japan. Each of these markets has had one or more major interruptions *that prevent their inclusion in long term studies*" [my emphasis].

In addition, think of the things that didn't happen in the last 50 years. There were no banking panics, no depressions, no civil wars, no constitutional crises, the cold war was not lost, and no missiles were fired over Berlin, Cuba, Korea, or Vietnam. If any of these things had happened, there undoubtedly would have been a calamitous decline in stock values. The statistical problem is *nonnormality*. Taking the standard deviation from a sample that did not include rare calamities, and calculating average return probabilities from a normal distribution may understate the true uncertainty. But investors, aware of that uncertainty, discount prices and hence leave high returns on the table.

We can cast the issue in terms of fundamental beliefs about the economy. Was it clear to people in 1945 (or 1871, or whenever the sample starts) and throughout the period that the average return on stocks would be 8 percent greater than that of bonds? If so, one would expect them to have bought more stocks, even considering the risk described by the 17 percent year-to-year variation. But perhaps it was not in fact obvious in 1945, that rather than slipping back into depression, the U.S. would experience a half century of growth never before seen in human history. If so, much of the equity premium was unexpected; good luck.

Time varying expected returns

Regressions of returns on price/dividend ratios

We are not only concerned with the *average* return on stocks but whether returns are expected to be unusually low at a time of high prices, such as the present. The first and most natural thing one might do to answer this question is to look at a regression forecast. To this end, table 4 presents regressions of returns on the price/dividend (P/D) ratio.

TABLE 4

OLS regressions of excess returns and dividend growth on VW P/D ratio

Horizon k (years)	$R_{t\to t+k} = a + b(P_t/D_t)$			$D_{t+k}/D_t = a + b(P_t/D_t)$		
	b	$\sigma(b)$	R^2	b	$\sigma(b)$	R^2
1	–1.04	(0.33)	0.17	–0.39	(0.18)	0.07
2	–2.04	(0.66)	0.26	–0.52	(0.40)	0.07
3	–2.84	(0.88)	0.38	–0.53	(0.43)	0.07
5	–6.22	(1.24)	0.59	–0.99	(0.47)	0.15

Notes: $R_{t\to t+k}$ indicates the k year return on the value weighted NYSE portfolio less the k year return from continuously reinvesting in Treasury bills; b = regression slope coefficient (defined by the regression equation above); $\sigma(b)$ = standard error of regression coefficient. Standard errors in parentheses use GMM to correct for heteroscedasticity and serial correlation.

The regression at a one-year horizon shows that excess returns are in fact predictable from P/D ratios, though the 0.17 R^2 is not particularly remarkable. However, at longer and longer horizons, the slope coefficients increase and larger and larger fractions of return variation can be forecasted. At a five-year horizon, 60 percent of the variation in stock returns can be forecasted ahead of time from the P/D ratio. (Fama and French, 1988b, is a famous early source for this kind of regression.)

One can object to dividends as the divisor for prices. However, price divided by just about anything sensible works about as well, including earnings, book value, and moving averages of past prices. There seems to be an additional business-cycle component of expected return variation that is tracked by the term spread or other business cycle forecasting variables, including the default spread and investment–capital ratio, the T-bill rate, the ratio of the T-bill rate to its moving average, and the dividend/earnings ratio. (See Fama and French, 1989, for term and default spreads, Campbell, 1987, for term spread, Cochrane, 1991c, for investment–capital ratios, Lamont, 1997, for dividend/earnings, and Ferson and Constantinides, 1991, for an even more exhaustive list with references). However, price ratios such as P/D are the most important forecasting variables, especially at long horizons, so I focus on the P/D ratio to keep the analysis simple. In a similar fashion, *cross-sectional* variation in expected returns can be very well described by the P/D ratio or (better) the ratio of market value to book value, which contains the price in its numerator. Portfolios of "undervalued" or "value" stocks with low price ratios outperform portfolios of "overvalued" or "growth" stocks with high price ratios. (See Fama and French, 1993).

Slow moving P/D and P/E

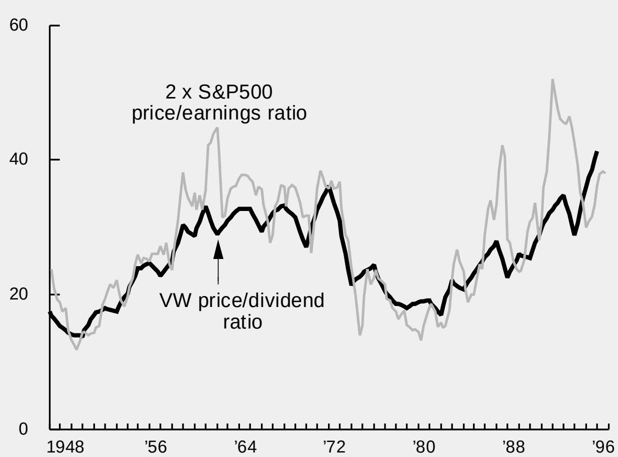
Figure 2 presents P/D and price/earnings (P/E) ratios over time. This graph emphasizes that price ratios are very *slow moving* variables. This is why they forecast long-horizon movements in stock returns.

The rise in forecast power with horizon is not a separate phenomenon. It results from the ability to forecast one period returns and the slow movement in the P/D ratio.[2] As an analogy, if it is ten degrees below zero in Chicago (low P/D ratio), one's best guess is that it will warm up a degree or so per day. Spring does come, albeit slowly. However, the weather varies a lot; it can easily go up or down 20 degrees in a day, so this forecast is not very accurate (low R^2). But the fact that it is ten degrees below zero signals that the temperature will rise a bit on average per day for many days. By the time we look at a six-month horizon, we forecast a 90 degree rise in temperature. The daily variation of 20 degrees is still there, but the change in temperature (90 degrees) that can be forecasted is much larger relative to the daily variation, implying a high R^2.

The slow movement in the P/D ratio also means that the ability to forecast returns is not the fabled alchemists' stone that turns lead into gold. A high P/D ratio means that prices will grow more slowly than dividends for a long time until the P/D ratio is reestablished, and vice versa. Trading on these signals—buying more stocks in times of low prices, and less in times of high prices—can raise (unconditional) average returns a bit, but not much more than 1 percent for the same standard

FIGURE 2

P/D and P/E ratios



Notes: VW is value weighted NYSE portfolio. Two times the S&P 500 P/E ratio is plotted so that the lines can be more easily compared.

deviation. If there were a 50 percent R^2 at a daily horizon, one could make a lot of money; but not so at a five-year horizon.

The slow movement of the P/D ratio also means that on a purely *statistical* basis, one can cast doubt on whether the P/D ratio really forecasts returns. What we really know, looking at figure 2 (figure 4 also makes this point), is that low prices relative to dividends and earnings in the 1950s preceded the boom market of the early 1960s; that the high P/D ratios of the mid-1960s preceded the poor returns of the 1970s; and that the low price ratios of the mid-1970s preceded the current boom. We also know that price ratios are very high now. In any real sense, there really are three data points. I do not want to survey the extensive statistical literature that formalizes this point, but it is there. Most importantly, it shows that the t-statistics one might infer from regressions such as table 4 are inflated; with more sophisticated tests, return predictability actually has about a 10 percent probability value before one starts to worry about fishing and selection biases.

What about repurchases? P/E and other forecasts

Is the P/D ratio still a valid signal? Perhaps increasing dividend repurchases mean that the P/D ratio will not return to its historical low values; perhaps it has shifted to a new mean so today's high ratio is not bad for returns. To address this issue, figure 2 plots the S&P 500 P/E ratio along with the P/D ratio. The two measures line up well. The P/E ratio forecasts returns almost as well as the P/D ratio. The P/E ratio, price/book value, and other ratios are also at historic highs, forecasting low returns for years to come. Yet they are of course immune to the criticism that the dividend–earnings relationship might be fundamentally different from the past.

Return forecasts

What do the regressions of table 4 say, quantitatively, about future returns? Figure 3 presents one-year returns and the P/D ratio forecast. Figure 4 presents five-year returns and the P/D ratio forecast. I include in-sample and out-of-sample forecasts in figures 3 and 4. To form the out-of-sample forecasts, I paired the regressions from table 4 with an autoregression of P/D_t,

$$P/D_{t+1} = \mu + \rho\, P/D_t + \delta_{t+1}.$$

Then, for example, since my data run through the end of 1996, the forecast returns for 1997 and 1998 are

$$E(R_{1997}) = a + b(P/D_{1996})$$
$$E(R_{1998}) = a + b(\mu + \rho\, P/D_{1996}),$$

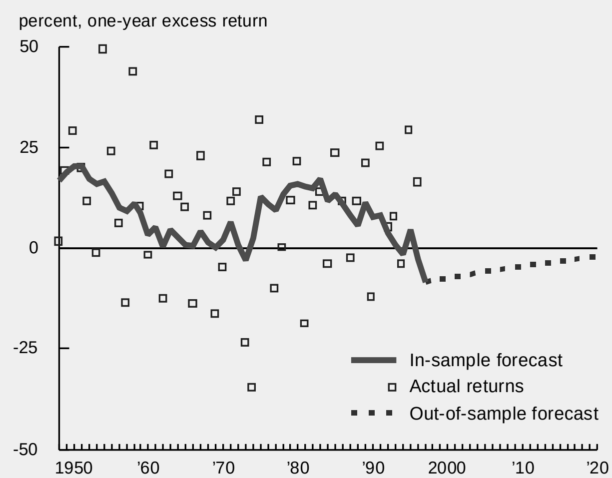
and so on.[3]

The one-year return forecast is extraordinarily pessimistic. It starts at a –8 percent excess return for 1997, and only very slowly returns to the estimated unconditional mean excess return of 8 percent. In ten years, the forecast is still –5 percent, in 25 years it is –1.75 percent, and it is still only 2.35 percent in 50 years. The five-year return forecasts are similarly pessimistic.

Of course, this forecast is subject to lots of uncertainty. There is uncertainty about what actual returns will be, given the forecasts. This will always be true: If one could precisely forecast the direction of stock prices, stocks would cease to be risky and would cease to pay a risk premium. There is also a great deal of uncertainty about the forecasts themselves.

FIGURE 3

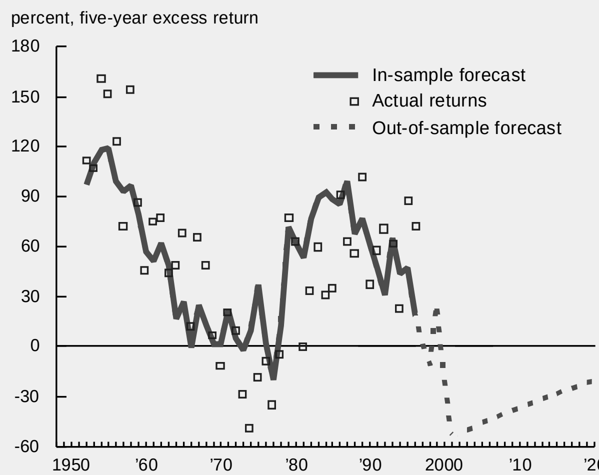
Actual and forecast one-year excess returns



Notes: One-year excess returns on the value weighted NYSE and forecast from a regression on the P/D ratio. Returns are plotted on the day of the forecast. For example, 1995 plots $a + b \times P/D_{1995}$ and the 1996 return. The out-of-sample forecast is made by joining $R_{t+1} = a + bP/D_t$ with $P/D_{t+1} = \mu + \rho P/D_t$.

FIGURE 4

Actual and forecast five-year excess returns



Notes: Five-year excess returns on value weighted NYSE (five-year returns minus five-year T-bill returns) and forecast from a regression of five-year returns on the P/D ratio. Returns are plotted on the last day; for example, 1996 plots the forecast $a + b \times P/D_{1991}$ and the return from 1991 to 1996. The out-of-sample forecast is formed for 1997–2001 from the observed P/D ratios, 1992–96, and for 2002 onward from $P/D_t = \mu + 0.98P/D_{t-1}$.

The forecasts attempt to measure *expected returns*, the quantity that investors must trade off against unavoidable risk in deciding how attractive an investment is, and they undoubtedly measure expected returns with error.

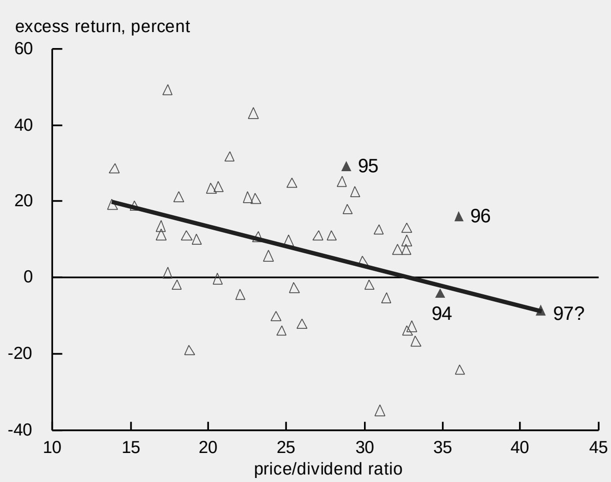
The plots of actual returns on top of the in-sample, one-year-ahead forecasts in figures 3 and 4 give one measure of the forecast uncertainty. One can see that year-to-year returns are quite likely to vary a lot given the forecast. Five-year returns track the forecast more closely, but here the chance of over-fitting is greater.

To get a handle on how reliable or robust the pessimistic forecast is, figure 5 gives a scatter plot of one-year returns and their forecasts based on P/D, together with the fitted regression line. The scatterplot indicates that the regression results are not spurious, or the result of a few outlying years.

The point marked "97?" is the P/D ratio at the end of 1996 together with the forecast return for 1997. We see immediately one source of trouble with the point forecast: the P/D ratio has never in the postwar period been as high as it is now. Extending historical experience to never-before seen values is always dangerous. One is particularly uncomfortable with a prediction that the market should earn *less* than the T-bill rate, given the strong theoretical presumption for a positive expected excess return.[4] One could easily draw a downward sloping line through the points, flattening out on the right, predicting a zero excess return for P/D ratios above 30 to 35, and never predicting a negative excess return. A nonlinear regression that incorporates this idea will fit about as well as the linear regression I have run. However, the scatterplot does not demand such a nonlinear relation either, so this is largely a matter of choice. In sum, while the scatterplot does suggest that the current forecast should be low, it does not give robust evidence that the forecast excess return should be negative.

FIGURE 5

One-year excess returns vs. one-year ahead forecast from P/D



What about the last few years of high returns?

The P/D ratios also pointed to low returns in 1995 and 1996. Anyone who took that advice missed out on a dramatic surge in the market, and some fund managers who took that advice are now unemployed. Doesn't this mean that the P/D signal should no longer be trusted?

