WHERE IS THE MARKET GOING? UNCERTAIN FACTS AND NOVEL THEORIES

John H. Cochrane

Working Paper **6207**

NBER WORKING PAPER SERIES

WHERE IS THE MARKET GOING? UNCERTAIN FACTS AND NOVEL THEORIES

John H. Cochrane

Working Paper 6207
http://www.nber.org/papers/w6207

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 1997

My research is supported by a grant from the NSF administered by the NBER, and by the University of Chicago Graduate School of Business. I thank George Constantinides, Andrea Eisfeldt and David Marshall for many helpful comments. This paper is part of NBER's research programs in Asset Pricing and Economic Fluctuations and Growth. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

Where is the Market Going? Uncertain Facts
and Novel Theories
John H. Cochrane
NBER Working Paper No. 6207
October 1997
JEL Nos. G12, E3
Asset Pricing and Economic Fluctuations
and Growth

ABSTRACT

Will the stock market provide high returns in the future as it has in the past? The average US stock return in the postwar period has been about 8% above treasury bill rates. But that average is poorly measured: The standard confidence interval extends from 3% to 13%. Furthermore, expected returns are low at times such as the present of high prices. Therefore, the statistical evidence suggests a period of low average returns, followed by a slow reversion to a poorly measured long term average.

I turn to a detailed survey of economic theory, to see if models that summarize a vast amount of other information shed light on stock returns. Standard models predict nothing like the historical equity premium. After a decade of effort, a range of drastic modifications to the standard model can account for the historical equity premium. It remains to be seen whether the drastic modifications and a high equity premium, or the standard model and a low equity premium, will triumph in the end. Therefore, economic theory gives one reason to fear that average excess returns will not return to 8% after the period of low returns signaled by today's high prices.

I conclude with a warning that low average returns does not imply one should change one's portfolio. Someone has to hold the market portfolio; one should only deviate from that norm if one is different from everyone else.

John H. Cochrane
Graduate School of Business
University of Chicago
1101 East 58th Street
Chicago, IL 60637
and NBER
john.cochrane@gsb.uchicago.edu

1 Issues

Over the last century, the stock market in the United States has yielded impressive returns to its investors. Will stocks continue to give impressive returns in the future? Are the high stock returns in the last century some sort of law of nature, a fundamental feature of advanced industrial economies? Or are they the opposite of the old joke on Soviet agriculture – 100 years of good luck? Instead, perhaps they are the result of features of the economy which have or will soon disappear as financial markets become less regulated, "globalized", securitized, integrated, or subject to some other buzzword?

The question is especially pressing now, given the recent rise in the market. (As I write, the Dow has just broken 8000). Do high prices now mean lower returns in the future, or have stocks finally achieved Irving Fisher's brilliantly mistimed 1929 prediction that stocks had achieved a "permanently high plateau". If stocks have reached a plateau, is it a rising plateau, or will stocks bounce around the 8000 level for many years, not crashing, but not yielding much above bonds to their investors?

This issue is on all of our minds as we allocate our pension plan monies. It is also an important public policy question. 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 will continue to provide returns in the future as it has in the past.

There are lots of prognosticators on these issues. In this article I'll summarize the academic, and if I dare say so, scientific, evidence on these issues.

2 Facts

2.1 Average returns and risk

The most obvious approach to these questions is of course statistical. What is the evidence on past stock and bond returns?

2.1.1 Average 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% after inflation. The S&P500 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 EW portfolio has earned more than 11%. Bonds by contrast seem a disaster. Long term government bonds earned only 1.7% after inflation despite a standard deviation (11%) more than half that of stocks (17%). Corporate bonds earn a slight premium over government bonds, but at 2.1% are still a tragedy compared to stocks. Treasury bills earn even less, 0.8% on average after inflation.

	VW	S&P	EW	GB	CB	TB
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

Table 1. Annual real returns on stocks and bonds 1947-1996. VW = value weighted NYSE, S&P = S&P500, EW = equally weighted NYSE, GB = 10 year government bond, CB = corporate bond, TB = 3 month treasury bills. All less CPI inflation. All data are from the Center for Research on Security Prices (CRSP) at the University of Chicago.

2.1.2 A reward for risk

Table 1 highlights a crucially important fact. *The high returns are only earned as a compensation for risk.* We cannot understand high stock returns merely as high "productivity of the American economy" (in economics language, high marginal productivity of capital), or impatience by consumers. Such high productivity or impatience would lead to high returns on bonds as well as stocks.

To understand average stock returns, and to think about whether they will continue, we have to understand *not* why the economy gives such high returns to saving – it doesn't – but why it gives such a high compensation for *bearing risk.* And the risk is substantial. A 17% 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% probability that the decline will be bigger than -8% or rise bigger than 26%.)

Horizon h (Years)	$\frac{E(R^e)}{h}$	$\frac{\sigma(R^e)}{\sqrt{h}}$	$\frac{E(R^e)}{\sqrt{h}\sigma(R^e)}$
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

Table 2. Statistics on long horizon excess return for value-weighted NYSE return. R^e = value weighted return less T-bill rate. All statistics in percent.

It is common fallacy to dismiss this risk as "short-run price fluctuation", and to argue that stock market risk becomes less and less in the long run. Table 2 addresses this fallacy.

The most common fallacy is to confuse the *annualized* or average return with the actual return. For example, the two year log or continuously compounded returns is just 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, just like coin flips, we would find the mean and variance scale the same way with horizon, $E\left(r_{0\to 2}\right) = 2E\left(r_{0\to 1}\right)$; and $\sigma^2\left(r_{0\to 2}\right) = 2\sigma^2\left(r_{0\to 1}\right)$. An investor who cared about mean and variance would invest the same fraction of his wealth in stocks for any return horizon. The variance of annualized returns does stabilize; $\sigma^2(\frac{1}{2}r_{0\to 2}) = \frac{1}{2}\sigma^2(r_{0\to 1})$, but you eat the total, not annualized return. Analogously, 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 1000 times, the average number of heads will almost certainly come out quite near 1/2. However, the risk of the bet is much larger: it only takes an average number of heads equal to 0.499 (i.e., 499/1000) to lose a dollar; if the average number of heads is 0.490, still very close to 0.5, you lose \$10. You care about dollars, not the fraction of heads; you care about total returns, not annualized rates.

Table 2 shows that mean returns and standard deviations scale with horizon just about as this independence argument suggests, out to 5 years. Returns are not in fact independent over time and estimates in Fama and French (1986), Poterba and Summers (1986) suggest that the variance grows a bit less slowly than the horizon for the first 5-10 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 restore the annualized return fallacy. Also bear in mind that long-horizon statistics are measured even less well than annual statistics; we have only 5 nonoverlapping 10 year samples in the postwar period for example.

The stock market is like a coin flip, but like a biased coin flip. Thus, even though mean and variance may grow at the same rate with horizon, the probablility 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 large 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?

2.1.3 Means vs. standard deviations; Sharpe ratio

Figure 1 presents means versus standard deviations. In addition to the portfolios listed in Table 1, I include 10 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 earns over 15% per year average real return, though at the cost of a huge standard deviation. Furthermore, one can form portfolios with very high 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 we start to consider economic models, it is easy to get them to produce higher mean returns (along with higher standard deviations) by considering claims to leveraged capital.

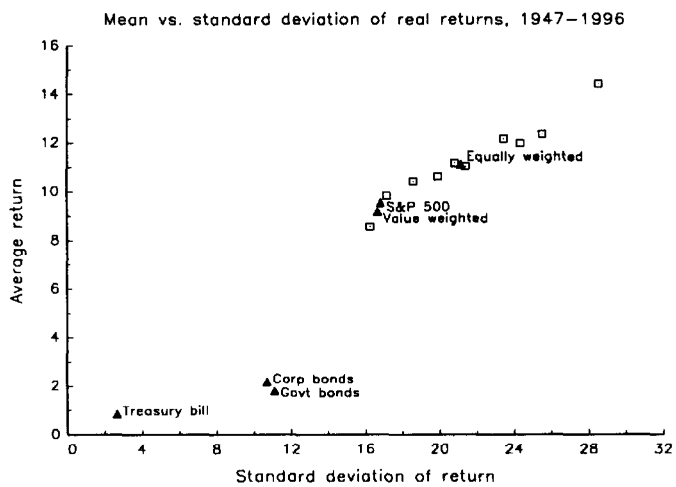

Figure 1: Mean vs. standard deviation of real returns, 1947-1996. TB = 3 month treasury bill, CB = corporate bonds, GB = 10 year government bonds, VW = NYSE value weighted stocks, EW = NYSE equally weighted stocks S&P = S&P500. Unmarked squares = NYSE size portfolios.

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 riskfree 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.