To answer this criticism, look at the figures again. They make clear that the returns for 1995 and 1996 and even another 20 percent or so return for 1997 are not so far out of line, despite a pessimistic P/D forecast, that we should throw away the regression based on the previous 47 years of experience. To return to the analogy, if it is ten degrees below zero in Chicago, that means spring is coming. But we can easily have a few weeks of 20 degree below weather before spring finally arrives. The graphs make vivid how large a 17 percent standard deviation really is, and to what extent the forecasts based on the P/D ratio mark long-term tendencies that are still subject to lots of short-term swings rather than accurate forecasts of year-to-year booms or crashes.

Another source of uncertainty about the forecast is how persistent the P/D ratio really is. If, for example, the P/D ratio had no persistence, then the low return forecast would only last a year. After that, it would return to the unconditional mean of 9 percent (8 percent over Treasury bills). Now, given a true value $\rho = 0.98$ in $P/D_t = \mu + \rho P/D_{t-1} + \delta_t$, the median ordinary least squares (OLS) estimate is 0.90, as I found in sample. That is why figure 3 uses the value $\rho = 0.98$. However, given this true value, the OLS estimate lies between 0.83 and 0.94 only 50 percent of the time and between 0.66 and 1.00 for 95 percent of the time. Thus, there is a huge range of uncertainty over the true value of ρ. The best thing that could happen to the forecast is if the P/D ratio were really less persistent than it seems. In this case, the near-term return forecast would be unchanged, but the long-term return forecast would return to 9 percent much more quickly.

Variance decomposition

When prices are high relative to dividends (or earnings, cash flow, book value, or some other divisor), one of three things must be true: 1) Investors expect dividends to rise in the future. 2) Investors expect returns to be low in the future. Future cash flows are discounted at a lower than usual rate, leading to higher prices. 3) Investors expect prices to keep rising forever, in a "bubble." This is not a theory, it is an accounting identity like 1=1: If the P/D ratio is high, either dividends must rise, prices must decline, or the P/D ratio must never return to its historical average. Which of the three options holds for our stock market?

Historically, *virtually all variation in P/D ratios has reflected varying expected returns*. At a simple level, table 4 makes this point with regressions of long-horizon dividend growth on P/D ratios to match the regressions of returns on P/D ratios. The dividend-growth coefficients are much smaller, typically one standard error from zero, and the R^2 values are tiny. Worse, the *signs* are wrong. To the extent that a high P/D ratio forecasts any change in dividends, it seems to forecast a small *decline* in dividends.

To be a little more precise, the identity

$$1 = R_{t+1}^{-1} R_{t+1} = R_{t+1}^{-1} \frac{P_{t+1} + D_{t+1}}{P_t}$$

yields, with a little algebra, the approximate identity

$$1) \quad p_t - d_t = \text{const.} + \sum_{j=1}^{\infty} \rho^j (\Delta d_{t+j} - r_{t+j}) + \lim_{j \to \infty} \rho^j (p_{t+j} - d_{t+j}),$$

where $\rho = P/D/(1 + P/D)$ is a constant of approximation, slightly less than one and lowercase letters denote logarithms (Campbell and Shiller, 1988). Equation 1 gives a precise meaning to my earlier statement that a high P/D ratio *must* be followed by high dividend growth Δd, low returns r, or a bubble.

Bubbles do not appear to be the reason for historical P/D ratio variation. Unless the P/D ratio grows faster than $1/\rho^j$, there is no bubble. It is hard to believe that P/D ratios can grow *forever*. Empirically, P/D ratios do not seem to have a trend or unit root over time.[5]

This still leaves two possibilities: are high prices signals of high dividend growth or low returns? To address this issue, equation 1 implies[6]

$$2) \quad var(p_t - d_t) = cov\left(p_t - d_t, \sum_{j=1}^{\infty} \rho^j \Delta d_{t+j}\right) - cov\left(p_t - d_t, \sum_{j=1}^{\infty} \rho^j r_{t+j}\right).$$

P/D ratios can *only* vary if they forecast changing dividend growth or if they forecast changing returns. Equation 2 has powerful implications. At first glance, it would seem a reasonable approximation that returns cannot be forecasted (the "random walk" hypothesis) and neither can dividend growth. But if this were the case, the P/D ratio would have to be a *constant*. Thus, the fact that the P/D ratio varies *at all* means that either dividend growth or returns can be forecasted.

This observation solidifies one's belief in P/D ratio forecasts of returns. Yes, the statistical evidence that P/D ratios forecast returns is weak. But P/D ratios *have* varied. The choice is, P/D ratios forecast returns *or* they forecast dividend growth. They have to forecast something. Given this choice and table 3, it seems a much firmer conclusion that they forecast returns.

Table 5 presents some estimates of the price-dividend variance decomposition (equation 2), taken from Cochrane (1991b). As one might suspect from table 4, table 5 shows that in the past *almost all variation in P/D ratios has been due to changing return forecasts.* (The rows of table 5 do not add up to exactly 100 percent because equation 2 is an approximation. The elements do not have to be between 0 and 100 percent. For example, –34, 138 occurs because high prices seem to forecast lower dividend growth. Therefore they must and do forecast really low returns. The real and nominal rows differ because P/D forecasts inflation in the sample.)

So much for history. What does it mean? Again, we live at a moment of historically unprecedented P/D, P/E, and other multiples. Perhaps *this time* high prices reflect high long-run dividend growth. If so, the prices have to reflect an unprecedented expectation of future dividend growth: the P/D ratio is about double its long-term average, so the level of dividends has to double, above and beyond its usual growth. However, if this time is at all like the past, high prices reflect low future returns.

TABLE 5

Variance decomposition of VW P/D ratio

	Dividends	Returns
Real	–34	138
Standard error	(10)	(32)
Nominal	30	85
Standard error	(41)	(19)

Notes: VW = value weighted NYSE. Table entries are the percent of the variance of the price/dividend ratio attributable to dividend and return forecasts, $100 \times cov(p_t-d_t, \Sigma_{j=1}^{15} \rho^j \Delta d_{t+j})/var(p_t-d_t)$ and similarly for returns.

The bottom line

Statistical analysis suggests that the long-term average return on broad stock market indexes is 8 percent greater than the T-bill rate, with a standard error of about 3 percent. High prices are related to low subsequent excess returns. Based on these patterns, the expected excess return (stock return less T-bill rate) is near zero for the next five years or so, and then slowly rising to the historical average. The large standard deviation of excess returns, about 17 percent, means that actual returns will certainly deviate substantially from the expected return. Finally, one always gets more expected return by taking on more risk.

Economics: Understanding the equity premium

Statistical analysis of past returns leaves a lot of uncertainty about future returns. Furthermore, it is hard to believe that average excess returns are 8 percent without knowing *why* this is so. Perhaps most important, no statistical analysis can predict if the future will be like the past. Even if the true expected excess return was 8 percent, did that result from fundamental or temporary features of the economy? Thus, we need an economic understanding of stock returns.

Economic theory and modeling is often portrayed as an ivory tower exercise, out of touch with the real world. Nothing could be further from the truth, especially in this case. Many superficially plausible stories have been put forth to explain the historically high return on stocks and the time-variation of returns. Economic models or theories make these stories explicit, check whether they are internally consistent, see if they can quantitatively explain stock returns, and check that they do not make wildly counterfactual predictions in other dimensions, for example, requiring wild variation in risk-free rates or strong persistent movements in consumption growth. Few stories survive this scrutiny.

We have a vast experience with economic theory; a range of model economies have formed the backbone of our understanding of economic growth and dynamic micro, macro, and international economics for close to 25 years. Does a large equity premium make sense in terms of such standard economics? Did people in 1947, and throughout the period, know that stocks were going to yield 8 percent over bonds on average, yet were rationally unwilling to hold more stocks because they were afraid of the 17 percent standard deviation or some other measure of stocks' risk? If so, we have "explained" the equity premium. If so, statistics from the past may well describe the future, since neither people's preferences nor the riskiness of technological opportunities seems to have changed dramatically. But what if it makes no sense that people should be so scared of stocks? In this case, it is much more likely that the true premium is small, and the historical returns were in fact just good luck.

The answer is simple: Standard economic models utterly fail to produce anything like the historical average stock return or the variation in expected returns over time. After ten years of intense effort, there is a range of drastic modifications to standard models that can explain the equity premium and return predictability and (harder still) are not inconsistent with a few obvious related facts about consumption and interest rates. However, these models are truly drastic modifications; they fundamentally change the description of the source of risk that commands a premium in asset markets. Furthermore, they have not yet been tested against the broad range of experience of the standard models. These facts must mean one of two things. Either the standard models are wrong and will change drastically, or the phenomenon is wrong and will disappear.

I first show how the standard model utterly fails to account for the historical equity premium (Sharpe ratio). The natural response is to see if perhaps we can modify the standard model. I consider what happens if we simply allow a very high level of risk aversion. The answer here, as in many early attempts to modify the standard model, is unsatisfactory. While one can explain the equity premium, easy explanations make strongly counterfactual predictions regarding other facts. The goal is to explain the equity premium in a manner consistent with the level and volatility of consumption growth (both about 1 percent per year), the predictability of stock returns described above, the relative lack of predictability in consumption growth, the relative constancy of real risk-free rates over time and across countries, and the relatively low correlation of stock returns with consumption growth. This is a tough assignment, which is only now starting to be accomplished.

Then I survey alternative views that do promise to account for the equity premium, without (so far) wildly counterfactual predictions on other dimensions. Each modification is the culmination of a decade-long effort by a large number of researchers. (For literature reviews, see Kocherlakota, 1996, and Cochrane and Hansen, 1992). The first model maintains the complete and frictionless market simplification, but changes the specification of how people feel about consumption over time, by adding *habit persistence* in a very special way that produces a strong *precautionary saving* motive. The second model abandons the perfect markets simplification. Here, uninsurable individual-level risks are the key to the equity premium. I will also discuss a part of an emerging view that the equity premium and time-variation of expected returns result from the fact that few people hold stocks. This view is not flushed out yet to a satisfactory model, but does give some insight.

Both modifications answer the basic question, "why are consumers so afraid of stocks?" in a similar way, and give a fundamentally different answer from the standard model's view that expected returns are driven by risks to wealth or consumption. The modifications both say that consumers are really afraid of stocks because stocks pay off poorly in *recessions*. In one case a recession means a time when consumption has recently fallen, no matter what its level. In the second case a recession is a time of unusually high cross-sectional (though not aggregate) uncertainty. In both cases the raw risk to wealth is not a particularly important part of the story.

The standard model

To say anything about dynamic economics, we have to say something about how people are willing to trade consumption in one moment and set of circumstances (state of nature) for consumption in another moment and set of circumstances. For example, if people were always willing to give up a dollar of consumption today for $1.10 in a year, then the economy would feature a steady 10 percent interest rate. It also might have quite volatile consumption, as people accept many such opportunities. If people did not care what the circumstances were in which they would get \$1.10, then all expected returns would be equal to the interest rate (risk neutrality). Of course, this is an extreme example. There certainly comes a point at which such willingness to substitute consumption becomes strained. If someone were going to consume \$1,000 this year but \$10,000 next year, it might take a bit more than a 10 percent interest rate to get him or her to consume even less this year.

To capture these ideas about people's willingness to substitute consumption, we use a *utility function* that gives a numerical "happiness" value for every possible stream of future consumption,

$$3) \quad U = E\int_{t=0}^{\infty} e^{-\rho t}u(C_t)dt.$$

E denotes expectation; C_t denotes consumption at date t; ρ is the subjective discount factor; and $e^{-\rho t}$ captures the fact that consumers prefer earlier consumption to later. The function $u(\cdot)$ is increasing and concave, to reflect the idea that people always like more consumption, but at a diminishing rate. The function $u(C) = C^{1-\gamma}$ is a common specification, with γ between 0 and 5. $\gamma = 0$ or $u(C) = C$ corresponds to *risk neutrality*, a constant interest rate ρ, and a perfect willingness to substitute across time. $\gamma = 1$ corresponds to $u(C) = ln(C)$, which is a very attractive choice since it implies that each doubling of consumption adds the same amount of happiness. For most asset pricing problems, writing the utility function over an infinite lifespan is a convenient simplification that makes little difference to the results. Economic models are often written in discrete time, in which case the utility function is

$$U = E\sum_{t=0}^{\infty} e^{-\rho t}u(C_t).$$

Dynamic economics takes this representation of people's preferences and mixes it with a representation of technological opportunities for production and investment. For example, the simplest model might specify that output is made from capital, $Y = f(K)$, output is invested or consumed, $Y = C+I$, and capital depreciates but is increased by investment, $K_{t+1} = (1\text{-}\delta)K_t + I_t$ in discrete time. To study business cycles, one adds detail, including at least labor, leisure, and shocks. To study monetary issues, one adds some friction that induces people to use money, and so on.

Despite the outward appearance of tension, this is a great unifying moment for macroeconomics. Practically all issues relating to business cycles, growth, aggregate policy analysis, monetary economics, and international economics are studied in the context of variants of this simple model. The remaining differences concern details of implementation.

Since this basic economic framework explains such a wide range of phenomena, what does it predict for the equity risk premium? Give the opportunity to buy assets such as stocks and bonds to a consumer whose preferences are described by equation 3, and figure out what the optimal consumption and portfolio decision is. (The appendix includes derivations of all equations.) The following conditions describe the optimal choice:

$$4) \quad r^f = \rho + \gamma E(\Delta c)$$

$$5) \quad E(r) - r^f = \gamma\, cov(\Delta c, r) = \gamma\, \sigma(\Delta c)\sigma(r)\, corr(\Delta c, r),$$

where Δc denotes the proportional change in consumption, r denotes a risky asset return, r^f denotes the risk-free rate, cov denotes covariance, $corr$ denotes correlation, and

$$\gamma \equiv -\frac{Cu''(C)}{u'(C)}$$

is a measure of *curvature* or *risk aversion*. Higher γ means that more consumption gives less pleasure very quickly; it implies that people are less willing to substitute less consumption now for more consumption later and to take risks.

Equation 5 expresses the most fundamental idea in finance. It says that the average excess return on any security must be proportional to the *covariance* of that return with marginal utility and, hence, consumption growth. This is because people value financial assets that can be used to smooth consumption over time and in response to risks. For example, a "risky" stock, one that has a high standard deviation $\sigma(r)$, may nonetheless command no greater average return $E(r)$ than the risk-free rate if its return is uncorrelated with consumption growth $corr(\Delta c,r) = 0$. If it yields any more, the consumer can buy just a little bit of the security, and come out ahead because the risk is perfectly diversifiable. Readers familiar with the capital asset pricing model (CAPM) will recognize the intuition; replacing wealth or the market portfolio with consumption gives the most modern and general version of that theory.