2.1.4 Standard errors

The average returns and Sharpe ratios look impressive. But are these true or just chance? One meaning of "chance" is this: suppose we look at T years of data from an economy that truly has an average return of $E(R)$. How likely is it that a 50 year sample has an average excess return of 8%? Equivalently, 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?

These questions are really unknowable at a deep level. Statistics provides an educated guess in the *standard error.* Assuming that each year's return is 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 quite deep: *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% 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% 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 $8-2\times2.4 = 3\%$ and $8+2\times2.4 = 13\%$ percent with 95% probability[1]. This is a wide band of uncertainty about the true market return, given 50 years of data! One can also see that 5 or 10 year averages are nearly useless; it takes a long time to statistically discern that the average return had 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.

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

Table 3. Standard errors and confidence interval widths, assuming returns are statistically independent with standard deviation $\sigma = 17\%$.

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

The standard errors are also the standard deviations of average returns over the next T years, and you can see that there is quite a lot of uncertainty about those returns, even if we accept the postwar sample as representative! For example, if we accept that the true mean excess return is and will continue to be 8%, the 5 year standard error of 7.6% means that there is still a good chance that the next 5 year return will average less than the treasury bill rate.

On the other hand, though we don't know the average return on stocks precisely, we do know something. The 2.4% standard error means that there is practically no chance that we see a 8% mean excess return given a true average return of zero, or even as high as 2-3%. The argument that *all* the past equity premium was luck faces very tough going against this simple statistical argument.

2.1.5 Selection and crashes

There are (at least) two subtle and important assumptions behind the above calculation however, and they do suggest ways in which the postwar average stock return might have been largely due to luck.

Argentina and the US looked very similar at the middle of the last century. Both were underdeveloped economies relative to the leaders, Britain and Germany, and had about the same per capita income. If Argentina had experienced the US's growth and stock returns, and vice versa, this article would probably be written in Spanish from the Buenos Aires Federal Reserve bank, and we would be puzzling over high Argentine stock returns.

The statistical danger is *selection* or *survival bias.* If you flip one coin 10 times, the chance of seeing 8 heads is low. But if you flip 10 coins 10 times, the chance that the coin with the greatest number of heads exceeds 8 heads is much larger. Does this story more closely capture the50 year return on US stocks? Brown, Goetzmann and Ross (1995) addresses this question quantitatively, and 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. We had no banking panics, and no depressions; no civil wars, no constitutional crises; we did not lose the cold war, no missiles were fired over Berlin, Cuba, Korea or Vietnam. If any of these things had happened, we would undoubtedly seen a calamitous decline in stock values. The statistical danger is *non-normality.* By taking the standard deviation from a sample that did not include rare calamities, and then calculating average return probabilities from a normal distribution, we are perhaps understating the true uncertainty. But investors, aware of that uncertainty, discount prices and hence leave high returns on the table.

The fundamental question we are trying to asses is this: was it clear to people in 1945 (or 1871, or whenever one starts the sample) and throughout the period that the average return on stocks would be 8% greater than that of bonds, subject to the 17% year to year variation? If so, then we face the challenge of explaining why people did not buy more stocks. But phrased this way, the answer is not so clear. Was it obvious in 1945 that the United States would not slip back into depression, but would instead experience a half century of growth never before seen in human history?

2.2 Time varying expected returns

2.2.1 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. Table 4 presents regressions of returns on the price-dividend ratio.

Horizon k	$R_{t\to t+k} = a + b(P_t/D_t)$			$D_{t+k}/D_t = a + b(P_t/D_t)$		
(years)	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

Table 4. OLS regressions of excess returns (value weighted NYSE - treasury bill rate) and dividend growth on VW price-dividend ratio. $R_{t\to t+x}$ indicates the x year return on the VW NYSE portfolio less the x year return from continuously reinvesting in treasury bills. Standard errors in parenthesis use GMM to correct for heteroskedasticity and serial correlation.

The regression at a one year horizon shows that excess returns are in fact predictable from price-dividend ratios. The 0.17 R^2 is not particularly remarkable. However, at longer and longer horizons larger and larger fractions of return variation are forecastable. At a 5 year horizon 60% of the variation in stock returns is forecastable ahead of time from the price/divided ratio! (Fama and French 1988 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 works about as well, including earnings, book value, moving averages of past prices. There seems to be an additional business-cycle component of expected return variance that is tracked by the term spread or similar business cycle forecasting variables including the default spread and investment/capital ratio, level of the T-bill rate, or the ratio of the T-bill rate to its moving average. (See Fama and French 1989 for term and default spreads, Cochrane 1991 for i/k, and the instruments in Ferson and

Constantinides 1991 for an 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 price-dividend ratio to keep the analysis simple.

In a similar fashion, *cross-sectional* variation in expected returns can be very well described by price-dividend ratio or (better) the ratio of market value to book value, which contains the price in its numerator. Portfolios of "undervalued" stocks with low price ratios outperform portfolios of "overvalued" stocks with high price ratios. (See Fama and French 1993 and references).

2.2.2 Slow moving P/D and P/E

Figure 2 presents the price-dividend and price earnings ratios. This graph emphasizes that price ratios are very *slow moving* variables. This is why they forecast *long-horizon* movements in stock returns.

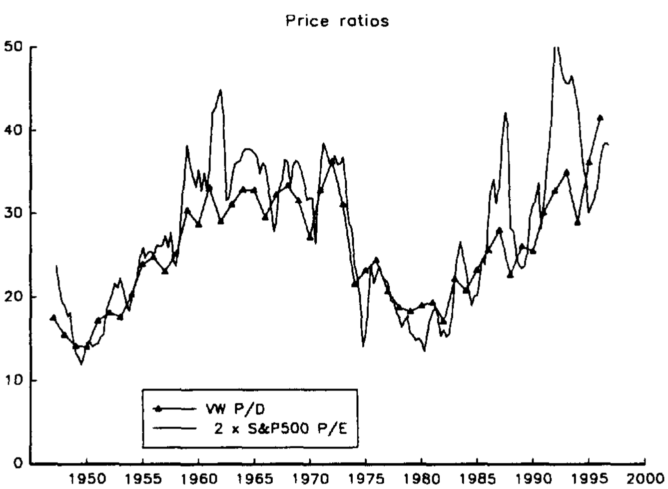

Figure 2: Price-dividend ratio of the value weighted NYSE portfolio and S&P500 price-earnings ratio. 2× the price-earnings ratio is plotted so that the lines can be more easily compared.

The rise in forecast power with horizon is not a separate phenomenon. It results from the one-period forecastability of returns and the slow movement in the price-dividend ratio[2]. As an analogy, if it is 10° below zero in Chicago, (low price-dividend ratio), one's

[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 means that long-horizon return regression coefficients and R^2 rise with the horizon, as in the table. See Hodrick (19xx) and Cochrane (1991a) for calculations.

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° in a day so the forecast is not very accurate. But the fact that it is 10° below signals that the temperature will rise on average one degree per day for many days. By the time we look at a 6 month horizon, we forecast a 90° rise in temperature. The daily variation of 20° is still there, but the forecastable change in temperature (90°) is much larger relative to the daily variation, implying a large R^2.

The slow movement in the price-dividend ratio also means that the forecastability of returns is not the fabled alchemists' stone that turns lead into gold. A high price-dividend ratio means that prices will grow more slowly than dividends for a long time until the price-dividend 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% for the same standard deviation. If we had a 50% R^2 at a daily horizon, we could make a lot of money; but not so at a 5 year horizon.

The slow movement of the price-dividend ratio also means that on a purely *statistical* basis, return forecastability is a very open question. What we really know, looking at Figure 2 (Figure 4 below also makes this point), is that low prices relative to dividends and earnings in the 50's preceded the boom market of the early 60's; that the high price-dividend ratios of the mid-60's preceded the poor returns of the 70's; that the low price ratios of the mid-70's preceded the current boom. And that price ratios are very high now. In any real sense, we really have 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% probability value before one starts to worry about fishing and selection biases.

2.2.3 What about repurchases? –P/E and other forecasts

Is the price-dividend ratio still a valid signal? Perhaps increasing dividend repurchases mean that the price-dividend ratio will not return to its historical low values; 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&P500 price-earnings ratio along with the price-dividend ratio. As you can see, the two measures line up well. The price-earnings ratio forecasts returns almost as well as the price-dividend ratio. The price-earnings 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 relation might be fundamentally different now than before.