The equity premium puzzle

To evaluate the equity premium, I transform equation 5 to

$$6)\quad \frac{E(r)-r^f}{\sigma(r)} = \gamma\sigma(\Delta c)\,corr(\Delta c, r).$$

The left-hand side is the Sharpe ratio. As I showed above, the (unconditional) Sharpe ratio is about 0.5 for the stock market, and it is robust to leveraging or choice of assets. The right-hand side of equation 6 says something very important. A high Sharpe ratio or risk premium must be the result of 1) high aversion to risk, γ, or 2) lots of risk, $\sigma(\Delta c)$. Furthermore, it can only occur for assets whose returns are correlated with the risks. This basic message will pervade the following discussion of much-generalized economic models. If the right-hand side of equation 6 is low, then the consumer should invest more in the asset with return r. Doing so will make the consumption stream more risky and more correlated with the asset return. Thus, as the consumer invests more, the right-hand side of equation 6 will approach the left-hand side.

The right-hand side of equation 6 is a prediction of what the Sharpe ratio should be. It does not come close to predicting the historical equity premium. The standard deviation of aggregate consumption growth is about 1 percent or 0.01. The correlation of consumption growth with stock returns is a bit harder to measure since it depends on horizon and timing issues. Still, for horizons of a year or so, 0.2 is a pretty generous number. $\gamma \approx 1$ or 2 is standard; $\gamma = 10$ is a very generous value. Putting this all together, $10 \times 0.01 \times 0.2 = 0.02$ rather than 0.5. At a 20 percent standard deviation, a 0.02 Sharpe ratio implies an average excess return for stocks of $0.02 \times 20 = 0.4\%$ (40 basis points) rather than 8 percent.

This devastating calculation is the celebrated "equity premium puzzle" of Mehra and Prescott (1985), as reinterpreted by Hansen and Jagannathan (1991). The failure is quantitative not qualitative, as Kocherlakota (1996) points out. Qualitatively, the right-hand side of equation 6 does predict a positive equity premium. The problem is in the numbers. This is a strong advertisement for quantitative rather than just qualitative economics.

Can we change the numbers?

The correlation of consumption growth with returns is the most suspicious ingredient in this calculation. While the correlation is undeniably low in the short run, a decade-long rise in the stock market should certainly lead to more consumption. In fact, the low correlation is somewhat of a puzzle in itself: Standard (one-shock) models typically predict correlations of 0.99 or more. Marshall and Daniel (1997) find correlations in the data up to 0.4 at a two-year horizon, and by allowing lags. But even plugging in a correlation of $corr(\Delta c,r) = 1$, $\sigma(\Delta c) = 0.01$ and $\gamma < 10$ implies a Sharpe ratio less than 0.1, or one-fifth the sample value.

A large literature has tried to explain the equity premium puzzle by introducing frictions that make T-bills "money-like," which artificially drive down the interest rate (for example, Aiyagari and Gertler, 1991). The highest Sharpe ratio occurs in fact when one considers short-term risk-free debt and money, since the latter pays no interest. Perhaps the same mechanism can be invoked for the spread between stocks and bonds. However, a glance at figure 1 shows that this will not work. High Sharpe ratios are pervasive in financial markets. One can recover a high Sharpe ratio from stocks alone, or from stocks less long-term bonds.

Time-varying expected returns

The consumption-based view with $u'(C) = C^{-\gamma}$ also has trouble explaining the fact that P/D ratios forecast stock returns. Consider the conditional version of equation 6

$$7)\quad \frac{E_t(r)-r_t^f}{\sigma_t(r)} = \gamma\sigma_t(\Delta c)\,corr_t(\Delta c, r),$$

where E_t, σ_t, $corr_t$ represent conditional moments. I showed above that the P/D ratio gives a strong signal about mean returns, $E_t(r)$. It does not however give much information about the standard deviation of returns. Figure 5 does suggest a slight increase in return standard deviation along with the higher mean return when P/D ratios decline—the leverage effect of Black (1976). However, the increase in standard deviation is much less than the increase in mean return. Hence, the Sharpe ratio of mean to standard deviation varies over time and increases when prices are low.

How can we explain variation in the Sharpe ratio? Looking at equation 7, it could happen if there were times of high and low consumption volatility, variation in $\sigma_t(\Delta c)$. But that does not seem to be the case; there is little evidence that aggregate consumption growth is much more volatile at times of low prices than high prices. The conditional correlation of consumption growth and returns could vary a great deal over time, but this seems unlikely, or more precisely like an unfathomable assumption on which to build the central understanding of time-varying returns.

What about the CAPM?

Finance researchers and practitioners often express disbelief (and boredom) with consumption-based models such as the above. Even the CAPM performs better: Expected returns of different portfolios such as those in figure 1 line up much better against their covariances with the market return than against their covariances with consumption growth. Why not use the CAPM or other, better-performing finance models to understand the equity premium?

The answer is that the CAPM and related finance models are *useless* for understanding the market premium. The CAPM states that the expected return of a given asset is proportional to its "beta" times the expected return of the market,

$$E(R^i) = R^f + \beta_{i,m}\left[E(R^m) - R^f\right].$$

This is fine if you want to think about an individual stock's return given the market return. But the average market return—the thing we are trying to explain, understand, and predict—is a *given* to the CAPM. Similarly, multifactor models explain average returns on individual assets, *given* average returns on "factor mimicking portfolios," including the market. Option pricing models explain option prices, *given* the stock price. To understand the market premium, there is no substitute for economic models such as the consumption-based model outlined above and its variants.

Highly curved power utility

Since we have examined all the other numbers on the right-hand side of equation 6, perhaps we should raise curvature γ. This is a central modification. All of macroeconomics and growth theory considers values of γ no larger than 2–3. To generate a Sharpe ratio of 0.5, $\gamma = 250$ is needed in equation 6. Even if $corr = 1$, γ must still equal 50. What's wrong with $\gamma = 50$ to 250? Although a high curvature γ explains the equity premium, it runs quickly into trouble with other facts.

Consumption and interest rates

The most basic piece of evidence for low γ is the relationship between consumption growth and interest rates. Real interest rates are quite stable over time (see the standard deviation in table 1) and roughly the same the world over, despite wide variation in consumption growth over time and across countries. A value of $\gamma = 50$ to 250 implies that consumers are essentially unwilling to substitute consumption over time; equivalently that variation in consumption growth must be accompanied by huge variations in interest rates that we do not observe.

Look again at the first basic relationship between consumption growth and interest rates, equation 4, reproduced here:

$$r_t^f = \rho + \gamma E_t(\Delta c).$$

High values of γ are troublesome even in understanding the level of interest rates. Average consumption growth and real interest rates are both about 1 percent. Thus, $\gamma = 50$ to 250 requires $\rho = -0.5$ to -2.5, or a –50 percent to –250 percent subjective discount rate. That's the wrong sign: People should prefer current to future consumption, not the other way around (Weil, 1989).[7]

The absence of much interest rate variation across time and countries is an even bigger problem. People save more and defer consumption in times of high interest rates, so consumption growth rises when interest rates are higher. However, $\gamma = 50$ means that a country with consumption growth 1 percentage point higher than normal must have real interest rates 50 percentage points higher than normal, and consumption 1 percentage point lower than normal must have real interest rates 50 percentage points lower than normal, implying huge negative interest rates—consumers pay financial intermediaries 48 percent to keep their money. This just does not happen.

This observation can also be phrased as a conceptual experiment, suitable for thinking about one's own preferences or for survey evidence on others' preferences. For example, what does it take to convince someone to skip a vacation? Take a family with $50,000 per year income, consumption equal to income, which spends $2,500 (5 percent) on an annual vacation. If interest rates are good enough, though, the family can be persuaded to skip this year's vacation and go on two vacations next year. What interest rate does it take to persuade the family to do this? The answer is $(\$52,500/\$47,500)^{\gamma} - 1$. For $\gamma = 250$ that is an interest rate of 3×10^{11}. For $\gamma = 50$, we still need an interest rate of 14,800 percent. I think most of us would defer the vacation for somewhat lower interest rates.

The standard use of low values for γ in macroeconomics is also important for delivering realistic quantity dynamics in macroeconomic models, including relative variances of investment and output, and for delivering reasonable speeds of adjustment to shocks.

Risk aversion

Economists have also shied away from high curvature γ on the basis that people do not seem that risk averse. After examining the argument, I conclude that there are fewer solid reasons to object to high risk aversion than to object to high aversion to intertemporal substitution via the consumption-interest rate relationships I examined above.

Surveys and thought experiments

Since Sharpe ratios are high for many assets, much analysis of risk aversion comes from simple thought experiments rather than data. For example, how much would a family pay per year to avoid a bet that led with equal probability to a $*y* increase or decrease in annual consumption for the rest of their lives? Table 6 presents some calculations of how much our family with $50,000 per year of income and consumption would pay to avoid various bets of this form.[8] For bets that are reasonably large relative to wealth, high γ means that families are willing to pay almost the entire amount of the bet to avoid taking it. For example, in the lower right-hand corner, the family with $\gamma = 250$ would rather pay $9,889 for sure than take a 50 percent chance of a $10,000 loss. This prediction is surely unreasonable, and has led most authors to rule out risk aversion coefficients over ten. Survey evidence for this kind of bet also finds low risk aversion, certainly below $\gamma = 5$ (Barsky, Kimball, Juster, and Shapiro, 1997), and even negative risk aversion if the survey is taken in Las Vegas.

Yet the results for small bets are not so unreasonable. The family might reasonably pay 5 cents to 25 cents to avoid a $10 bet. We are all risk neutral for small enough bets. For small bets,

$$\frac{\text{amount willing to pay to avoid bet}}{\text{size of bet}} \approx \gamma \frac{\text{size of bet}}{\text{consumption}}.$$

Thus, for any γ, the amount one is willing to pay is an arbitrarily small fraction of the bet for small enough bets. For this reason, it is easy to cook numbers of conceptual experiments like table 6 by varying the size of the bet and the presumed wealth of the family. Significantly, I only used *local* curvature above; γ represented the derivative $\gamma = -Cu''(C)/u'(C)$. In asking how much the family would pay to avoid a $10,000 bet, we are asking for the response to a very, very non-local event.

The main lesson of conceptual experiments and laboratory and survey evidence of simple bets is that people's answers to such questions routinely violate expected utility. This observation lowers the value of this source of evidence as a measurement of risk aversion. As a similar cautionary note, Barsky et al. report that whether an individual partakes in a wide variety of risky activities correlates poorly with the level

TABLE 6

Amount family would pay to avoid an even bet

	Risk aversion γ				
Bet	**2**	**10**	**50**	**100**	**250**
$10	$0.00	$0.01	$0.05	$0.10	$0.25
100	0.20	1.00	4.99	9.94	24
1,000	20	99	435	665	863
10,000	2,000	6,921	9,430	9,718	9,889

Notes: I assume the family has a constant $50,000 per year consumption and an even chance of winning or losing the indicated net, per year.

of risk aversion inferred from a survey. In the end, surveys about hypothetical bets that are far from the range of everyday experience are hard to interpret.

Microeconomic evidence

Microeconomic observations might be a more useful measure of risk aversion, that is, evidence from people's actual behavior in their daily activities. For example, the numbers in table 6 could be matched with insurance data. People are willing to pay substantially more than actuarially fair values to insure against car theft or house fires. What is the implied risk aversion? But even if there are other markets whose prices reflect less risk aversion than stocks, this leaves open the question: If people are risk-neutral in other markets, why do they become risk averse in the stockbroker's office? Perhaps the risk aversion people display in the stock broker's office should be the fact and the (possibly) low risk aversion displayed in other offices should be the puzzle.

Portfolio calculations

A common calibration of risk aversion comes from simple portfolio calculations (see Friend and Blume, 1975). Following the principle that the last dollar spent should give the same increase in happiness in any alternative use, the marginal value of wealth should equal the marginal utility of consumption,[9] $V_w(W,.) = u_c(C)$. Therefore, if we assume returns are independent over time and no other variables are important for the marginal value of wealth, $V_w(W)$, equation 6 can also be written as

$$8)\quad \frac{E(r)-r^f}{\sigma(r)} = \frac{-WV_{WW}}{V_W}\sigma(\Delta w)\,corr(\Delta w, r).$$

The quantity $-WW_{ww}/V_w$ is in fact a better measure of risk aversion than $-Cu_{cc}/u_c$, since it represents aversion to bets over wealth rather than bets over consumption; most bets observed are paid off in dollars. For a consumer who invests entirely in stocks, $\sigma(\Delta w)$ is the standard deviation of the stock return and $corr(\Delta w,r) = 1$. To generate a Sharpe ratio of 0.5, it seems that we only need risk aversion equal to 3,

$$\frac{-WV_{WW}}{V_W} = \frac{0.5}{0.17} \cong 3.$$

The Achilles heel of this calculation is the hidden simplifying assumption that returns are independent over time, so no variables other than wealth show up in V_w. If this were the case, consumption would move one-for-one with wealth, and $\sigma(\Delta c) = \sigma(\Delta w)$. If wealth doubles and nothing else has changed, the consumer would double consumption. The calculation, therefore, hides a model with the drastically counterfactual implication that consumption growth has a 17 percent standard deviation!

The fact that consumption has a standard deviation so much lower than that of stock returns suggests that returns are not independent over time (as is already known from the return on P/D regressions) and/or that other state variables must be important in driving stock returns. If some other state variable, z, is allowed—representing subsequent expected returns, labor income, or some other measure of a consumer's overall opportunities—the substitution $V_w(W,z) = u_c(C)$ in equation 6 adds another term to equation 8,

$$\frac{E(r)-r^f}{\sigma(r)} = \frac{-WV_{WW}}{V_W}\sigma(\Delta w)\,corr(\Delta w, r) + \frac{zW_{Wz}}{V_W}\sigma(z)\,corr(z,r).$$

The Sharpe ratio may be driven not by consumers' risk aversion and the wealth-riskiness of stocks, but by stocks' exposure to other risks.

In the current context, this observation just tells us that portfolio-based calibrations of risk aversion do not work, because they implicitly assume independent returns and, hence, consumption growth as volatile as returns. Below, I introduce plausible candidates for the variable z that can help us to understand high Sharpe ratios. The fact that consumption is so much less volatile than shock returns indicates that the other state variables must account for the bulk of the equity premium.

Overall, the evidence against high risk aversion is not that strong, and it is at least a possibility to consider. This argument does not rescue the power utility model with $\gamma = 50$ to 250—that ship sank on the consumption-interest rate shoals. However, other models with high risk aversion can be contemplated.

New utility functions and state variables

If changing the parameter γ in $u'(C) = C^{-\gamma}$ does not work, perhaps we need to change the *functional form*. Changing the form of $u(C)$ is not a promising avenue. As I have stressed by using a continuous time derivation, only the derivatives of $u(C)$ really matter; hence quite similar results are achieved with other functional forms. A more promising avenue is to consider other arguments of the utility function, or *nonseparabilities.*

Perhaps how people feel about eating more today is affected not just by how much they are already eating, but by other things, such as how much they ate yesterday or how much they worked today. Then, the covariance of stock returns with these other variables will also determine the equity premium. Fundamentally, consumers use assets to smooth marginal utility. Perhaps today's marginal utility is related to more than just today's consumption.

Such a modification is a fundamental change in how we view stock market risk. For example, perhaps more leisure raises the marginal utility of consumption. Stocks are then feared because they pay off badly in recessions when employment is lower and leisure is higher, not because consumers are particularly averse to the risk that stocks decline per se. Formally, our fundamental equation 6 is derived from

$$E_t(r) - r_t^f = cov_t(\Delta u_c, r),$$

and substituting $u_c = \partial u(c)/\partial c$. If I substitute $u_c = \partial u(c,x)/\partial c$ instead, then u_c will depend on other variables x as well as c, and

$$E(r) - r^f = \gamma\, cov(\Delta c, r) + \frac{u_{cx}}{u_c} cov(x, r).$$

Since the first covariance does not account for much premium, we will have to rely heavily on the latter term to explain the premium.

There is a practical aim to generalizing the utility function as well. As illustrated in the last section, one parameter γ did two things with power utility: It controlled how much people are willing to substitute consumption over time (consumption and interest rates) and it controlled their attitudes toward risk. The choice $\gamma = 50$ to 250 was clearly a crazy representation of how people feel about consumption variation over time, but perhaps not so bad a representation of risk aversion. Maybe a modification of preferences can disentangle the two attitudes.

State separability and leisure

With the latter end in mind, Epstein and Zin (1989) started an avalanche of academic research on utility functions that relax state-separability. The expectation E in the utility function of equation 3 sums over states of nature, for example

$$U = prob(\text{rain}) \times u(C \text{ if it rains}) + prob(\text{shine}) \times u(C \text{ if it shines}).$$

"Separability" means that one adds utility across states, so the marginal utility of consumption in one state is unaffected by what happens in another state. But perhaps the marginal utility of a little more consumption in the sunny state of the world is affected by the level of consumption in the rainy state of the world.

Epstein and Zin and Hansen, Sargent, and Tallarini (1997) propose recursive utility functions of the form

$$U_t = C_t^{1-\gamma} + e^{-\rho} f\left[E_t f^{-1}(U_{t+1})\right].$$

If $f(x) = x$, this expression reduces to power utility. These utility functions are not state-separable, and do conveniently distinguish risk aversion from intertemporal substitution among other modifications. However, this research is only starting to pay off in terms of plausible models that explain the facts (Campbell, 1996, is an example) so I will not review it here.

Perhaps leisure is the most natural extra variable to add to a utility function. It is not clear *a priori* whether more leisure enhances the marginal utility of consumption (why bother buying a boat if you are at the office all day and cannot use it?) or vice versa (if you have to work all day, it is more important to come home to a really nice big TV), but we can let the data speak on this matter. However, explicit versions of this approach have not been very successful to date. (Eichenbaum, Hansen, and Singleton, 1988, for example). On the other hand, recent research has found that adding labor income growth as an extra ad-hoc factor can be useful in explaining the cross section of average stock returns (Jagannathan and Wang, 1996; Reyfman, 1997). Though not motivated by an explicit utility function, the facts in this research may in the future be interpretable as an effect of leisure on the marginal utility of consumption.

Force of habit

Nonseparabilities over *time* have been more useful in empirical work. Anyone who has had a large pizza dinner knows that yesterday's consumption can have an impact on today's appetite. Might a similar mechanism apply over a longer time horizon? Perhaps people get accustomed to a standard of living, so a fall in consumption hurts after a few years of good times, even though the same level of consumption might have seemed very pleasant if it arrived after years of bad times. This view at least explains the perception that recessions are awful events, even though a recession year may be just the second or third best year in human history rather than the absolute best. Law, custom, and social insurance insure against falls in consumption as much as low levels of consumption.

Following this idea, Campbell and Cochrane (1997) specify that people slowly develop habits for higher or lower consumption. Technically, they replace the utility function $u(C)$ with

$$9)\quad u(C-X)=(C-X)^{1-\eta},$$

where X represents the level of habits. In turn, habit X adjusts slowly to the level of consumption.[10] (I use the symbol η for the power, because curvature and risk aversion no longer equal η.) This specification builds on a long tradition in the microeconomic literature (Duesenberry, 1949, and Deaton, 1992) and recent asset-pricing literature (Constantinides, 1990; Ferson and Constantinides, 1991; Heaton, 1995; and Abel, 1990).

When a consumer has such a habit, local curvature depends on how far consumption is above the habit, as well as the power η,

$$\gamma_t \equiv \frac{-C_t\, u_{cc}}{u_c}=\eta\frac{C_t}{C_t-X_t}.$$

As consumption falls toward habit, people become much less willing to tolerate further falls in consumption; they become very risk averse. Thus, a low power coefficient η (Campbell and Cochrane use η = 2) can still mean a high and time-varying curvature.

Recall our fundamental equation 6 for the Sharpe ratio,

$$\frac{E_t(r)-r_t^f}{\sigma_t(r)}=\gamma_t\sigma_t(\Delta c)\,corr_t(\Delta c, r).$$

High curvature γ_t means that the model can explain the equity premium, and curvature that varies over time, as consumption rises in booms and falls toward habit in recessions, means that the model can explain a time-varying and countercyclical (high in recessions, low in booms) Sharpe ratio, despite constant consumption volatility, $\sigma_t(\Delta c)$, and correlation, $corr_t(\Delta c, r)$.

So far so good, but doesn't raising curvature imply high and time-varying interest rates? In the Campbell-Cochrane model, the answer is no. The reason is precautionary saving. Suppose times are bad and consumption is low relative to habit. People want to borrow against future, higher consumption and this should drive up interest rates. However, people are also much more risk averse in bad times; they want to save more in case tomorrow might be even worse. These two effects balance.

The precautionary saving motive also makes the model more plausibly consistent with variation in consumption growth across time and countries. The interest rate in the model adds a precautionary savings motive term to equation 4,

$$r^f=\rho+\eta E(\Delta c)-\frac{1}{2}\left(\frac{\eta}{\bar{S}}\right)^2\sigma^2(\Delta c),$$

where $\bar{S}$ denotes the steady state value of $(C–X)/C$, about 0.05. The power coefficient, η = 2, controls the relationship between consumption growth and interest rates, while the curvature coefficient, $\gamma = \eta\, C/(C–X)$, controls the risk premium. Thus, this habit model allows high risk aversion with low aversion to intertemporal substitution and is consistent with the consumption and interest rate data and a sensible value of ρ.

Campbell and Cochrane create a simple artificial economy with these preferences. Consumption growth is independent over time and real interest rates are constant. They calculate time series of stock prices and interest rates in the artificial economy and subject them to the standard statistical analysis reviewed above. The artificial data replicate the equity premium (0.5 Sharpe ratio). The ability to forecast returns from the P/D ratio and the P/D variance decomposition are both quite like the actual data. The standard deviation of returns rises a bit when prices decline, but less than the rise in mean returns, so a low P/D ratio forecasts a higher Sharpe ratio. Artificial data from the model also replicate much of the low observed correlation between consumption growth and returns, and the CAPM and ad-hoc multifactor models perform better than the power utility consumption-based model in the artificial data.

The model also provides a good account of P/D fluctuations over the last century, based entirely on the history of consumption. However, it does not account for the currently high P/D ratio. This is because the model generates a high P/D ratio when consumption is very high relative to habit and, therefore, risk aversion is low. Measured consumption has been increasing unusually slowly in the 1990s.

Like other models that explain the equity premium and return predictability, this one does so by fundamentally changing the story of why consumers are afraid of holding stocks. From equation 9, the marginal utility of consumption is proportional to

$$u_c = C_t^{-\eta}\left(\frac{C_t - X_t}{C_t}\right)^{-\eta}.$$

Thus, consumers dislike low consumption as before, but they are also afraid of recessions, times when consumption, whatever its level, is low relative to the recent past as described by habits. Consumers are afraid of holding stocks not because they fear the wealth or consumption volatility per se, but because bad stock returns tend to happen in recessions, times of a recent belt-tightening.

This model fulfills a decade-long search kicked off by Mehra and Prescott (1985). It is a complete-markets, frictionless economy that replicates not only the equity premium but also the predictability of returns, the nearly constant interest rate, and the near-random walk behavior of consumption.

Habit models with low risk aversion

The individuals in the Campbell-Cochrane model are highly risk averse. They would respond to surveys about bets on wealth much as the $\gamma = 50$ column of table 6. The model does not give rise to a high equity premium with low risk aversion; it merely disentangles risk aversion and intertemporal substitution so that a high risk aversion economy can be consistent with low and constant interest rates, and it generates the predictability of stock returns.

Constantinides (1990) and Boldrin, Christiano, and Fisher (1995) explore habit persistence models that can generate a large equity premium without large risk aversion. That is, they create artificial economies in which consumers simultaneously shy away from stocks with a very attractive Sharpe ratio of 0.5, yet would happily take bets with much lower rewards.

Suppose a consumer wins a bet or enjoys a high stock return. Normally, the consumer would instantly raise consumption to match the new higher wealth level. But consumption is addictive in these models: Too much current consumption will raise the future habit level and blunt the enjoyment of future consumption. Therefore, the consumer increases consumption slowly and predictably after the increase in wealth. Similarly, the consumer would borrow to slowly decrease consumption after a decline in wealth, avoiding the pain of a sudden loss at the cost of lower long-term consumption.

The fact that the consumer will choose to spread out the consumption response to wealth shocks means that the consumer is not averse to wealth bets. If consumption responds little to a wealth shock, then marginal utility of consumption, $u_c(C)$, also responds little, as does the marginal value of wealth, $V_W(W,\cdot) = u_c$. Risk aversion to wealth bets is measured by the change of marginal utility when wealth changes ($\partial \ln V_W/\partial \ln W = -WV_{WW}/V_W$).

The argument is correct, but shows the problem with these models. The change in consumption in response to wealth is not eliminated, it is simply deferred. Thus, these models have trouble with long-run behavior of consumption and asset returns.

If consumption growth is considered independent over time (formally, an endowment economy), which is a good approximation to the data, the model must feature strong interest rate variation to keep consumers from trying to adapt consumption smoothly to wealth shocks. For example, consumers all want to save if wealth goes up, thereby slowly increasing consumption. For consumption growth to remain unpredictable, we must have a strong decline in the interest rate at the same time as the wealth increase. Of course, interest rates are in fact quite stable and, if anything, slightly positively correlated with stock returns.

Alternatively, interest rates may be fixed to be constant over time as in Constantinides (1990) (formally, linear technologies). But then there is no force to stop consumers from slowly and predictably raising consumption after a wealth shock. Thus, such models predict counterfactually that consumption growth is positively auto-correlated over time, and that long-run consumption growth shares the high volatility of long- and short-run wealth.

The Campbell-Cochrane habit model avoids these long-run problems with precautionary savings. In response to a wealth shock, consumers with the Campbell-Cochrane version of habit persistence would also like to save more for intertemporal substitution reasons, but they also feel less risk averse and so want to save less for precautionary savings reasons. The balance means that consumption can be a random walk with constant interest rates, consistent with the data. However, it means that consumption does move right away, so u_c and V_W are affected by the wealth shock, and the consumers are highly risk averse. In this model, wealth (stock prices) comes back toward consumption after a shock, so that long-run wealth shares the low volatility of long- and short-run consumption, and high stock prices forecast low subsequent returns.

A finance perspective

To get a high equity premium with low risk aversion, we need to find some crucial characteristic that separates stock returns from wealth bets. This is a difficult task. After all, what are stocks if not a bet? The answer must be some additional state variable. Having a stock pay off badly must tell you additional bad news that losing a bet does not; therefore, people shy away from stocks even though they would happily take a bet with the same mean and variance.

In the context of perfect-markets models without leisure or other goods, the only real candidates for extra state variables are variables that describe changes in expected returns. If stock prices rise, the consumer learns something important that is not learned from winning a bet: The consumer learns that future stock returns are going to be lower. The trouble is the sign of this relationship. Lower returns in the future are bad news.[11] Stocks act as a hedge for this bad news; they go *up* just at the time one gets bad news about future returns. This consideration makes stocks *more* desirable than pure bets. Thus, considering time variation in expected returns requires even *more* risk aversion to explain the equity premium.

Consistency with individual consumption behavior

The low risk aversion models face one more important hurdle: microeconomic data. Suppose an individual receives a wealth shock (wins the lottery), not shared by everyone else. For aggregate wealth shocks, we could appeal to interest rate variation to avoid the prediction that consumption would grow slowly and predictably. However, interest rates can't change in response to an individual wealth shock. Thus, we are stuck with the prediction that the individual's consumption will increase slowly and predictably. The huge literature on individual consumption (see Deaton, 1992, for survey and references) almost unanimously finds the opposite. People who receive windfalls consume too much, too soon, and have typically spent it all in a few years. The literature abounds with "liquidity constraints" to explain the "excess sensitivity" of consumption. The Campbell-Cochrane model avoids the prediction of slowly increasing consumption by specifying an *external* habit; each person's habit responds to everyone else's consumption, related to the need to "keep up with the Joneses," as advocated by Abel (1990). Though the external specification has little impact on aggregate consumption and prices, it means that individual consumption responds fully and immediately to individual wealth shocks, because there is no need for individuals to worry about habit formation. The downside is, again, high risk aversion.

In the end, there is currently no representative agent model with low risk aversion that is consistent with the equity premium, the stability of real interest rates, nearly unpredictable consumption growth, and return predictability of the correct sign.

Heterogeneous agents and idiosyncratic risks

In the above discussion, I did not recognize any difference between people. Everyone is different, so why bother looking at representative agent or complete market models? While making an assumption such as "all people are identical" seems obviously foolish, it is not foolish to hope that we can use aggregate behavior to make sense of aggregate data, without explicitly taking account of the differences between people. While differences are there, one hopes they are not relevant to the basic story. However, seeing the difficulties that representative agent models face, perhaps it is time to see if the (aggregate) equity premium does in fact surface from differences between people rather than common behavior.

The empirical hurdle

Idiosyncratic risk explanations face a big empirical challenge. Look again at the basic Sharpe ratio equation 6,

$$\frac{E(r)-r^f}{\sigma(r)}=\gamma\sigma(\Delta c)corr(\Delta c,r).$$

This relationship should hold for every (any) consumer or household. At first sight, thinking about individuals seems promising. After all, individual consumption is certainly more variable than aggregate consumption at 1 percent per year, so we can raise $\sigma(\Delta c)$. However, this argument fails quantitatively. First, it is inconceivable that we can raise $\sigma(\Delta c)$ enough to account for the equity premium. For example, even if individual consumption has a standard deviation of 10 percent per year, and maintaining a generous limit $\gamma < 10$, we still predict a Sharpe ratio no more than $10 \times 0.1 \times 0.2 = 0.2$. To explain the 0.5 Sharpe ratio with risk aversion $\gamma = 10$, we have to believe that individual consumption growth has a 25 percent per year standard deviation; for a more traditional $\gamma = 2.5$, we need 100 percent per year standard deviation. Even 10 percent per year is a huge standard deviation of consumption growth. Remember, we are considering the risky or uncertain part of consumption growth. Predictable increases or decreases in consumption due to age and life-cycle effects, expected raises, and so on do not count. We are also thinking of the flow of consumption (nondurable goods, services) not the much more variable purchases of durable goods, such as cars and houses.