2.2.4 Return forecasts

So, what do the regressions of Table 4 say, quantitatively, about future returns? Figure 3 presents one-year returns and the price-dividend ratio forecast. Figure 4 presents 5 year returns and the price-dividend ratio forecast. I include in-sample and out-of-sample forecasts in figures 3 and 4. To form the latter 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 runs through the end of 1996, the forecast returns for 1997 and 1998 are

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

and so on.[3]

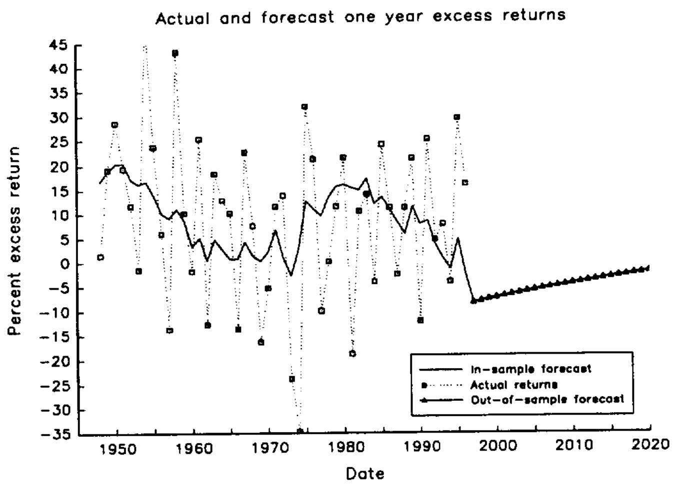

Figure 3: One year excess returns on the value weighted NYSE portfolio and forecast from a regression on the price/dividend ratio. Returns are plotted on the day of the foreacast; for example 1995 plots $a + b \times P/D_{1995}$ and the 1996 return. The out-of sample forecast is made by joining $R_{t+1} = a + bP/D_t$ with $P/D_{t+1} = \mu + \rho P/D_t$.

The forecast is extraordinarily pessimistic. It starts at a -8% excess return for 1997, and only very slowly returns to the estimated unconditional mean excess return of 8%. In 10 years, the forecast is still -5%, in 25 years it is -1.75%, and it is still only 2.35% in 50 years! The 5 year return forecasts are similarly pessimistic.

[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 estiamate by using $\hat{\rho} = 0.98$.

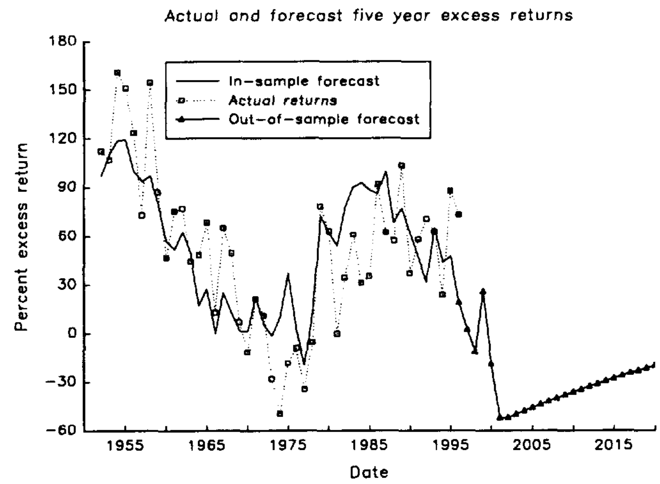

Figure 4: Five year excess returns on the value weighted NYSE portfolio (5 year returns - 5 year treasury bill returns) and forecast from a regression of 5 year returns on the price/dividend ratio. Returns are plotted on the last day; for example 1996 plots the forecast $a + b \times P/D_{1991}$ and the return from 1991 to 1996. The out of sample forecast is formed for 1997-2001 from the observed price/dividend ratios 1992-1996, and for 2002 on from $P/D_t = \mu + 0.98P/D_{t-1}$.

Of course, this forecast is subject to lots of uncertainty. There is uncertainty about what actual returns will be, given the forecasts. More subtly, there is also a great deal of uncertainty and doubt about the forecasts themselves. We can never hope to precisely forecast the direction of stock prices; if we could they would cease to be risky and cease to pay a premium for risk. The forecasts attempt to measure *expected returns*, the quantity that investors must trade off against unavoidable risk in deciding how attractive an investment is. We have to ask how accurately the forecasts measure expected returns.

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 overfitting is larger.

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 gives one comfort 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 price-dividend ratio has never in the postwar period been as high as it is now. Extending historical experience to never-before seen values is always a dangerous proposition. 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 price-dividend ratios above 30-35, 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.

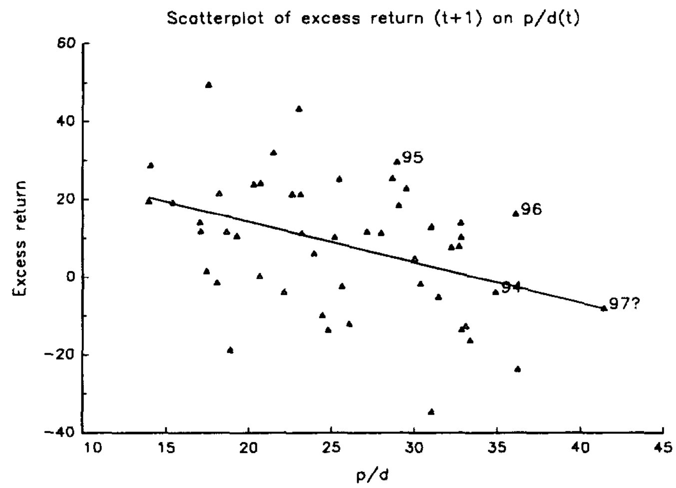

Figure 5: Scatterplot of one-year excess returns, and one year ahead forecast from the price-dividend ratio.

2.2.5 What about the last few years of high returns?

The price-dividend ratios also pointed to low returns in 1995 and 1996, and they were wrong. 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 fact mean that the price-dividend signal should no longer be trusted?

To answer this criticism, look at the figures again. They make clear that the returns of 95 and 96 and even another 20% or so return in 97 are not so far out of line, despite a pessimistic price-dividend forecast, that we should throw away the regression based on the previous 47 years of experience. To return to the analogy, if it is 10° below zero in

[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.

Chicago that means spring is coming. But we can easily have a few weeks of 20° below weather before spring finally arrives. The graphs make vivid how large a 17% standard deviation really is, and to what extent the forecasts based on the price-dividend 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 substantial uncertainty about the forecast is how persistent the price-dividend ratio really is. If, for example, the price-dividend 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% (7% 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 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 only lies between 0.83 and 0.94 50% of the time, and 0.66 and 1.00 95% 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 price-dividend ratio were really less persistent than it seems. In this case, the near-term return forecast would be unchanged, but the long-term forecast would return to 9% much more quickly.

2.3 Variance decomposition

When prices are high relative to dividends (or earnings, cashflow, 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 cashflows are discounted at a lower than usual rate, leading to higher prices. 3) Investors expect prices to keep rising forever, in a "bubble." This statement is not a theory, it is an accounting identity like 1=1: If the price-dividend ratio is high, either dividends must rise, prices must decline, or the price-dividend ratio must never return to its historical average. The open question is, which of the three options holds for our stock market? Are prices high now because investors expect future earnings, dividends etc. to rise, because they expect low returns in the future, or because they expect prices to go on rising forever?

The first place to stop in meditating on this question is history. And historically, we find that *virtually all variation in price-dividend ratios has reflected varying expected returns.*

At a simple level, Table 4 includes regressions of long-horizon dividend growth on price-dividend ratios to match the regressions of returns on price-dividend ratios. The coefficients in the dividend growth case are much smaller, typically one standard error from zero, and the R^2 are tiny. Worse, the *signs* are wrong. To the extent that a high price-dividend 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

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

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.) The identity (1) gives a precise meaning to my earlier statement that a high price-dividend 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 price-dividend ratio variation. Unless the price-dividend ratio grows faster than $1/\rho^j$, there is no bubble. It is hard to believe that price-dividend ratios can grow *forever*. Empirically, price-dividend 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]

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

In words, price-dividend ratios can *only* vary if they forecast changing dividend growth or of they forecast changing returns.

This is a powerful equation. At first glance, it would seem a reasonable approximation that returns are unforecastable (the "random walk" hypothesis) and that dividend growth is not forecastable either. But if this were the case, the price-dividend ratio would have to be a *constant.* Thus the fact that the price-dividend ratio varies *at all* means that either dividend growth or returns must be forecastable.

This observation solidifies one's belief in price-dividend ratio forecasts of returns. Yes, the statistical evidence that price-dividend ratios forecast returns is weak. But price-dividend ratios *have* varied. So the choice is not, price-dividend ratios forecast returns or they forecast nothing. The choice is, price-dividend ratios forecast returns *or* they forecast dividend growth. They have to forecast *something.* Given this choice and Table 4, it seems a much firmer conclusion that they forecast returns.