More fundamentally, the addition of idiosyncratic risk lowers the correlation between consumption growth and returns, which lowers the predicted Sharpe ratio. Idiosyncratic risk is, by its nature, idiosyncratic. If it happened to everyone, it would be aggregate risk. Idiosyncratic risk cannot therefore be correlated with the stock market, since the stock market return is the same for everyone.

For a quantitative example, suppose that individual consumption of family i, Δc^i, is determined by aggregate consumption, Δc^a, and idiosyncratic shocks (such as losing your job), ε^i,

$$\Delta c^i = \Delta c^a + \varepsilon^i.$$

For the risk ε^i to average to zero across people, we must have $E(\varepsilon^i) = 0$ and $E(\varepsilon^i \mid \Delta c^a) = E(\varepsilon^i \mid r) = 0$. Then, the standard deviation of individual consumption growth does increase with the size of idiosyncratic risk,

$$\sigma^2(\Delta c^i)=\sigma^2(\Delta c^a)+\sigma^2(\varepsilon^i).$$

But the correlation between individual consumption growth and aggregate returns declines in exact proportion as the standard deviation $\sigma(\Delta c^i)$ rises,

$$\frac{E(r)-r^f}{\sigma(r)}=\gamma\frac{cov(\Delta c^a+\varepsilon^i,r)}{\sigma(r)}=\gamma\frac{cov(\Delta c^a,r)}{\sigma(r)}.$$

Therefore, *the equity premium is completely unaffected by idiosyncratic risk.*

The theoretical hurdles

The theoretical challenge to idiosyncratic risk explanations is even more severe. We can easily construct models in which consumers are given lots of idiosyncratic income risk. But it is very hard to keep consumers from insuring themselves against those risks, producing a very steady consumption stream and a low equity premium.

Start by handing out income to consumers; call it "labor income" and make it risky by adding a chance of being fired. Left to their own devices, consumers would come up with unemployment insurance to share this risk, so we have to close down or limit markets for labor income insurance. Then, consumers who lose their jobs will borrow against future income to smooth consumption over the bad times, achieving almost the same consumption smoothness. We must shut down these markets too.

However, nothing stops our borrowing-constrained consumers from saving. They build up a stock of durable goods, government bonds, or other liquid assets that they can draw down in bad times and again achieve a very smooth consumption stream (Telmer, 1993, and Lucas, 1994). To shut down this avenue for consumption-smoothing, we can introduce large transactions costs and ban from the model the simple accumulation of durable goods. Alternatively, we can make idiosyncratic shocks permanent, because borrowing and saving can only insure against transitory income. If losing your job means losing it forever, there is no point in borrowing and planning to pay it back when you get a new job.

Heaton and Lucas (1996a) put all these ingredients together, calibrating the persistence of labor income shocks from microeconomic data. They also allow the cross-sectional variance of shocks to increase in a downturn, a very helpful ingredient suggested by Mankiw (1986) that I discuss in detail in the next section. Despite all of these ingredients, their model explains at best one-half of the observed excess average stock return (and this much only with no net debt). It also predicts counterfactually that interest rates are as volatile as stock returns, and that individuals have huge (10 percent to 30 percent, depending on specification) consumption growth uncertainty.

A model that works

Constantinides and Duffie (1996) construct a model in which idiosyncratic risk can be tailored to generate any pattern of aggregate consumption and asset prices; it can generate the equity premium, predictability, relatively constant interest rates, smooth and unpredictable aggregate consumption growth, and so forth. Furthermore, it requires no transactions costs, borrowing constraints, or other frictions, and the individual consumers can have any nonzero value of risk aversion.

As mentioned earlier, if consumers are given idiosyncratic income that is correlated with the market return, they will trade away that risk. Constantinides and Duffie therefore specify that the *variance* of idiosyncratic risk rises when the market declines. Variance cannot be traded away. In addition, if marginal utility were linear, an increase in variance would have no effect on the average level of marginal utility. The interaction of cross-sector variance correlated with the market and nonlinear marginal utility produces an equity premium.

The Constantinides-Duffie model and the Campbell-Cochrane model are quite similar in spirit, though the Constantinides-Duffie model is built on incomplete markets and idiosyncratic risks, while the Campbell-Cochrane model is in the representative-agent frictionless and complete market tradition.

First, both models make a similar, fundamental change in the description of stock market risk. Consumers do not fear the loss of wealth of a bad market return so much as they fear a bad return in a recession, in one model a time of heightened labor market risk and in the other a fall of consumption relative to the recent past. This recession state variable or risk factor drives most expected returns.

Second, both models require high risk aversion. While Constantinides and Duffie's proof shows that one can dream up a labor income process to rationalize the equity premium for any risk aversion coefficient, I argue below that even vaguely plausible characterizations of actual labor income uncertainty will require high risk aversion to explain the historical equity premium.

Third, both models provide long-sought demonstrations that it is possible to rationalize the equity premium in their respective class of models. This is particularly impressive in Constantinides and Duffie's case. Many researchers (myself included) had come to the conclusion that the effort to generate an equity premium from idiosyncratic risk was hopeless.

The open question in both cases is empirical. The stories are consistent; are they right? For Constantinides and Duffie, does actual individual labor income behave as their model requires in order to generate the equity premium? The empirical work remains to be done, but here are some of the issues.

A simple version of the model

Each consumer i has power utility,

$$U = E\sum_t e^{-\rho t} C_{it}^{1-\gamma}.$$

Individual consumption growth, C_{it}, is determined by an independent, idiosyncratic normal (0,1) shock, η_{it},

$$10)\quad \ln\left(\frac{C_{it}}{C_{i,t-1}}\right)=\eta_{it}y_t-\frac{1}{2}y_t^2,$$

where y_t is the cross-sectional standard deviation of consumption growth at any moment in time. y_t is specified so that a *low* market return, R_t, gives a high cross-sectional variance of consumption growth,

$$11)\quad y_t=\sigma\left[\ln\left(\frac{C_{it}}{C_{it-1}}\right)\middle|R_t\right]$$
$$=\sqrt{\frac{2}{\gamma(\gamma+1)}}\sqrt{\ln\frac{1}{R_t}+\rho}.$$

Since η_{it} determines consumption *growth*, the idiosyncratic shocks are permanent, which I argued above was important to keep consumers from smoothing them away.

Given this structure, the individual is exactly happy to consume C_{it} and hold the stock (we can call C_{it} income I_{it} and prove the optimal decision rule is to set $C_{it} = I_{it}$.). The first-order condition for an optimal consumption-portfolio decision

$$1=E_{t-1}\left[e^{-\rho}\left(\frac{C_{it}}{C_{it-1}}\right)^{-\gamma}R_{t+1}\right]$$

holds, exactly.[12]

The general model

The actual Constantinides-Duffie model is much more general than the above example. They show that the idiosyncratic risk can be constructed to price exactly a large collection of assets, not just one return as in the example, and they allow uncertainty in aggregate consumption. Therefore, they can tailor the idiosyncratic risk to exactly match the Sharpe ratio, return forecastability, and consumption-interest rate facts as outlined above.

In the general model, Constantinides and Duffie define

$$12)\quad y_t=\sqrt{\frac{2}{\gamma(\gamma+1)}}\sqrt{\ln m_t+\rho+\gamma\ln\frac{C_t}{C_{t-1}}},$$

where m_t is a strictly positive discount factor[13] that prices all assets under consideration. That is, m_t satisfies

$$14)\quad 1=E_{t-1}\left[m_tR_t\right]\ \text{for all}\ R_t.$$

Then, they let

$$\begin{aligned}C_{it}&=\delta_{it}C_{it}\\ \delta_{it}&=\delta_{it-1}\exp\left[\eta_{it}y_t-\frac{1}{2}y_t^2\right].\end{aligned}$$

Following exactly the same argument in the text, we can now show that

$$1=E_{t-1}\left[e^{-\rho}\left(\frac{C_{it}}{C_{it-1}}\right)^{-\gamma}R_{t+1}\right]$$

for all the assets priced by m.

Microeconomic evaluation and risk aversion

Like the Campbell-Cochrane model, this model could be either a new view of stock market (and macroeconomic) risk or just a theoretically interesting existence proof. The first question is whether the microeconomic picture painted by this model is correct, or even plausible. Is idiosyncratic risk large enough? Does idiosyncratic risk really rise when the market falls, and enough to account for the equity premium? Do people really shy away from stocks because stock returns are low at times of high labor market risk?

This model does not change the empirical puzzle discussed earlier. To get power utility consumers to shun stocks, they still must have tremendously volatile consumption growth or high risk aversion. The point of this model is to show how consumers can get stuck with high consumption volatility in equilibrium, already a difficult task.

More seriously than volatility itself, consumption growth variance also represents the amount by which the distribution of individual consumption and income spreads out over time, since the shocks must be permanent and independent across people. The 10 percent to 50 percent volatility ($\sigma(\Delta c)$) that is required to reconcile the Sharpe ratio with low risk aversion means that the distribution of consumption (and income) must also *spread out* by 10 percent to 50 percent per year.

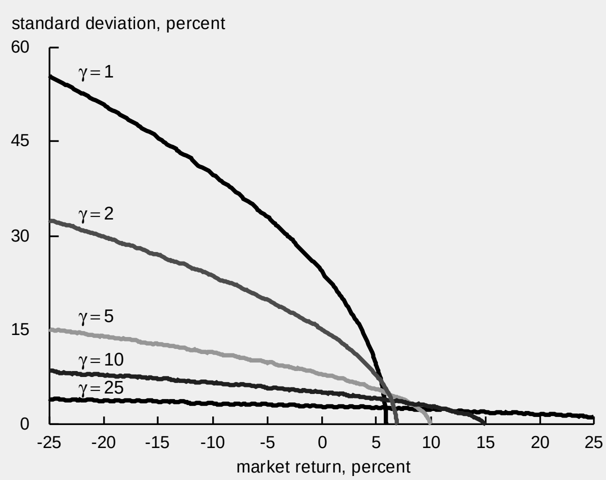
Constantinides and Duffie show how to avoid the implication that the overall distribution of income spreads out, by limiting inheritance and repopulating the economy each year with new generations that are born equal. But the distribution of consumption must still spread out within each generation to achieve the equity premium with low risk aversion. Is this plausible? Deaton and Paxson (1994) report that the cross-sectional variance of log consumption within an age cohort rises from about 0.2 at age 20 to 0.6 at age 60. This means that the cross-sectional standard deviation of consumption rises from $\sqrt{0.2}$ = 0.45 or 45 percent at age 20 to $\sqrt{0.6}$ = 0.77 or 77 percent at age 60 (77 percent means that an individual one standard deviation better off than the mean consumes 77 percent more than the mean consumer). This works out to about 1 percent per year, not 10 percent or so.

Finally, the cross-sectional uncertainty about individual income must not only be large, it must be higher when the market is lower. This risk factor is after all the central element of Constantinides and Duffie's explanation for the market premium. Figure 6 shows how the cross-sectional standard deviation of consumption growth varies with the market return and risk aversion in my simple version of Constantinides and Duffie's model. If we insist on low (γ = 1 to 2) risk aversion, the cross-sectional standard deviation of consumption growth must be extremely sensitive to the level of the market return. Looking at the γ = 2 line for example, is it plausible that a year with 5 percent market return would show a 10 percent cross-sectional variation in consumption growth, while a mild 5 percent decline in the market is associated with a 25 percent cross-sectional variation?

The Heaton and Lucas (1996a) model can be regarded as an empirical assessment of these issues. Rather than constructing a labor income process that generates an equity premium, they calibrated the labor income process from microeconomic data. They found less persistence and less increase in cross-sectional variation with a low market return than specified by Constantinides and Duffie, which is why their model predicts a low equity premium with low risk aversion. Of course, this view is at best preliminary evidence. They did not test the exact Constantinides-Duffie specification as a special case, nor did they test whether one can reject the Constantinides-Duffie specification.

FIGURE 6

Cross-sectional standard deviation of consumption growth



Notes: Cross-sectional standard deviation of individual consumption growth as a function of the market return in the Constantinides-Duffie model.

The plot is the variable $y_t = \sqrt{\frac{2}{\gamma(\gamma+1)}}\sqrt{\ln\frac{1}{R_t}+\rho+\gamma\ln\frac{C_t}{C_{t-1}}}$. Parameter values are ρ = 0.05, $\ln C_t/C_{t-1}$ = 0.01, and γ and $\ln R_{t+1}$ as graphed.

All of these empirical problems are avoided if we allow high risk aversion rather than a large risk to drive the equity premium. The γ = 25 line in figure 6 looks possible; a γ = 50 line would look even better. With high risk aversion, we do not need to specify highly volatile individual consumption growth, spreading out of the income distribution, or dramatic sensitivity of the cross-sectional variance to the market return. As in any model, a high equity premium must come from a large risk, or from large risk aversion. Labor market risk correlated with the stock market does not seem large enough to account for the equity premium without high risk aversion.

Segmented markets

All these models try to answer the basic question, if stocks are so attractive, why have people not bought more of them? So far, I have tried to find representations of people's preferences or circumstances, or a description of macroeconomic risk, in which stocks aren't that attractive after all. Then the high Sharpe ratio is a compensation for risk.

Instead, we could argue that stocks really are attractive, but a variety of market frictions keep people from buying them. This approach yields some important insights. First of all, stock ownership has been quite concentrated. The vast majority of American households have not directly owned any stock or mutual funds. One might ask whether the consumption of people who do own stock lines up with stock returns. Mankiw and Zeldes (1991) find that stockholders do have consumption that is more volatile and more correlated with stock returns than non-stockholders. But it is still not volatile and correlated enough to satisfy the right-hand side of equation 6 with low risk aversion.

Heaton and Lucas (1996b) look at individual asset and income data. They find that the richest households, who own most of the stocks, also get most of their income from proprietary business income. This income is likely to be more correlated with the stock market than is individual labor income. Furthermore, they find that among rich households, those with more proprietary income hold fewer stocks in their portfolios. This paints an interesting picture of the equity premium: In the past most stocks were held by rich people, and most rich people were proprietors whose other income (and consumption) was quite volatile and covaried strongly with the market. This is a hard crowd to sell stocks to, so they have required a high risk premium. The Campbell-Cochrane and Constantinides-Duffie models specify that stock market risk is spread as evenly as possible through the population, whereas if the risk is shared among a small group of people, higher rewards will have to be offered to offset that risk.