Having seen equation (2), one is hungry for estimates. Table 5 presents some, taken from Cochrane (1991b). As one might suspect from Table 4, Table 5 shows that in the past *almost all variation in price-dividend ratios is due to changing return forecasts.* (The rows of Table 5 do not add up to exactly 100% because equation (2) is an approximation. The elements do not have to be between 0 and 100%. For example, -34, 138 occurs because high prices seem to forecast lower dividend growth. Therefore they must and do forecast really low returns, and returns must account for more than 100% of price-dividend variation.)

[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 possiblity, but the current consensus does seem to be that bubbles do not explain price variation.

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

	Dividends	Returns
Real	-34	138
std. error	10	32
Nominal	30	85
std. error	41	19

Table 5. Variance decomposition of value-weighted NYSE price-dividend ratio. Table entries are the percent of the variance of the price-dividend ratio attributable to dividend and return forecasts, $100\times cov(p_t-d_t, \sum_{j=1}^{15}\rho^j\Delta d_{t+j})/var(p_t-d_t)$ and similarly for returns.

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

2.4 The Bottom line

Statistical analysis suggests that the long-term average return on broad stock market indices is 8% greater than the treasury bill rate, with a standard error of about 3%. High prices are related to low subsequent excess returns. Based on these patterns, the expected excess return (stock return less treasury bill rate) is near zero for the next 5 years or so, and then slowly rising to the historical average. The large standard deviation of excess returns, about 17%, means that actual returns will certainly deviate substantially from the expected return. Finally, one always gets more expected return by taking on more risk.

3 Economics: Understanding the equity premium

The statistical analysis I have presented so far should and does leave one very unsatisfied. Statistical analysis of past returns leaves a lot of uncertainty about future returns. Furthermore, we can't really believe that average excess returns are 8% until we understand *why* this is so: one should always distrust statistical associations with no known underlying mechanism. Perhaps most importantly, no statistical analysis can tell us if the future will be like the past. Even if the true expected excess return was 8%, did that fact result from fundamental, or from temporary features of the economy? We have to try for an *economic* understanding of stock returns to answer these questions.

Economic theory and modeling has an undeserved image 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. By an "economic model" or "theory" all I mean is the exercise of making such a story explicit, checking whether it is internally consistent, checking whether it can quantitatively explains stock returns, and then quantitatively checking that it does not make wildly counterfactual predictions in other dimensions, for example checking that to explain the stock return it does not require wild variation in riskfree 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 that 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?

By "makes sense," I mean this: Did people in 1947, and throughout the period, know that stocks were going to yield 8% over bonds on average, yet they were rationally unwilling to hold more stocks because they were afraid of the 17% standard deviation or some other measure of stocks' risk? If so then 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 expected returns over time. After 10 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 command 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 first 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 couterfactual predictions regarding other facts. The goal is to explain not only the equity premium, but to do so in a manner consistent with the level and volatility of consumption growth (both about 1% per year), the predictability of stock returns described above, the relative lack of predictability in consumption growth, the relative constancy of real riskfree rates over time and across countries, and the relatively low correlation of stock returns with consumption growth. This is a tough assignment, that 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. I will not try to review the entire literature and detail contributions along the way. Kocherlakota 1996 and Cochrane and Hansen 1992 are two literature reviews. 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 they 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.

3.1 The standard model

To say anything about dynamic economics, we have to say something about how people are willing to trade consumption at one moment and set of circumstances (state of nature) for consumption at 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% interest rate. It also might have quite volatile consumption. If people did not care what was the set of circumstances in which they get $1.10, then all expected returns will be equal to the interest rate (risk neutrality). However, 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% interest rate to get him to consume even less this year.

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

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

E denotes expectation. C_t denotes consumption at date t. ρ is the "subjective discount factor"; the term $e^{-\rho t}$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, $K_{t+1} = (1 - \delta)K_t + I_t$ 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 squabbles concern details of implementation; what exact ingredients are important for specific phenomena.

Since this basic economic framework explains such a wide range of phenomena, let's ask what it predicts for the equity risk premium. Give the opportunity to buy assets such as stocks and bonds to a consumer whose preferences are described by (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:

$$r^f = \rho + \gamma E\left(\Delta c\right) \tag{4}$$

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

where Δc denotes the proportional change in consumption, r denotes a risky asset return, r^f denotes the riskfree 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. It results from the fact that 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 CAPM (capital asset pricing model) will recognize the intuition; replacing wealth or the market portfolio with consumption gives the most modern and general version of that theory.

3.1.1 The equity premium puzzle

Now, we are ready to evaluate the equity premium. Transform equation (5) to

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

The left hand side is the Sharpe ratio. As I showed above, the Sharpe ratio is about 0.5 for the stock market, and is a good quantity on which to focus, because it is robust to leveraging or choice of assets.

The right hand side of (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. The right hand side of (6) is a prediction of what the Sharpe ratio should be. If this prediction is low, then the consumer should invest more in the asset with return r. Doing so will make his consumption stream more risky, and more correlated with the asset return. Thus, as he invests more, the right hand side of (6) will approach the left hand side.

The right hand side of (6) predicts nothing even close to the historical equity premium. The standard deviation of aggregate consumption growth is about 1% 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. Putting this all together, equation (6) shows that the standard model, with γ even as high as 10, doesn't produce anything like the historical Sharpe ratio. $10 \times 0.01 \times 0.2 = 0.02$ rather than 0.5; at a 20% standard deviation this ratio implies an average return for stocks of $0.02 \times 20 = 0.4\%$ (40 basis points) rather than 9%.

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 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.

3.1.2 Can we change the numbers?

The correlation of consumption growth and returns is the most suspicious ingredient in this calculation. While it 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. Daniel and Marshall (1997) find correlations in the data up to 0.4 at a 2 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 treasury bills "money-like" and artificially drive down the interest rate (for example Aiyagari and Gertler 1991). The highest Sharpe ratio occurs in fact when one considers short-term riskfree 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 fix 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.

3.1.3 Time-varying expected returns

The consumption based view with $u'(C) = C^{-\gamma}$ also has trouble reproducing the forecastability of stock returns documented above. Consider the conditional version of equation (6),

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

where $E_t, \sigma_t, corr_t$ represent conditional moments. We found that the price-dividend 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 price-dividend ratios decline – the "leverage effect" of Black (1976). However, all that evidence comes from three outliers, and 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? Look at (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. We could imagine that the conditional correlation of consumption growth and returns varies a great deal over time. But this seems unlikely, or more precisely like an unfathomable assumption on which to build our central understanding of time-varying returns.

Since the standard model seems not to explain the asset pricing facts, one natural response is to modify it, or think about other economic frameworks.

3.2 What about the CAPM?

Finance researchers and practitioners often express disbelief and indeed 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. The consumption-based model looks like a huge ivory-castle waste of time in comparison to the empirical fit of multifactor asset pricing models, to say nothing of the dramatic successes of options pricing models such as Black and Scholes (1972). Why not use the CAPM or other, better-performing finance models to understand the equity premium?

The answer is that the CAPM and 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 the return of stock i 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. The Black-Scholes model explains 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.

3.3 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. In order to generate a Sharpe ratio of 0.5, we need $\gamma = 250$ in equation (6). Even if we allow $corr = 1$, we still need $\gamma = 50$.

What's wrong with $\gamma = 50$ to 250? Why is the low numerical value of γ so entrenched with the standard model? The answer is, though a high curvature γ explains the equity premium, it runs quickly into trouble with other facts.

3.3.1 Consumption and interest rates

The most basic piece of evidence for low γ is the relation between consumption growth and interest rates. Real interest rates are quite stable over time (see Table 1), and roughly the same the world over, despite variation in consumption growth over time and across people and countries. $\gamma = 50$ 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 relation between consumption growth and interest rates, equation (4), reproduced here:

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

High values of γ give us trouble when just trying to understand the level of interest rates. Average consumption growth and real interest rates are both about 1% Thus, $\gamma = 50$ to 250 requires $\rho = -0.5$ to -2.5, or a -50% to -250% subjective discount factor. 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. But $\gamma = 50$ means that a country or a boom with consumption growth even 1% higher than normal must have real interest rates 50 percentage points higher than normal, and consumption 1% lower than normal should be accompanied by real interest rates of 50 percentage points lower than normal, which typically means huge negative interest rates – you pay them 48% to keep your money. We don't see anything like this.

We can also phrase this issue as a conceptual experiment, suitable for thinking about one's own preferences or for survey evidence on others' preferences. For example, ask

[7] Several ways around this argument do exist. Kocherlakota (199x) 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 (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, $r^f = \rho + \gamma E(\Delta c) - \frac{1}{2}\gamma^2\sigma^2(\Delta c)$. 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% interest rate with 1% consumption growth. What's going on? The model $u'(C) = C^{-250}$ is extroardinarily sensitive to the probability of consumption declines. The second model gives slightly higher weight to those probabilities. Rather than rescue the model, in my evaluation, this example shows how special it is: interest rates as well as all asset prices depend only on the probabilities assigned to extremely rare events.

what does it take to convince someone to skip a vacation. Take a family with $50,000 per year income, consumption equal to income, and which spends $2,500 (5%) 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%. I think most of us would give in and 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, output etc. and for delivering reasonable speeds of adjustment to shocks.

3.3.2 Risk aversion

Economists have also shied away from high curvature γ on the objection that people don't seem that averse to risks. I will conclude that we have much less solid reasons to object to high risk aversion than we do to object to high aversion to intertemporal substitution via the consumption - interest rate relations I examined above. To reach this conclusion, I outline and examine the case against high risk aversion.

Surveys and thought experiments

Since Sharpe ratios are high everywhere, much of the aversion to risk aversion comes from simple thought experiments rather than data. Let us ask, 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].

	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
$1000	20	99	435	665	863
$10000	2000	6921	9430	9718	9889

Table 6. Amount that a family with constant $50,000 per year consumption would pay per year to avoid an even bet of gaining or losing the indicated amount, also per year.

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

[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.

the lower right hand corner, the family would rather pay \$9,889 for sure than take a 1/2 chance of a \$10,000 loss! This prediction is surely unreasonable, and has lead most authors to rule out risk aversion coefficients over 10. Survey evidence for this kind of bet 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 5-25¢ to avoid a \$10 bet. More generally, 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, the amount 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. More deeply, we 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 thing one learns from conceptual experiments, and laboratory and survey evidence of simple bets is that people's answers to such questions routinely violate expected utility. This fact lowers the value of this source of evidence as a *measurement* of risk aversion. Barsky et al. also report that whether an individual partakes in a wide variety of risky activities correlates poorly with the level of risk aversion inferred a survey. In the end, surveys about what one would do with hypothetical bets, far from the range of everyday experience are hard to interpret.

Microeconomic evidence

One wishes instead for some microeconomic observations from which to infer risk aversion, evidence from people's actual behavior in their daily activities. For example, one could match numbers such as those in Table 6 with insurance data. People are willing to pay substantially above actuarially fair values to insure against car theft or house fires. What is the implied risk aversion? But even if we were to find other markets whose prices reflect less risk aversion than stocks, this leaves open the gaping question: if they are so risk-neutral in the survey taker or insurance broker's office, why do they seem to get so risk averse in the stock broker's office? Perhaps the carefully examined 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! A challenge I issue at conferences: If you think risk aversion is so low, why aren't *you* holding a highly leveraged position in the market?

Portfolio calculations

A common calibration of risk aversion comes from simple portfolio calculations. Friend and Blume (1975) is an oft-cited early source for this kind of calculation. 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, $V_W(W)$, we can also write (6) as

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

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 we observe are paid off in wealth. For a consumer who invests entirely in stocks, $\sigma(\Delta w)$ is the standard deviation of the *stock* return, and $corr(\Delta w, r) = 1$. Now to generate a Sharpe ratio of 0.5 we only need

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

What's wrong with this calculation? The Achilles heel 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 however, then consumption would move one-for-one with wealth, and $\sigma(\Delta c) = \sigma(\Delta w)$. If your wealth doubles and nothing else has changed, you would double consumption. This calculation in fact hides a "model", and that model has the drastically counterfactual implication that consumption growth has a 17% standard deviation!

The fact that consumption has a standard deviation so much lower than that of stock returns tells us that returns are *not* independent over time (as we already know from the return on p/d regressions), and/or that other state variables *must* be important in driving stock returns. But the minute we allow some other state variable z – 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 (6) leads not to (8), but to

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

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

For now, this observation just tells us that portfolio-based calibrations of risk aversion don't work, because they implicitly assume independent returns and hence consumption

[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,

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

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.

growth as volatile as returns. Below, I will start to think about plausible candidates for the variable z that can help us to understand high Sharpe ratios.

Risk aversion?

Overall, I conclude that the evidence against high *risk aversion* is not that strong, and it is at least a possibility we should consider. This observation does not rescue the power utility model with $\gamma = 50$ to 250 – that ship sank on the consumption-interest rate shoals. But we can at least contemplate other models with high risk aversion.

3.4 New utility functions and state variables

If changing the *parameter* γ in $u'(C) = C^{-\gamma}$ doesn't 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 one gets qualitatively the same results even in discrete time with other functional forms. A more promising avenue is to consider other arguments of the utility function, or *non-separabilities*.

Perhaps how you feel about an extra bit of consumption today is affected not just by how much you are already eating, but by other things: how much you ate yesterday, or how much you worked today. Then, the covariance of stock returns with these other variables will also determine the 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. Then, stocks may be 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, we get to our fundamental equation (6) from

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

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

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

Since the first covariance is a bust, we will have to rely heavily on the latter to explain the premium.

There is a practical aim to generalizing the utility function as well. As you saw in the last section, the 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) as well as their attitudes toward *risk*. $\gamma = 50$ 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. Perhaps a modification of preferences can help us to disentangle the two attitudes.

3.4.1 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 (3) sums over states of nature, e.g.

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

"Separability" means one adds 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

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

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's not clear a priori whether more leisure enhances the marginal utility of consumption (why bother buying a boat if you're at the office all day and can't use it) or vice versa (if you have to work all day, it's more important to come home to a really nice big TV). However, we can let the data speak on this matter. Explicit versions of this approach have not been very successful to date. (Eichenbaum, Hansen and Singleton 1989 for example). On the other hand, recent research has found that adding labor income 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 via a utility function, the facts in this research may in the future be interpretable as an effect of leisure on the marginal utility of consumption.

3.4.2 The force of habits

Nonseparabilities over *time* have been more useful in empirical work. Anyone who has had a large pizza dinner or smoked a cigarette knows that what you consumed yesterday can have an impact on how you feel about more consumption today. Might a similar mechanism apply for consumption in general and at a longer time horizon? Perhaps we get used to an accustomed 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 thought can at least explain 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 also 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

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

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

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

Here is the central idea. 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 $\eta = 2$) 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 γ_t 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 didn't we just learn that raising curvature implies high and time-varying interest rates? In the Campbell-Cochrane model the answer is, no. The reason

[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,

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

Campbell and Cochrane specify that habits are "external"; your neighbor's consumption raises your habit. This is a simplification, since it means 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.

is *precautionary saving*. Suppose we are in a bad time, in which consumption is low relative to habit. People want to borrow against future, higher, consumption, and this force should drive up interest rates. However, people are also much more risk averse when consumption is low. This consideration induces them to *save* more, in order to build up assets against the event that tomorrow might be even worse.

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 $\eta = 2$ controls the relation between consumption growth and interest rates, while the *curvature* coefficient $\gamma = \eta\frac{C}{C-X}$ controls the risk premium. Thus this habit model allows high "risk aversion" with low "aversion to intertemporal substitution", and it *is* consistent with the consumption and interest rate data.

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 return forecastability from the price-dividend ratio and 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 price-dividend ratio forecasts a higher Sharpe ratio. Artificial data from the model also replicates 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 price-dividend fluctuations over the last century. However, it does not account for the current high price-dividend ratio. The reason is simple: the model generates a high price-dividend 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 for 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.

3.4.3 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 above. 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 (1997) explore habit persistence models that can generate a large equity premium without large risk aversion, i.e. 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 reward.

Suppose a consumer wins a bet, or enjoys a high stock return. Normally, he would instantly raise consumption to match his 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, he increases consumption slowly and predictably after the increase in wealth. Similarly, he would borrow to slowly decrease consumption after a decline in wealth, trading an even lower eventual life-style to avoid the pain of a sudden loss.

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 $u_c(C)$ responds little too, and the marginal value of wealth $V_W(W, .) = u_c$ also responds little. Risk aversion to wealth bets is measured by the change of marginal utility when wealth changes. $(\partial \ln V_W / \partial \ln W = -W V_{WW} / V_W)$.

The argument is correct, but shows the problem with these models: What about the long run? 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 one insists that consumption growth must be independent over time (formally, an "endowment economy"), which is a good approximation to the data, then the model must feature strong interest rate variation to keep consumers from trying to adapt smoothly to wealth shocks. For example, consumers all want to save if wealth goes up, thereby slowly increasing consumption. If consumption growth is to be 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, one may fix interest rates 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, models such as this predict counterfactually that consumption growth will be positively 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 pathologies 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. This balance means that consumption can be a random walk with constant interest rates, consistent with the data. But alas, it also 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) come back towards 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

Let's look at the low risk aversion issue from 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 something important about your opportunities that losing a bet does not; you therefore shy away from stocks even though you 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 candidate for extra state variables are variables that describe changes in expected returns. If stock prices rise, you do learn something important that you would not learn from winning a bet: You learn 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]. Now, 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 means that we need *even more* risk aversion to explain the equity premium!