These views are still not sorted out quantitatively. We don't know why rich stockholders don't buy even more stocks, given low risk aversion and the tyrannical logic of equation 6. We don't know why only rich people held stocks in the first place: The literature shows that even quite high transactions costs and borrowing constraints should not be enough to deter people with low risk aversion from holding stocks.

If these segmented market views of the past equity premium are correct, they suggest that the future equity premium will be much lower. Transaction costs are declining through financial deregulation and innovation. The explosion in tax-deferred pension plans and no-load mutual funds means more and more people own stocks, spreading risks more evenly, driving up prices, and driving down prospective returns. Equation 6 will hold much better for the average consumer in the future. One would expect to see a lower equity premium. One would also expect consumption that is more volatile and more closely correlated with the stock market, which will result in a fundamental change in the nature and politics of business cycles.

Technology and investment

So far, I have tried to rationalize stock returns from the consumer's point of view: Does it make sense that consumers should not have tried to buy more stocks, driving stock returns down toward bond returns? I can ask the same questions for the firm: Do firms' investment decisions line up with stock prices as they should?

The relative prices of apples and oranges are basically set by technology, the relative number of apples versus oranges that can be grown on the same acre of land. We do not need a deep understanding of consumers' desires to figure out what the price should be. If technology is (close to) linear, it will determine relative prices, while preferences will determine quantities. Does this argument work for stocks?

Again, there is a standard model that has served well to describe quantities in growth, macroeconomics, and international economics. The standard model consists of a production function by which output, Y, is made from capital, K, and labor, L, perhaps with some uncertainty, θ, together with an accumulation equation by which investment I turns into new capital in the future. In equations, together with the most common functional forms,

$$\text{15)}\quad \begin{aligned} Y_t &= f(K_t, L_t, \theta_t) = \theta_t K_t^{\alpha} L_t^{1-\alpha} \\ K_{t+1} &= (1-\delta)K_t + I_t \\ Y_t &= C_t + I_t. \end{aligned}$$

It was well known already in the 1970s that this standard, "neoclassical" model would be a disaster at describing asset pricing facts. It predicts that stock prices and returns should be extremely stable. To see this, invest an extra dollar, reap the extra output that the additional capital will produce, and then invest a bit less next year. This action gives a physical or investment return. For the technology described in equation 15, the investment return is

$$16)\quad R^I_{t+1} = 1 + f_k(K_{t+1}, L_{t+1}, \theta_{t+1}) - \delta$$
$$= 1 + \alpha \frac{Y_{t+1}}{K_{t+1}} - \delta.$$

With the share of capital $\alpha \approx 1/3$, an output-capital ratio $Y/K \approx 1/3$, and depreciation $\delta \approx 10\%$, we have $R^I \approx 6\%$, so average equity returns are easily within the range of plausible parameters. The trouble lies with the variance. Capital is quite smooth, so even if output varies 3 percent in a year, the investment return only varies by 1 percent, far below the 17 percent standard deviation of stock returns. The basic problem is the absence of price variation. The capital accumulation equation shows that installed capital, K_t, and uninstalled capital, I_t, are perfect substitutes in making new capital, K_{t+1} Therefore, they must have the same price—the price of stocks relative to consumption goods must be exactly 1.0.

The obvious modification is that there must be some difference between installed and uninstalled capital. The most natural extra ingredient is an adjustment cost or irreversibility: It is hard to get any work done on the day the furniture is delivered, and it is hard to take paint back off the walls and sell it. To recognize these sensible features of investment, we can reduce output during periods of high investment or make negative investment costly by modifying equation 15 to

$$17)\quad Y_t = f(K_t L_t, \theta_t) - c(I_t, \cdot).$$

The dot reminds us that other variables may influence the adjustment or irreversibility cost term. A common specification is

$$Y_t = \theta_t K_t^\alpha L_t^{1-\alpha} - \frac{a}{2}\frac{I_t^2}{K_t}.$$

Now, there is a difference between installed and uninstalled capital, and the price of installed capital can vary. Adding an extra unit of capital tomorrow via extra investment costs $1 – \partial c(\cdot) \partial I$ units of output today, while an extra unit of capital would give (1–δ) units of capital tomorrow. Hence the price of capital in terms of output is

$$18)\quad P_t = \frac{1-\delta}{1 - \partial c / \partial i} \approx 1 + \frac{\partial c}{\partial I} - \delta$$
$$= 1 + a\frac{I_t}{K_t} - \delta,$$

where the last equality uses the quadratic functional form. (This is the q theory of investment. With an asymmetric *c* function, this is the basis of the theory of irreversible investment. Abel and Eberly [1996] give a recent synthesis with references.)

Equation 18 shows that stock prices are expected to be high when investment is high, or firms are expected to issue stock and invest when stock prices are high. The investment return is now

$$19)\quad R^I_{t+1} = (1-\delta)\frac{1 + f_k(t+1) - c_k(t+1) + c_i(t+1)}{1 + c_i(t)}$$
$$= (1-\delta)\frac{1 + \alpha\frac{Y_{t+1}}{K_{t+1}} - \frac{a}{2}\frac{I^2_{t+1}}{K^2_{t+1}} + a\frac{I_{t+1}}{K_{t+1}}}{1 + a\frac{i_t}{K_t}}$$
$$\approx 1 + \alpha\frac{Y_{t+1}}{K_{t+1}} - \delta + a\left(\frac{I_{t+1}}{K_{t+1}} - \frac{I_t}{K_t}\right).$$

Comparing equation 19 with equation 16, the investment return contains a new term proportional to the change in investment. Since investment is quite volatile, this model can be consistent with the volatility of the market return. In equation 18, the last term adds price changes to the model of the investment return.

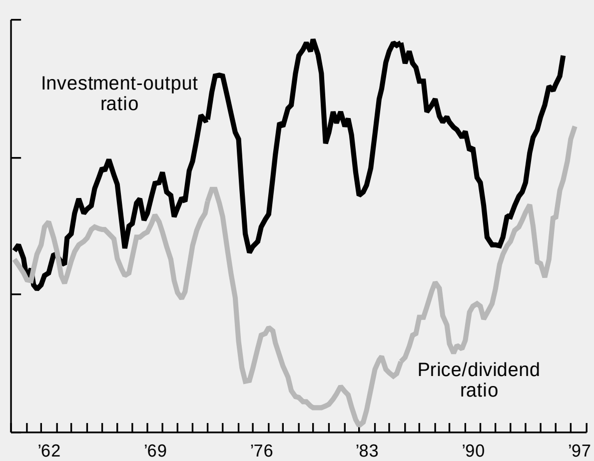
How does all this work? Figure 7 presents the investment-output ratio along with the value weighted P/D ratio. (The results are almost identical using an investment-capital ratio with capital formed from depreciated past investment.) Equation 18 suggests that these two series should move together. The cyclical movements in investment and stock prices do line up pretty well. The longer-term variation in P/D is not mirrored in investment: This simple model does not explain why investment stayed robust in the late 1970s despite dismal stock prices. However, the recent surge in the market is matched by a surge in investment.

This kind of model has been subject to an enormous formal empirical effort, which pretty much confirms the figure. First, the model is consistent with a good deal of the cyclical variation in investment and stock returns, both forecasts and *ex-post*. (See, for example, Cochrane, 1991c.) It does not do well with longer-term trends in the P/D ratio. Second, early tests relating investment to interest rates that imposed a constant risk premium did not work (Abel, 1983). The model only works at all if one recognizes that most variation in the cost of capital comes from time-varying expected stock returns with relatively constant interest rates. Third, the model in equation 18 taken literally allows no residual. If prices deviate one iota from the right-hand side of equation 18, then the model is statistically rejected—we can say with perfect certainty that it is not a literal description of the data-generating mechanism. There is a residual in actual data of course, and the residual can be correlated with other variables such as cash flow that suggest the presence of financial frictions (Fazzari, Hubbard, and Peterson, 1988). Finally, the size of the adjustment cost, a, is the subject of the same kind of controversy that surrounds the size of the risk aversion coefficient, γ. From equation 19 and the fact that investment growth has standard deviation of about 10 percent, $a \approx 2$ is needed to rationalize the roughly 20 percent standard deviation of stock returns. With $I/Y \approx 15\%$ and $Y/K \approx 33\%$, and hence $I/K \approx 1/20$, a value $a \approx 2$ means that adjustment costs relative to output are $\frac{a}{2}\frac{I}{K}\frac{I}{Y} = \frac{2}{2}\left(\frac{1}{20}\right) \text{ x } 15\% = 0.75\%$, which does not seem unreasonable. However, estimates of a based on regressions, Euler equations, or other techniques often result in much higher values, implying that implausibly large fractions of output are lost to adjustment costs.

This model does not yet satisfy the goal of determining the equity premium by technological considerations alone. Current specifications of technology allow firms to transform resources over time but not across states of nature. If the firm's own stock is undervalued, it can issue more and invest. However, if the interest rate is low, there is not much one can say about what the firm should do without thinking about the price of residual risk, and hence a preference approach to the equity premium. Technically, the marginal rate of transformation across states of nature is undefined.

FIGURE 7

Value weighted portfolio P/D and investment



Notes: Investment-output ratio and P/D of value weighted NYSE. Investment = gross fixed investment, output = gross domestic output. Series are stretched to fit on the same graph.

Implications of the recent surge in investment and stock prices

The association between stock returns and investment in figure 7 verifies that at least one connection between stock returns and the real economy works in some respects as it should. This argues against the view that stock market swings are due entirely to waves of irrational optimism and pessimism. It also verifies that the flow of money into the stock market does at least partially correspond to new real assets and not just price increases on existing assets.

In particular, the surge in stock prices since 1990 has been accompanied by a surge in investment. If expected stock returns and the cost of capital are low, then investment should be high. Statistically, high investment-output or investment-capital ratios also forecast low stock returns (Cochrane, 1991c). Thus, high investment provides additional statistical and economic evidence for the view that expected stock returns are quite low.

General equilibrium

To really understand the equity premium, we need to combine the utility function and production function modifications to construct complete, explicit economic models that replicate the asset pricing facts. This effort should also preserve, if not enhance, our understanding of the broad range of dynamic microeconomic, macroeconomic, international, and growth facts underpinning the standard models. The task is challenging. Anything that affects the relationship between consumption and asset prices will affect the relationship between consumption and investment. Asset prices mediate the consumption-investment decision and that decision lies at the heart of any dynamic macroeconomic model. We have learned a bit about how to go about this task, but have developed no completely satisfactory model as yet.

Jermann (1997) tried putting habit persistence consumers in a model with a neoclassical technology like equation 15, which is almost completely standard in business-cycle models. The easy opportunities for intertemporal transformation provided by that technology meant that the consumers used it to dramatically smooth consumption, destroying the prediction of a high equity premium. To generate the equity premium, Jermann added an adjustment cost technology like equation 17, as the production-side literature had proposed. This modification resulted in a high equity premium, but also large variation in risk-free rates.

Boldrin, Christiano, and Fisher (1995) also add habit-persistence preferences to real business cycle models with frictions in the allocation of resources to two sectors. They generate about one-half of the historical Sharpe ratio. They find some quantity dynamics are improved over the standard model. However, their model still predicts highly volatile interest rates and persistent consumption growth.

To avoid the implications of highly volatile interest rates, I suspect we will need representations of technology that allow easy transformation across time but not across states of nature, analogous to the need for easy intertemporal substitution but high risk aversion in preferences. Alternatively, the Campbell-Cochrane model above already produces the equity premium with constant interest rates, which can be interpreted as a linear production function. Models with this kind of precautionary savings motive may not be as severely affected by the addition of an explicit production technology.

Hansen, Sargent, and Tallarini (1997) use non-state-separable preferences similar to those of Epstein and Zin in a general equilibrium model. They show that a model with standard preferences and a model with non-state-separable preferences can predict the same path of quantity variables (such as output, investment, and consumption) but differ dramatically on asset prices. This example offers some explanation of how the real business cycle and growth literature could spend 25 years examining quantity data in detail and miss all the modifications to preferences that we seem to need to explain asset pricing data. It also means that asset price information is crucial to identifying preferences and calculating welfare costs of policy experiments. Finally, it offers hope that adding the deep modifications necessary to explain asset pricing phenomena will not demolish the success of standard models at describing the movements of quantities.

Implications

The standard economic models, which have been used with great success to describe growth, macroeconomics, international economics, and even dynamic microeconomics, do not predict the historical equity premium, let alone the predictability of returns. After ten years of effort, a range of deep modifications to the standard models show promise in explaining the equity premium as a combination of high risk aversion and new risk factors. Those modifications are now also consistent with the broad facts about consumption, interest rates, and predictable returns. However, the modifications have so far only been aimed at explaining asset pricing data. We have not yet established whether the deep modifications necessary to explain asset market data retain the models' previous successes at describing quantity data.

The modified models are drastic revisions to the macroeconomic tradition. In the Campbell-Cochrane model, for example, strong time-varying precautionary savings motives balance strong time-varying intertemporal substitution motives. Uncertainty is of first-order importance in this model; linearizations near the steady state and dynamics with the shocks turned off give dramatically wrong predictions about the model's behavior. The costs of business cycles are orders of magnitude larger than in standard models. In the Constantinides-Duffie model, one has to explicitly keep track of micro-economic heterogeneity in order to say anything about aggregates.

The new models are also a drastic revision to finance. We are used to thinking of aversion to wealth risk, as in the CAPM, as a good starting place or first-order approximation. But this view cannot hold. To justify the equity premium, people must be primarily averse to holding stocks because of their exposure to some other state variable or risk factor, such as recessions or changes in the investment opportunity set. To believe in the equity premium, one has to believe that these stories are sensible.

Finally, every quantitatively successful current story for the equity premium still requires astonishingly high risk aversion. The alternative, of course, is that the long-run equity premium is much smaller than the average postwar 8 percent excess return. The standard model was right after all, and historically high U.S. stock returns were largely due to luck or some other transient phenomenon.

Faced with the great difficulty economic theory still has in digesting the equity premium, I think the wise observer shades down the estimate of the future equity premium even more than suggested by the statistical uncertainty documented above.

Portfolio implications

In sum, the long-term average stock return may well be lower than the postwar 8 percent average over bonds, and currently high prices are a likely signal of unusually low expected returns. It is tempting to take a sell recommendation from this conclusion. There is one very important caution to such a recommendation. *On average, everyone has to hold the market portfolio*. The average person does not change his or her portfolio at all. For every individual who keeps money out of stocks, someone else must have a very long position in stocks. Prices adjust until this is the case. Thus, one should only hold less stocks than the average person if one is different from everyone else in some crucial way. It is not enough to be bearish, one must be more bearish than everyone else.