Consistency with individual consumption behavior

The low risk aversion models face one more supreme 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. But unless you're Bill

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

Gates, interest rates won't change in response to an individual wealth shock. 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 if anything *too much, too soon*, and have 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 this prediction by specifying an *external* habit; each person's habit responds to everyone else's consumption. They feel the need to 'keep up with the Joneses' as advocated by Abel (1990). Though this 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 form that high current returns mean low future returns. I can't say the goal is impossible, but I have outlined the challenges that such a model faces.

3.5 Heterogeneous agents and idiosyncratic risks

In the above discussion, I did not recognize any difference between people. Technically, I reviewed "representative agent" or "complete market" models. We are all different, so why bother looking at such models?

There is a strong reason to do so. Ultimate reductionism is not successful in finance and economics any more than it is in other branches of science – one does not try to figure out if a drug will cure a disease by starting with quantum mechanics. While stating an assumption such as "all people are identical" seems obviously foolish, it is not so foolish to hope that we can understand aggregates with aggregates, without having explicitly to take account of the differences between people. For example, consider a statement like "interest rates rise when future prospects of the economy seem brighter, because people would like to borrow more, to finance investments, and people want to save less because there is less need to provide for future income." This statement ties interest rates to aggregate behavior, without recognizing differences between people. While differences are there, they are hopefully not relevant to the basic story.

It would be awfully nice if we could tell stories like this one in order to describe the behavior of aggregate quantities. The fact that representative agent models have dominated growth, macroeconomics, international trade and so on from the beginning is testimony to the power of such arguments.

But if one dislikes the extensive modification that representative agent models have had to go through just to explain the equity premium and return predictability with low and constant interest rates, perhaps we should give up on this two century old paradigm and investigate how individual idiosyncracies may lie at the bottom of an aggregate phenomenon, the equity premium.

3.5.1 The empirical hurdle.

Idiosyncratic risk explanations face a big empirical challenge. Go back to the basic Sharpe ratio equation (6) which I repeat here,

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

This relation 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% 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% 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% per year standard deviation; for a more traditional $\gamma = 2.5$, we need 100% per year standard deviation! Even 10% 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 (losing your job, etc.) 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 $E(\varepsilon^i|\Delta c^a) = E(\varepsilon^i|r) = 0$. 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 $\sigma(\Delta c^i)$ rises.

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

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

3.5.2 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 for themselves a very steady consumption stream and resulting in a model that predicts a low equity premium.

Start by handing out income to consumers; we could 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. A natural approach is to shut down these markets too. Shutting down both markets is of course sensible; unemployment insurance is not perfect, and after one's home equity is exhausted it is hard to borrow. The point is that we will likely have to include such frictions in a theoretical model to make idiosyncratic risk have any bite.

But nothing stops our borrowing-constrained consumers from saving, and that is what they do next: 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, borrowing and saving can only insure against transitory income. If we make idiosyncratic shocks permanent, then consumers cannot use borrowing and saving – instruments that smooth consumption over *time* – to smooth consumption across states of nature. If losing your job means losing it forever, there is no point in smoothing consumption by borrowing, and planning to pay it back when you get a new job.

Heaton and Lucas (1996) 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 1/2 of the observed excess average stock return. It also predicts counterfactually that interest rates are as volatile as stock returns, and that individuals have huge (10-30%, depending on specification) consumption growth uncertainty. Heaton and Lucas have to set the net stock of debt to zero to even getting 1/2 of the observed average stock return. If they allow a stock of government bonds, then households can smooth consumption by holding the stock of bonds; the interest rate volatility falls but so does the equity premium.

Even if all this theory worked, all we would accomplish is to keep consumers from using markets to smooth away any idiosyncratic income shocks we give them. We are still faced with the empirical problem above: Even if we can write down a model in which consumers are stuck with, say, 10% annual standard deviation of idiosyncratic consumption growth, to first order that fact does *nothing* to help explain the equity premium.

3.5.3 A model that works

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

As I argued above, if we give consumers "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. In addition, if marginal utility were linear, an increase in variance would have no effect on the average level of marginal utility. Therefore, Constantinides and Duffie specify non-quadratic power utility, and the interaction of these two features produces an equity premium.

The Constantinides-Duffie model and the Campbell-Cochrane model are in fact quite similar in spirit, though the Constantinides-Duffie model is built on incomplete markets and idiosyncratic risks, while the Campbell-Cochrane model is firmly 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 much the loss of wealth of a bad market return per se. They fear that loss of wealth because it tends to come in recessions, in one case defined as times of heightened labor market risk, and in the other case defined as a fall of consumption relative to its recent past. This recession state-variable or risk-factor drives most variation in 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 will 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 existence proof is particularly stunning in Constantinides and Duffie's case. Many authors (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 case 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 I lay out some of the issues below.

A simple version of the model

Here is a very simplified version of the Constantinides-Duffie 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},

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

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

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

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 $C_{it} = I_{it}$.) His 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 I outlined above.

[12]To prove this assertion, just substitute in 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 $E[f(\eta y)] = E[E(f(\eta y|y))]$. Now, with η normal, $E(\exp[-\gamma\eta_{it}y_t] \mid y_t) = \exp\left[\frac{1}{2}\gamma^2 y_t^2\right]$. Therefore, we have

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

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

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

In the general model, Constantinides and Duffie define

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

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

$$1 = E_{t-1}\left[m_t R_t\right] \text{ for all } R_t. \tag{14}$$

Then, they let

$$C_{it} = \delta_{it} C_t$$

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

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 could be either a new view of stock market (and macroeconomic) risk, or just an extremely clever 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 of stock returns are low at times of high labor market risk?

This model does not change the empirical puzzle discussed in section 3.5.1. To get power utility consumers to shun stocks, they still must have tremendously volatile consumption growth or high risk aversion, and the calculations of section 3.5.1 apply to

[13] Astute readers will notice the possibility that the square root term in (11) and (12) might be negative. Constantinides and Duffie rule out this possibility by assuming that the discount factor m satisfies

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

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 $m_t = \max\left[a + bR_t, e^{\rho}\left(\frac{C_t}{C_{t-1}}\right)^{\gamma}\right]$ to satisfy (14). However, neither this construction nor a discount factor satisfying (13) is guranteed to exist for a given set of assets. The restriction (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. Ledoit and Bernardo (1997) show that restrictions like (13) are equivalent to restrictions on the maximum gain/loss ratio available from the set of 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 postive discount factor that prices a single asset return $1 = E(R^{-1}R)$, but does not necessarily satisfy restriction (13). For high R, we can have very negative $\ln 1/R$. This is why the lines in Figure 6 below run into the horizontal axis at high R.

individual consumption in the Constantinides-Duffie model, for example equation (10). 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 10-50% volatility ($\sigma(\Delta c)$) that we require to reconcile the Sharpe ratio with low risk aversion means that the distribution of consumption (and income) must also spread out by 10-50% 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 by 10-50% per year if we are to have 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 $\sqrt{0.2} = .45$ or 45% at age 20 to $\sqrt{0.6} = .77$ or 77% at age 60. (77% means that an individual one standard deviation better off than the mean consumes 77% more than the mean consumer.) We are back to about 1% per year.

Finally, and most crucially, 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 ($\gamma = 1$ 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 $\gamma = 2$ line for example, is it plausible that a year with 5% market return would show a 10% cross-sectional variation in consumption growth, while a mild 5% decline in the market is associated with a 25% cross-sectional variation?

One can in fact regard the Heaton and Lucas (1986) model as an empirical assessment of these issues. Rather than construct a labor income process that would generate 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. Of course, this view is at best preliminary evidence. They did not nest the exact Constantinides-Duffie specification as a special case, nor did they test whether one can reject the Constantinides-Duffie specification.

All of these empirical problems are avoided if we allow high risk aversion rather than a large risk to drive the equity premium. The $\gamma = 25$ line in Figure 6 looks possible; a $\gamma = 50$ 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*.

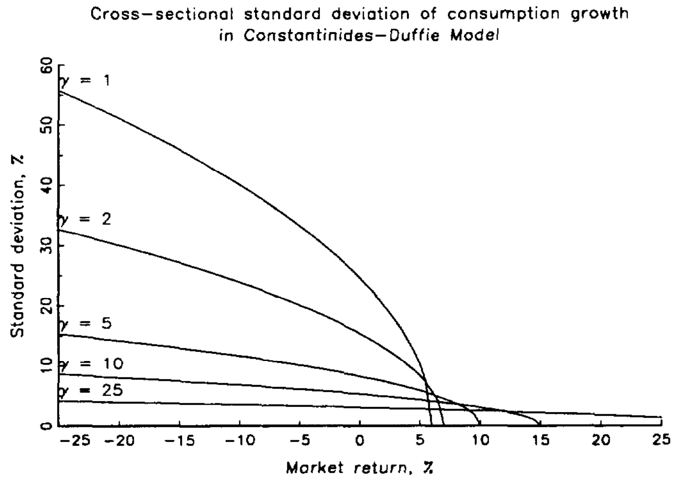

Figure 6: Cross-sectional standard deviation of individual consumption growth as a function of the market return in the Constantinides-Duffie model. The plot is the variable $y_t = \sqrt{\frac{2}{\gamma(\gamma+1)}}\sqrt{\ln\frac{1}{R_t} + \rho + \gamma\ln\frac{C_t}{C_{t-1}}}$. Parameter values are $\rho = 0.05$, $\ln C_t/C_{t-1} = 0.01$, and γ and $\ln R_{t+1}$ as graphed.

Labor market risk correlated with the stock market does not seem large enough to account for the equity premium without high risk aversion.

3.5.4 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, we 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.

We could instead 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 in fact have consumption that is more volatile and more correlated with stock returns than that of non-stockholders. But their consumption is still not volatile and correlated enough to satisfy the right hand side of (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. True to this prediction, they find that among rich households, those with more proprietary income hold fewer stocks in their portfolios. Thus they paint an interesting picture of the equity premium: In the past most stock was held by rich people, and most rich people were proprietors whose other incomes (and consumption) are quite volatile and covary strongly with the market. This is a hard crowd to sell stocks to, so they have required a high risk premium. More formally, our models specify that stock market risk is spread as evenly as possible through the population. If a risk is only shared in a small group of people, it will have to give higher rewards for 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 long literature I reviewed in section 3.5.2 shows that even quite high transactions costs, borrowing constraints, and so forth 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. All transactions costs are declining through financial deregulation and innovation. The explosion in tax-deferred pension plans, no-load mutual funds, and so forth means that everyone can own stocks, and more and more people are in fact doing so, driving up prices and driving down prospective returns. Equation (6) will hold much better for the average consumer in the future. In part we will see with a lower equity premium. We also should see more volatile consumption, better correlated with the market, which will be a fundamental change in the nature of business cycles.

3.6 Technology and investment

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

This is an attractive exercise. For example, the relative price of apples and oranges is basically set by technology, the relative number of apples vs. oranges that can be grown on the same acre of land. We don't need psychological studies or a deep understanding of consumers' desires to figure out what the price should be. Technically, if technology is (close to) linear, it will determine relative prices while preferences will determine quantities. Can we make the same arguments 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 in to new capital in the future. In equations, together with the most common functional forms,

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

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 (15), the investment return is

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

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% in a year, the investment return only varies by 1%, far below the 17% 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's hard to get any work done on the day the furniture is delivered, and it's awfully 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 (15) to

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

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 $(1-\delta)$ units of capital tomorrow. Hence the price of capital in terms of output is

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

where the last equality uses the quadratic functional form. (This is the famous 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 we expect stock prices to be high when investment is high, or equivalently, we expect firms to issue stock and invest a lot when stock prices are high. The investment return is now

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

Comparing (19) with (16), we see that the investment return contains a new term proportional to the change in investment. Since investment is quite volatile, for a large enough value of the parameter a, this model can be consistent with the volatility of the market return. Looking at equation (18), we see that the last term adds price changes to our model of the investment return.

How does all this work? Figure 7 presents the investment-output ratio along with the value-weighted price-dividend 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 in fact 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 70s despite dismal stock prices. The recent surge in the market is matched by a surge in investment, a fact that I will return to shortly.

This kind of model has been subject to an enormous formal empirical effort, which pretty much confirms the simple 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 price-dividend ratio. Second, early tests related investment to interest rates, imposing a constant risk premium, and they 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 (18) taken literally allows *no* residual. If prices deviate one iota from the right hand side of (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 cashflow 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 about 10%, we see that we need $a \approx 2$ to

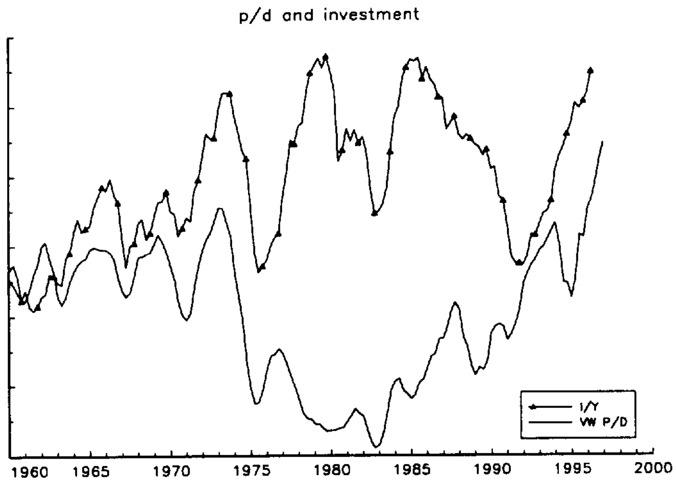

Figure 7: Investment - output ratio, and price - dividend ratio of value-weighted portfolio. Investment = gross fixed investment, output = gross domestic product. The series are stretched to fit on the same graph.

rationalize the 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 typically $\frac{a}{2}\frac{I}{K}\frac{I}{Y} = \frac{2}{2}\left(\frac{1}{20}\right) \times 15\% = 0.75\%$ 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.

Alas, this model does not yet fulfill our hope of being able to determine the equity premium by technological considerations alone. The trouble is that 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. But if the interest rate is low, there is not much we 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.

3.6.1 Implications of the recent surge in investment and stock prices

For our purpose, there are two important features of Figure 7. First, the association between stock returns and investment verifies that at least one connection between stock returns and the real economy works in some respects as it should. This observation argues against the view that stock market swings are due entirely to waves of irrational optimism and pessimism. It is also comforting to verify 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.

Second, the surge in stock prices since 1990 has been accompanied by a surge in investment. If expected stock returns and hence the cost of capital are low, then investment should be high and it is. 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 in fact quite low.

3.7 General equilibrium

We cannot say we really understand the equity premium until we put the utility function and production function modifications together, to construct complete explicit economic models that replicate the asset pricing facts. Such efforts should also at least preserve if not enhance our ability to understand the broad range of dynamic microeconomic, macroeconomic, international and growth facts that the standard models were constructed around. This effort is likely to be very challenging. Anything that affects the relation between consumption and asset prices will drastically affect the relation between consumption and investment, since asset prices mediate the consumption-investment decision, and that decision lies at the heart any dynamic macroeconomic model. The effort is just beginning; we have learned a bit about how to go about this task, but no completely satisfactory model as yet.

Jermann (1997) tried putting habit persistence consumers in a model with a standard technology like (15), which is almost completely standard in real 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 (17), as the production-side literature had found necessary. This modification resulted in a high equity premium, but also large variation in riskfree rates.

Boldrin, Christiano and Fisher (1997) also added habit-persistence preferences to real business cycle models with frictions in the allocation of resources to two sectors. They generate about 1/2 the historical Sharpe ratio. They find some quantity dynamics are improved over the standard model. However, they still predict 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 $f(K)$. Models with this kind 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 a beautiful observational equivalence result: A model with standard preferences and a model with non-state-separable preferences can predict the same path of *quantity* variables (output, investment, consumption, etc.) but differ dramatically on *asset prices.* This result offers one explanation of how the real business cycle and growth literature could go on for 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.

4 Implications

The standard economic models, models that have been used with great success to describe growth, macroeconomics, international economics, and even dynamic microeconomics, do not predict anything like the historical equity premium, let alone the predictability of returns. After 10 years of effort, a range of deep modifications to the standard models shows the promise to explain 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 finished the task of going back to see if the deep modifications necessary to explain asset market data retain the models' previous successes at describing quantity data.

The modified models are indeed 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 microeconomic 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 we have seen that 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, you have 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% excess return. The standard model was right after all, and the historical stock returns in the US were largely 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 his estimate of the future equity premium even more than suggested by the statistical uncertainty documented above.