In the economic models that generate the equity premium, every investor is exactly happy to hold his or her share of the market portfolio, no more and no less. The point of the models is that the superficial attractiveness of stocks is balanced by a well-described source of risk, so that people are just willing to hold them. Similarly, the time variation in the equity premium does not necessarily mean one should attempt to time the market, buying more stocks at times of high expected returns and vice versa. Every investor in the Campbell-Cochrane model, for example, holds exactly the same portfolio all the time, while buy and sell signals come and go. In the peak of a boom they are not feeling very risk averse, and put their money in the market despite its low expected returns. In the bottom of a bust, they feel very risk averse, but the high expected returns are just enough to keep their money in the market.

To rationalize active portfolio strategies, such as pulling out of the market at times of high price ratios, you have to ask, who is there who is going to be more in the market than average now? And, what else are you going to do with the money?

More formally, it is easy to crank out portfolio advice, solutions to optimal portfolio problems given objectives like the utility function in equation 3. Assuming low risk aversion, and no labor income or other reason for time-varying risk exposure or risk aversion, solutions typically suggest large portfolio shares in equities and a strong market timing approach, sometimes highly leveraged and sometimes (now) even short. (See Barberis, 1997; Brandt, 1997.) If everyone followed this advice, however, the equity premium and the predictable variation in expected returns would disappear. Everyone trying to buy stocks would simply drive up the prices; everyone trying to time the market would stabilize prices. Thus, the majority of investors must be solving a different problem, deciding on their portfolios with different considerations in mind, so that they are always just willing to hold the outstanding stocks and bonds at current prices. Before going against this crowd, it is wise to understand why the crowd seems headed in a different direction.

Here, a good macroeconomic model of stock market risk could be extremely useful. The models describe why average consumers are so afraid of stocks and why that fear changes over time. Then, individuals in circumstances that make them different from everyone else can understand why they should behave differently from the crowd. If you have no habits or are immune to labor income shocks, in other words if you are unaffected by the state variables or risk factors that drive the stock market premium, by all means go your own way. However, the current state of the art is not advanced enough to provide concrete advice along these lines.

The last possibility is that one thinks one is smarter than everyone else, and that the equity premium and predictability are just patterns that are ignored by other people. This is a dangerous stance to take. *Someone* must be wrong in the view that he or she is smarter than everyone else. Furthermore, this view also suggests that the opportunities are not likely to last. People do learn. The opinions in this article are hardly a secret. We could interpret the recent run-up in the market as the result of people finally figuring out how good an investment stocks have been for the last century, and building institutions that allow wide participation in the stock market. If so, future returns are likely to be much lower, but there is not much one can do about it but sigh and join the parade.

APPENDIX: DERIVATIONS

Variance decomposition

Massaging an identity,

$$1 = R_{t+1}^{-1} R_{t+1}$$

$$1 = R_{t+1}^{-1} \frac{P_{t+1} + D_{t+1}}{P_t}$$

$$\frac{P_t}{D_t} = R_{t+1}^{-1} \frac{P_{t+1} + D_{t+1}}{D_t}$$

$$\frac{P_t}{D_t} = R_{t+1}^{-1} \left(\frac{P_{t+1}}{D_{t+1}} + 1 \right) \frac{D_{t+1}}{D_t}$$

$$20) \quad \frac{P_t}{D_t} = \sum_{j=1}^{\infty} \prod_{k=1}^{j} R_{t+k}^{-1} \frac{D_{t+k}}{D_{t+k-1}} + \lim_{j \to \infty} \left(\prod_{k=1}^{j} R_{t+k}^{-1} \right) \frac{P_{t+j}}{D_{t+j}}.$$

This equation shows how price-dividend ratios are exactly linked to subsequent returns, dividend growth, or a potential bubble. It is convenient to approximate this relation. We can follow Cochrane (1991b) and take a Taylor expansion now, or follow Campbell and Shiller (1986) and Taylor and expand the first equation in 20 to

$$p_t - d_t = \Delta d_{t+1} - r_{t+1} + \rho(p_{t+1} - d_{t+1})$$

and then iterate to

$$p_t - d_t = \sum_{j=1}^{\infty} \rho^j (\Delta d_{t+j} - r_{t+j}) + \lim_{j \to \infty} \rho^j (p_{t+j} - d_{t+j}).$$

Consumption-portfolio equations

I develop the consumption-portfolio problem in continuous time. This leads to a number of simplifications that can also be derived as approximations or specializations to the normal distribution in discrete time. A security has price P, dividend Ddt and thus instantaneous rate of return $dP/P + D/Pdt$. The utility function is

$$E_t \int e^{-\rho s} u(C_{t+s}) ds.$$

The first-order condition for an optimal consumption-portfolio choice is

$$u'(C_t)P_t = E_t \int e^{-\rho s} u'(C_{t+s}) D_{t+s} ds + E_t\left[e^{-\rho k} u'(C_{t+k}) P_{t+k} \right].$$

Letting the time interval shrink to zero, we have

$$0 = E_t[d(\Lambda P)] + D_t dt$$

where

$$\Lambda_t \equiv e^{-\rho t} u'(C_t).$$

Expanding the second moment, and dividing by ΛP

$$0 = E_t\left(\frac{dP}{P}\right) + \frac{D}{P} dt + E_t\left(\frac{d\Lambda}{\Lambda}\right) + E_t\left[\frac{d\Lambda}{\Lambda} \frac{dP_t}{P_t}\right].$$

Applying this basic condition to a risk-free asset,

$$r_t^f dt = -E_t\left[\frac{d\Lambda}{\Lambda}\right] = \rho dt - E_t\left[\frac{du'(C)}{u'(C)}\right] =$$
$$\rho dt - \frac{Cu''(C)}{u'(C)} E_t\left[\frac{dC}{C}\right]$$
$$r_t^f dt = \rho dt + \gamma E_t\left[\frac{dC}{C}\right].$$

This establishes equation 4. For any other asset,

$$0 = E_t\left(\frac{dP}{P}\right) + \frac{D}{P}dt - r_t^f dt = -E_t\left[\frac{d\Lambda}{\Lambda}\frac{dP_t}{P_t}\right].$$

Using Ito's lemma on Λ, we have

$$E_t\left[\frac{d\Lambda}{\Lambda}\frac{dP_t}{P_t}\right] = \frac{Cu''(C)}{u'(C)} E_t\left(\frac{dC}{C}\frac{dP}{P}\right).$$

Finally, using the symbols

$$r = \frac{dP}{P} + \frac{D}{P}dt, r^f = r^f dt,$$
$$\gamma = \frac{-Cu''(C)}{u'(C)}, \Delta c = \frac{dC}{C}$$

we have equation 5,

$$E_t(r) - r^f = -\gamma \operatorname{cov}_t[\Delta c, r] =$$
$$-\gamma \sigma_t(\Delta c)\sigma_t(r)\rho_t(\Delta c, r).$$

I drop the t subscript in the text where it is not important to keep track of the difference between conditional and unconditional moments.

Risk aversion calculations

What is the amount x that a consumer is willing to pay every period to avoid a bet that either increases consumption by y every period or decreases it by the same amount? The answer is found from the condition

$$\sum_j \delta^j u(C-x) = \frac{1}{2}\sum_j \delta^j u(C+y) +$$
$$\frac{1}{2}\sum_j \delta^j u(C-y).$$

Using the power functional form,

$$(C-x)^{1-\gamma} = \frac{1}{2}(C+y)^{1-\gamma} + \frac{1}{2}(C-y)^{1-\gamma},$$

and solving for x,

$$x = C - \left[\frac{1}{2}(C+y)^{1-\gamma} + \frac{1}{2}(C-y)^{1-\gamma}\right]^{\frac{1}{1-\gamma}}.$$

This equation is easier to solve in ratio form; the fraction of consumption that the family would pay is related to the fractional wealth risk by

$$\frac{x}{C} = 1 - \left[\frac{1}{2}\left(1+\frac{y}{C}\right)^{1-\gamma} + \frac{1}{2}\left(1-\frac{y}{C}\right)^{1-\gamma}\right]^{\frac{1}{1-\gamma}}.$$

This equation is the basis for the calculations in table 5.

For small risks, we can approximate

$$u(C-x) = \frac{1}{2}[u(C+y)+u(C-y)]$$
$$-u'(C)x \approx \frac{1}{2}u''(C)y^2$$
$$\frac{x}{C} \approx \frac{-Cu''(C)}{u'(C)}\left(\frac{y}{C}\right)^2 = \gamma\left(\frac{y}{C}\right)^2$$
$$\frac{x}{C} \approx \gamma\left(\frac{y}{C}\right)^2$$
$$\frac{x}{y} \approx \gamma\left(\frac{y}{C}\right).$$

NOTES

[1]More formally, we can only reject hypotheses that the true return is less than 3 percent or greater than 13 percent with 95 percent probability.

[2]A bit more formally, if you start with a regression of log returns on the P/D ratio, $r_{t+1} = a + b\ p/d_t + \varepsilon_{t+1}$, and a similar autoregression of the P/D ratio, $p/d_t = \mu + \rho\ \ p/d_{t-1} + \delta_t$, then you can calculate the implied long-horizon regression statistics. The fact that ρ is a very high number implies that long-horizon return regression coefficients and R^2 rise with the horizon, as in the table. See Hodrick (1992) and Cochrane (1991a) for calculations.

[3]The OLS regression estimate of ρ, is 0.90. However, this estimate is severely downward biased. In a Monte Carlo replication of the regression, a true coefficient $\rho = 0.90$ resulted in an estimate $\hat{\rho}$, with a mean of 0.82, a median of 0.83, and a standard deviation of 0.09. Assuming a true coefficient of 0.98 produces an OLS estimator $\hat{\rho}_{OLS}$, with median 0.90. I therefore adjust for the downward bias of the OLS estimate by using $\hat{\rho} = 0.98$.

[4]To generate a negative expected excess return, we have to believe that the market return is negatively conditionally correlated with the state variables that drive excess returns, for example consumption growth. This is theoretically possible, but seems awfully unlikely.

[5]Craine (1993) does a formal test of price/dividend stationarity and connects the test to bubbles. My statements are a superficial dismissal of a large literature. A lot of careful attention has been paid to the bubble possibility, but the current consensus seems to be that bubbles, as I have defined them here, do not explain price variation.

[6]Eliminate the last term, multiply both sides by $(p_t - d_t) - E(p_t - d_t)$ and take expectations.

[7]Several ways around this argument do exist. Kocherlakota (1990) defends a preference for later consumption. Kandel and Stambaugh (1991) note that the argument hinges critically on the definition of Δc. If we define Δc as the proportional change in consumption $\Delta c = (C_{t+\Delta t} - C_t)/C_t$ as I have (or, more properly, $\Delta c = dC/C$ in continuous time; see the appendix), then we obtain equation 4. However, if we define Δc as the change in log consumption, $\Delta c = \ln (C_{t+\Delta t}/C_t)$ or more properly $\Delta c = d(\ln C)$, we obtain an additional term, $r^f = \rho + \gamma E(\Delta c) - 1/2\gamma^2\sigma^2(\Delta c)$. For $\gamma < 100$ or so, the choice does not matter. The last term is small, since $E(\Delta c) \approx \sigma(\Delta c) \approx 0.01$. However, since γ is squared, the second term can be large with $\gamma = 250$, and can take the place of a negative ρ in generating a 1 percent interest rate with 1 percent consumption growth. What's going on? The model $u'(C) = C^{-250}$ is extraordinarily sensitive to the probability of consumption declines. The second model gives slightly higher weight to those probabilities. Rather than rescue the model with $\gamma = 250$, in my evaluation, this example shows how special it is: It says that interest rates as well as all asset prices depend only on the probabilities assigned to extremely rare events.

[8]I specify bets on annual consumption to sidestep the objection that most bets are bets on wealth rather than bets on consumption. As a first-order approximation, consumers will respond to lost wealth by lowering consumption at every date by the same amount. More sophisticated calculations yield the same qualitative results.

[9]The value function is formally defined as the achieved level of expected utility. It is a function of wealth because the richer you are, the happier you can get, if you spend your wealth wisely. The value can also be a function of other variables such as labor income or expected returns that describe the environment. Thus,

$$V(W_t,\cdot) = \max_{\{c_{t+s}\}} E_t \int_{s=0}^{\infty} e^{-\rho s} u(c_{t+s})\, ds \ \ s.t.(\text{constraints}).$$

The dot reminds us that there can be other arguments to the value function. $V_W = u_c$ is the "envelope" condition, and follows from this definition.

[10]Precisely, define the "surplus consumption ratio," $S = (C - X)/C$, and denote $s = \ln S$. Then s adapts to consumption following a discrete-time "square root process"

$$s_{t+1} - s_t = -(1-\phi)(s_t - \bar{s}) + \left[\frac{1}{\bar{S}}\sqrt{1 - 2(s - \bar{s})} - 1\right](c_{t+1} - c_t - g).$$

Taking a Taylor approximation, this specification is locally the same as allowing log habit x to adjust to consumption,

$$x_{t+1} \approx const. + \phi x_t + (1-\phi) c_{t+1}.$$

Campbell and Cochrane specify that habits are "external": Your neighbor's consumption raises your habit. This is a simplification, since it means that each consumption decision does not take into account its habit-forming effect. They argue that this assumption does not greatly affect aggregate consumption and asset price implications, though it is necessary to reconcile the unpredictability of individual consumption growth.

[11]Technically, this assertion depends on the form of the utility function. For example, with log utility, consumers don't care about future returns. In this statement I am assuming risk aversion greater than 1. See Campbell (1996).

[12]To prove this assertion, just substitute for C_{it} and take the expectation:

$$1 = E_{t-1} \exp\left[-\rho - \gamma \eta_{it} y_t + \frac{1}{2}\gamma y_t^2 + \ln R_{t+1}\right].$$

Since η is independent of everything else, we can use $E[f(\eta y)] = E[f(\eta y) \mid y)]$. Now, with η normal, E (exp $[-\gamma \eta_{it} y_t] \mid y_t) = \exp [\tfrac{1}{2}\gamma^2 y_t^2]$. Therefore, we have

$$1 = E_{t-1} \exp\left[-\rho + \frac{1}{2}\gamma^2 y_t^2 + \frac{1}{2}\gamma y_t^2 + \ln R_{t+1} \right]$$

$$1 = E_{t-1} \exp\left[-\rho + \frac{1}{2}\gamma(\gamma+1)\left(\frac{2}{\gamma(\gamma+1)}\right)\left(\ln\frac{1}{R_t} + \rho\right) + \ln R_{t+1} \right]$$

$$1 = E_{t-1} 1.$$

[13]There is a possibility that the square root term in equations 11 and 12 might be negative. Constantinides and Duffie rule out this possibility by assuming that the discount factor m satisfies

$$13)\quad \ln m_t \geq -\rho - \gamma \ln \frac{C_t}{C_{t-1}}$$

in every state of nature, so that the square root term is positive.