4.0.1 Before you sell.

In sum, the long-term average stock return may well be lower than the postwar 8% 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 portfolio at all. For every individual who keeps his 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 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 "market time", buying more stocks at times of high expected returns and vice versa. Every investor in the Campbell-Cochrane or Constantinides-Duffie models (for example) holds exactly the same market 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, pulling out of the market at times of high price ratios such as the present, you have to think, who is there who is going to be more *in* the market than average now? 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 (3). Assuming low risk aversion, and no labor income or other reason for time-varying risk exposure or risk aversion, solutions to such problems typically suggest large portfolio shares in equities, and a strong market timing approach, sometimes highly leveraged and sometimes (now) even short. (See Barberis 1997 and Brandt 1997.) But if everyone followed this advice, the equity premium and the predictable variation in expected returns would disappear. Everyone trying to buy stocks would simply drive up their prices; everyone trying to market time 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, an individual in a circumstance that *is different* from everyone else can in fact understand why he should behave differently from the crowd. If you have no "habits", or a if you have a sinecure such that you are immune to labor income shocks; if you are unaffected by the "state variables" or "risk factors" that drive the stock market premium, then by all means go your own way. Obviously, the current state of the art is not well-enough advanced to provide solid advice along these lines, but the question is worth asking.

The last possibility of course is that one thinks one is smarter than everyone else; that the equity premium and predictability are just patterns that are ignored by other people. This is a dangerous stance to take. Unlike the children of Lake Woebegone, we can't all be above average. Someone is wrong in the view that he's 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, though spelled out at scholarly length, are hardly a secret. We might 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 which 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.

5 References

Abel, Andrew B. 1983, "Optimal Investment under Uncertainty," *American Economic Review* 73, 228-33.

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

Abel, Andrew B., and Janice C. Eberly, 1996, "Optimal Investment with Costly Reversibility" *Review of Economic Studies* 63, 581-594.

Barberis, Nicholas 1997 "Investing for the Long Run when Returns are Predictable" CRSP working paper 439.

Aiyagari, S. Rao and Mark Gertler, 1991 "Asset Returns with Transactions Costs and Undiversifiable Individual Risk," *Journal of Monetary Economics* 27 311-331.

Barsky, Robert B., Miles S. Kimball, F. Thomas Juster and Matthew D. Shapiro, 1997, Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Survey, NBER Working Paper No. 5213.

Brandt, Michael, 1997, "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach," Ph.D. dissertation, University of Chicago.

Black, Fischer, 1976, "Studies of Stock Price Volatility Changes," *Proceedings of the 1976 Meetings of the Business and Economic Statistics Section, American Statistical Association,* 177-181.

Black, Fischer and Myron Scholes, 1973, "The Valuation of Options and Corporate Liabilities," *Journal of Political Economy,* 81, 637-654.

Boldrin, Michele, Lawrence J. Christiano, and Jonas D.M. Fisher, 1997, "Habit Persistence, Mean Asset Returns and Real Business Cycle Models", Manuscript.

Brown, Stephen J., William N. Goetzmann and Stephen A. Ross, 1995, "Survival," *Journal of Finance* 50 (3) 853-73

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

Campbell, John Y., 1996, "Understanding Risk and Return," *Journal of Political Economy* 104, 298-345.

Campbell, John Y. and John H. Cochrane, 1997, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," Working paper.

Campbell, John Y. and Robert J. Shiller, 1988a, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies* 1, 195-227.

Cochrane, John H., 1991a, "Volatility Tests and Efficient Markets: A Review Essay," *Journal of Monetary Economics* 127 463-485.

Cochrane, John H., 1991b, "Explaining the Variance of Price-Dividend Ratios," *Review of Financial Studies* 15, 243-280.

Cochrane, John H. 1991c, "Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations" *Journal of Finance* 146 207-234.

Cochrane, John H. and Lars Peter Hansen, 1992, "Asset Pricing Lessons for Macroeconomics," in Olivier Blanchard and Stanley Fischer, Eds. *1992 NBER Macroeconomics Annual.*

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

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

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

Craine, Roger, 1993, "Rational Bubbles: A Test," *Journal of Economic Dynamics and Control* 17 829-46.

Daniel, Kent and David Marshall, 1997, "Consumption-Based Modeling of Long-Horizon Returns" Working paper.

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

Deaton, Angus, and Christina Paxson, 1994, "Intertemporal Choice and Inequality", *Journal of Political Economy* 102437-467.

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

Eichenbaum, Martin, Lars Peter Hansen and Kenneth J. Singleton, 1988, "A Time Series Analysis of Representative Agent Models of Consumption and Leisure Choice Under Uncertainty" *Journal of Business and Economic Statistics* 103 51-78

Epstein, Larry G. and Stanley E. Zin,1989, "Substitution, Risk Aversion and the Temporal Behavior of Asset Returns," *Journal of Political Economy* 99, 263-286.

Fama, Eugene F. and Kenneth R. French, 1988, "Permanent and Temporary Components of Stock Prices,"*Journal of Political Economy,* 96 246-73.

Fama, Eugene F. and Kenneth R. French, 1988, "Dividend Yields and Expected Stock Returns", *Journal of Financial Economics* 22, 3-27.

Fama, Eugene F. and Kenneth R. French, 1989, "Business Conditions and Expected Returns on Stocks and Bonds", *Journal of Financial Economics* 25, 23-49.

Fama, Eugene F. and Kenneth R. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics* 33, 3-56.

Fazzari, Steven M.. Robert Glenn Hubbard, Robert Glenn and Bruce C. Petersen,1988, "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*; 0 141- 95.

Ferson, Wayne E. and George Constantinides, 1991, "Habit Persistence and Durability in Aggregate Consumption: Empirical Tests", *Journal of Financial Economics* 29, 199–240.

Friend, I. and M. Blume (1975), "The Demand for Risky Assets," *American Economic Review* 55, 900-922.

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

Hansen, Lars Peter, Thomas J. Sargent, and Thomas D. Tallarini, 1997, "Robust Permanent Income and Pricing" manuscript, University of Chicago

Heaton, John C., 1995, "An Empirical Investigation of Asset Pricing with Temporally Dependent Preference Specifications", *Econometrica* 63, 681–717.

Heaton, John and Deborah J. Lucas 1996 "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing" *Journal of Political Economy* 104 443-487.

Heaton, John and Deborah Lucas 1996 "Saving Behavior and Portfolio Choice: Which Risks Matter?" Northwestern University Working Paper.

Hodrick, Robert, 1992 "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement" *Review of Financial Studies* 5 357-86.

Jagannathan, Ravi and Z. Wang, 1996, "The Conditional CAPM and the Cross-Section of Expected Returns", *Journal of Finance* 51, 3–53. (Also updated in Staff Report Staff Report 208, Federal Reserve Bank of Minneapolis, 1997.)

Jermann, Urban, 1997, "Asset Pricing in Production Economies," Forthcoming, *Journal of Monetary Economics*

Kandel, Shmuel and Robert F. Stambaugh, 1990, "Expectations and Volatility of Consumption and Asset Returns," *Review of Financial Studies* 3, pp. 207-232.

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

Kocherlakota, Narayana R., 1990, "On the 'Discount' Factor in Growth Economies," *Journal of Monetary Economics,* 25 43- 47.

Kocherlakota, Narayanna, 1996 "The Equity Premium: it's Still a Puzzle," *Journal of Economic Literature* 34, 42-71

Bernardo, Antonio and Olivier Ledoit, 1996, "Gain, Loss and Asset Pricing," Anderson Graduate School of Management Working Paper #16-96.

Lucas, Deborah J.1994, "Asset Pricing with Undiversifiable Income risk and Short Sales Constraints: Deepening the Equity Premium Puzzle" *Journal of Monetary Economics* 34, 325-41.

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

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

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

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

Poterba, James and Lawrence J. Summers, 1988 "Mean Reversion in Stock Prices," *Journal of Financial Economics,* 22, 27- 59.

Reyfman, Alexander 1997, "Labor Market Risk and Expected Asset Returns," Ph.D. Dissertation (Ch. 3), University of Chicago

Rietz, Thomas A., 1988, "The Equity Risk Premium: A Solution", *Journal of Monetary Economics* 22, 117-131.

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

Telmer, Chris I., 1993, "Asset Pricing Puzzles and Incomplete Markets," *Journal of Finance* 48, 1803-32.

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

6 Appendix: Derivations

6.1 Variance decomposition

Massaging an identity,

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

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 expand (20) to

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

and then iterate to

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

6.2 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 $dP/P + Ddt$. 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

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

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 riskfree 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)

$$E_t(r) - r^f = cov_t [\Delta c, r] = \sigma_t(\Delta c) \sigma_t(r) \rho_t(\Delta c, r).$$

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

6.3 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

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

Using the power functional form,

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

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 6

For small risks, we can approximate

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

$$-u'(C)x \approx \frac{1}{2}u''(C)y^2$$

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

$$\frac{x}{C} \approx \gamma\left(\frac{y}{C}\right)^2$$

or

$$\frac{x}{y} \approx \gamma\left(\frac{y}{C}\right).$$