We can sometimes construct such discount factors by picking parameters a, b in $m_t = \max\,[a + bR_t, e^{-\rho}\,(C_t/C_{t-1})^{-\gamma}]$ to satisfy equation 14. However, neither this construction nor a discount factor satisfying equation 13 is guaranteed to exist for a given set of assets. The restriction in equation 13 is a tighter form of the familiar restriction that $m_t \geq 0$ is equivalent to the absence of arbitrage in the assets under consideration. Presumably, this restriction is what rules out markets for individual labor income risks in the model. The example $m = 1/R$ that I use is a positive discount factor that prices a single asset return $1 = E(R^{-1}R)$, but does not necessarily satisfy the restriction in equation 13. For high R, we can have very negative $\ln 1/R$. This is why the lines in figure 6 run into the horizontal axis at high R.

REFERENCES

Abel, Andrew B., "Optimal investment under uncertainty," *American Economic Review*, Vol. 73, No. 1, March 1983, pp. 228–233.

__________, "Asset prices under habit formation and catching up with the Joneses," *American Economic Review Papers and Proceedings*, Vol. 80, No. 2, May 1990, pp. 38–42.

Abel, Andrew B., and Janice C. Eberly, "Optimal investment with costly reversibility," *Review of Economic Studies*, Vol. 63, 1996, pp. 581–594.

Aiyagari, S. Rao, and Mark Gertler, "Asset returns with transactions costs and undiversifiable individual risk," *Journal of Monetary Economics*, Vol. 27, No. 3, June 1991, pp. 311–331.

Barberis, Nicholas, "Investing for the long run when returns are predictable," University of Chicago, Center for Research on Security Prices, working paper, No. 439, 1997.

Barsky, Robert B., Miles S. Kimball, F. Thomas Juster, and Matthew D. Shapiro, "Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement survey," National Bureau of Economic Research, working paper, No. 5213, 1997.

Bernardo, Antonio, and Olivier Ledoit, "Gain, loss and asset pricing," University of California at Los Angeles, Anderson Graduate School of Management, working paper, No. 16-96, 1996.

Black, Fischer, "Studies of stock price volatility changes," *Proceedings of the 1976 Meetings of the Business and Economic Statistics Section, American Statistical Association*, 1976, pp. 177–181.

Boldrin, Michele, Lawrence J. Christiano, and Jonas D.M. Fisher, "Asset pricing lessons for modeling business cycles," NBER, working paper, No. 5262, 1995.

Brandt, Michael, "Estimating portfolio and consumption choice: A conditional Euler equations approach," University of Chicago, Ph.D. dissertation, 1997.

Brown, Stephen J., William N. Goetzmann, and Stephen A. Ross, "Survival," *Journal of Finance*, Vol. 50, No. 3, 1995, pp. 853–873

Campbell, John Y., "Stock returns and the term structure," *Journal of Financial Economics*, Vol. 18, 1987, pp. 373–399.

__________, "Understanding risk and return," *Journal of Political Economy*, Vol. 104, April 1996, pp. 298–345.

Campbell, John Y., and John H. Cochrane, "By force of habit: A consumption-based explanation of aggregate stock market behavior," University of Chicago, working paper, 1997. [Revision of National Bureau of Economic Research, working paper, No. 4995, dated January 1, 1995.]

Campbell, John Y., and Robert J. Shiller, "The dividend–price ratio and expectations of future dividends and discount factors," *Review of Financial Studies*, Vol. 1, 1988, pp. 195–227.

Cochrane, John H., "Volatility tests and efficient markets: A review essay," *Journal of Monetary Economics*, Vol. 27, June 1991a, pp. 463–485.

__________, "Explaining the variance of price–dividend ratios," *Review of Financial Studies*, Vol. 15, 1991b, pp. 243–280.

__________, "Production-based asset pricing and the link between stock returns and economic fluctuations" *Journal of Finance*, Vol. 41, March 1991c, pp. 207–234.

Cochrane, John H., and Lars Peter Hansen, "Asset pricing lessons for macroeconomics," in *1992 NBER Macroeconomics Annual*, Olivier Blanchard and Stanley Fischer (eds.), 1992, pp. 1115–1165.

Constantinides, George M., "Theory of valuation: Overview," in *Theory of Valuation: Frontiers of Modern Financial Theory, Volume I*, Sudipto Bhattahcharya and George M. Constantinides (eds.), Totowa NJ: Rowman and Littlefield, 1989, pp. 1–23.

__________, "Habit formation: A resolution of the equity premium puzzle," *Journal of Political Economy*, Vol. 98, June 1990, pp. 519–543.

Constantinides, George M., and Darrell Duffie, "Asset pricing with heterogeneous consumers," *Journal of Political Economy*, Vol. 104, April 1996, pp. 219–240.

Craine, Roger, "Rational bubbles: A test," *Journal of Economic Dynamics and Control*, Vol. 17, 1993, pp. 829–846.

Deaton, Angus S., *Understanding Consumption*, New York: Oxford University Press, 1992.

Deaton, Angus, and Christina Paxson, "Intertemporal choice and inequality," *Journal of Political Economy*, Vol. 102, June 1994, pp. 437–467.

Duesenberry, James S., *Income, Saving, and the Theory of Consumer Behavior*, Cambridge, MA: Harvard University Press, 1949.

Eichenbaum, Martin, Lars Peter Hansen, and Kenneth J. Singleton, "A time series analysis of representative agent models of consumption and leisure choice under uncertainty," *Journal of Business and Economic Statistics*, Vol. 103, February 1988, pp. 51–78.

Epstein, Larry G., and Stanley E. Zin, "Substitution, risk aversion and the temporal behavior of asset returns," *Journal of Political Economy*, Vol. 99, 1989, pp. 263–286.

Fama, Eugene F., and Kenneth R. French, "Permanent and temporary components of stock prices," *Journal of Political Economy*, Vol. 96, April 1988a, pp. 246–273.

__________, "Dividend yields and expected stock returns," *Journal of Financial Economics*, Vol. 22, October 1988b, pp. 3–27.

__________, "Business conditions and expected returns on stocks and bonds," *Journal of Financial Economics*, Vol. 25, 1989, pp. 23–49.

__________, "Common risk factors in the returns on stocks and bonds," *Journal of Financial Economics*, Vol. 33, February 1993, pp. 3–56.

Fazzari, Steven M., Robert Glenn Hubbard, Robert Glenn, and Bruce C. Petersen, "Financing constraints and corporate investment," *Brookings Papers on Economic Activity*, Vol. 0, 1988, pp. 141–195.

Ferson, Wayne E., and George Constantinides, "Habit persistence and durability in aggregate consumption: Empirical tests," *Journal of Financial Economics*, Vol. 29, October 1991, pp. 199–240.

Friend, I., and M. Blume, "The demand for risky assets," *American Economic Review*, Vol. 55, 1975, pp. 900–922.

Hansen, Lars Peter, and Ravi Jagannathan, "Restrictions on intertemporal marginal rates of substitution implied by asset returns," *Journal of Political Economy*, Vol. 99, 1991, pp. 225–262.

Hansen, Lars Peter, Thomas J. Sargent, and Thomas D. Tallarini, "Robust permanent income and pricing" University of Chicago, manuscript, 1997.

Heaton, John C., "An empirical investigation of asset pricing with temporally dependent preference specifications," *Econometrica*, Vol. 63, May 1995, pp. 681–717.

Heaton, John, and Deborah J. Lucas, "Evaluating the effects of incomplete markets on risk sharing and asset pricing," *Journal of Political Economy*, Vol. 104, June 1996a, pp. 443–487.

__________, "Saving behavior and portfolio choice: Which risks matter?," Northwestern University, working paper, 1996b.

Hodrick, Robert, "Dividend yields and expected stock returns: Alternative procedures for inference and measurement," *Review of Financial Studies*, Vol. 5, No. 3, 1992, pp. 357–386.

Jagannathan, Ravi, and Z. Wang, "The conditional CAPM and the cross-section of expected returns," *Journal of Finance*, Vol. 51, March 1996, pp. 3–53. [Updated in *Staff Report*, No. 208, Federal Reserve Bank of Minneapolis, 1997.]

Jermann, Urban, "Asset pricing in production economies," *Journal of Monetary Economics*, 1997 forthcoming.

Kandel, Shmuel, and Robert F. Stambaugh, "Expectations and volatility of consumption and asset returns," *Review of Financial Studies*, Vol. 3, No. 2, 1990, pp. 207–232.

__________, "Asset returns and intertemporal preferences," *Journal of Monetary Economics*, Vol. 27, February 1991, pp. 39–71.

Kocherlakota, Narayanna R., "On the 'discount' factor in growth economies," *Journal of Monetary Economics*, Vol. 25, January 1990, pp. 43–47.

__________, "The equity premium: It's still a puzzle," *Journal of Economic Literature*, Vol. 34, March 1996, pp. 42–71.

Lamont, Owen, "Earnings and Expected Returns," National Bureau of Economic Research, working paper, No. 5671, 1996.

Lucas, Deborah J., "Asset pricing with undiversifiable income risk and short sales constraints: Deepening the equity premium puzzle," *Journal of Monetary Economics*, Vol. 34, December 1994, pp. 325–341.

Mankiw, N. Gregory, "The equity premium and the concentration of aggregate shocks," *Journal of Financial Economics*, Vol. 17, September 1986, pp. 211–219.

Mankiw, N. Gregory, and Stephen P. Zeldes, "The consumption of stockholders and nonstockholders," *Journal of Financial Economics*, Vol. 29, March 1991, pp. 97–112.

Marshall, David, and Kent Daniel, "The equity premium puzzle and the risk-free rate puzzle at long horizons," *Macroeconomic Dynamics*, Vol. 1, No. 1, 1997 forthcoming, pp. 452–484.

Mehra, Rajnish, and Edward Prescott, "The equity premium puzzle," *Journal of Monetary Economics*, Vol. 15, March 1985, pp. 145–161.

Merton, Robert C., "An intertemporal capital asset pricing model," *Econometrica*, Vol. 41, September 1973, pp. 867–887.

Poterba, James, and Lawrence J. Summers, "Mean reversion in stock prices," *Journal of Financial Economics*, Vol. 22, October 1988, pp. 27–59.

Reyfman, Alexander, "Labor market risk and expected asset returns," University of Chicago, Ph.D. dissertation, 1997, Ch. 3.

Rietz, Thomas A., "The equity risk premium: A solution," *Journal of Monetary Economics*, Vol. 22, July 1988, pp. 117–131.

Shiller, Robert J., "Consumption, asset markets, and macroeconomic fluctuations," *Carnegie Rochester Conference Series on Public Policy*, Vol. 17, Autumn 1982, pp. 203–238.

Telmer, Chris I., "Asset pricing puzzles and incomplete markets," *Journal of Finance*, Vol. 48, December 1993, pp. 1803–1832.

Weil, Philippe, "The equity premium puzzle and the risk-free rate puzzle," *Journal of Monetary Economics*, Vol. 24, November 1989, pp. 401–421.

Call for Papers

Payments Systems in the Global Economy: Risks and Opportunities

The Federal Reserve Bank of Chicago invites research and policy-oriented papers for submission to the 34th Annual Conference on Bank Structure and Competition to be held at the Westin Hotel in Chicago, Illinois, May 6–8, 1998.

The goal of the conference series is to generate an ongoing dialogue on policy issues affecting financial services. This year, the major theme of the conference will be developments in payments systems.

Technology is significantly altering the means available to initiate and settle transactions. Since payments technology and the structure of the financial services industry are changing rapidly, it is appropriate to evaluate how the payments system is evolving and what issues, including public policy issues, are raised by this evolution. For example, what is the current state of payments activity and the resulting levels of risk? What different problems are there for the paper-based, small-value electronic, and large-value electronic payments systems? Why do some industry participants argue for the expansion of real-time gross settlement arrangements, while others emphasize the efficiency aspects of netting arrangements? What is the evidence of systemic risk in the interbank market? Is there a need for regulatory oversight of this market? If so, are there moral hazard implications resulting from this involvement? Is systemic risk increasing with the increase in cross-border transactions? Is there a need for international coordination by supervisory agencies to prevent such problems? How can liquidity constraints best be addressed in the settlement of international transactions? Do the appropriate means to address these issues differ between developed and transition economies?

In the U.S., the Federal Reserve has recently undertaken a thorough analysis of the appropriate role the central bank should have in payments (the Rivlin Committee). Should it be exclusively a regulatory role? Is there a need for an operational role? What is the justification of any oversight role of private payments systems? What alternative arrangements exist internationally? Similarly, what are the causes of the significant differences in payment modes across countries? Why has the long-projected demise of the paper check in the U.S. not materialized?

There is also the question of the appropriate criteria for central bank provision of direct access to the large-dollar payments mechanism. Should access be limited to banks, expanded to other financial related firms, or expanded beyond the financial sector directly to commerce? Is this becoming a moot point as technological advances erode past protections given to banks?

Finally, there are issues related to the current and projected use of new payment vehicles, including smart cards, debit cards, stored-value cards, the Internet, and various electronic means to replace paper transactions. What are the preferred payment methods of the future and what are the related, if any, regulatory issues concerning security, efficiency, and monetary policy? Since the technology required to utilize these new payment vehicles is already in place, why have these developments not been more readily accepted and more frequently utilized? Why does the public not see the same benefits from these new developments that policy setters see?

The 1998 conference will focus on these and related questions. As in past years, the conference welcomes papers on other traditional issues regarding financial structure policy, including:

- safety net reform,
- systemic risk,
- risk management,
- industry consolidation,
- antitrust in banking,
- financial issues associated with credit access, community development, and fair lending issues,
- the unique role of commercial banks in intermediation,
- universal vs. narrow banks and alternative bank structures,
- issues related to international financial regulation, and
- the structure and regulation of the financial services industry.

Because of its success last year, as the last session of the conference we will again have featured industry experts provide an overview of the status of a particularly topical regulation. Presenters will discuss the purpose, structure, problems, and proposed changes associated with the regulation. A detailed discussion of the regulation will be available for inclusion in the *Proceedings* of the Conference. We will also hear from industry participants on how they perceive the regulation. Proposals for this session are also welcomed.

If you would like to present a paper at the conference, please submit four copies of the completed paper or an abstract with your name, address, and telephone number, and those of any coauthors, by December 19, 1997. Correspondence should be addressed to:

Conference on Bank Structure and Competition
Research Department
Federal Reserve Bank of Chicago
230 South LaSalle Street
Chicago, Illinois 60604-1413

For additional information, contact Douglas Evanoff at 312-322-5814 or devanoff@frbchi.org.

To order copies of any of these issues, or to receive a list of other publications, telephone 312-322-5111, fax 312-322-5515, or write to

Public Information Center
Federal Reserve Bank of Chicago
P.O. Box 834
Chicago, IL 60690-0834

ECONOMIC PERSPECTIVES and other Bank publications are available on the World Wide Web at http://www.frbchi.org.