The very basic investment and constant-growth models from introductory finance courses can be used to interpret the long-run unconditional historical data on returns. So, let’s begin with the basic model:

\[
\frac{E_{t+1}}{E_t} = 1 + [(b)(ROE)],
\]

where

- \( E \) = earnings
- \( b \) = the retention rate
- \( ROE \) = return on equity

So that, with investment at time \( t \) denoted by \( I_t \),

\[
ROE = \frac{E_{t+1} - E_t}{I_t}
\]

and

\[
b = \frac{I_t}{E_t}
\]

therefore, the growth rate of earnings is

\[
(b)(ROE) = \frac{E_{t+1} - E_t}{E_t}.
\]

This model implies that the growth rate in earnings is the retention rate times the return on equity, \( b \cdot ROE \). In discussing the models, I would like to stress an important point: If you are interpreting the growth in earnings as being the retention rate times the return on equity, you have to be very careful when you are working with historical data. For example, does the retention rate apply only to dividends or to dividends and other payouts, such as share repurchases? The distinction is important because those proportions change in the more recent period. And if you make that distinction, you have to make a distinction between aggregate dividends and per share dividends because the per share numbers and the aggregate numbers will diverge. In working with the historical data, I have attempted to correct for that aspect.
Table 1 gives the arithmetic average data for growth rates in GNP, earnings, and dividends for two periods: 1951–2000 and 1972–2000. (I used the 1972–2000 period because it mirrors the same period shown in Figure 1.) The earnings growth rates are so much more volatile than the dividend growth rates. And because of the volatility effect on arithmetic averages, GNP and earnings exhibit very similar growth rates from the early 1970s to the present. Dividends (and Table 1 shows the growth rate of actual dividends, not payouts) have grown much less than earnings for two reasons: First, dividends are less volatile, and second, dividend substitution is occurring. Corporations are not providing shareholders the same constant fraction of earnings (in the form of dividends) that they were in the past.

Despite the 1972–2000 data, it seems to me that earnings are not going to grow as fast as or faster than GNP in the future. This notion seems to be consistent with long-term historical data, and it fits my view of how politics works on the economy. If you accept that notion, it has immediate implications for the future.

First, under any reasonable underlying assumptions about inflation, equity risk premiums cannot be much more than 3 percent (300 bps) because the earnings growth rate is constrained unconditionally in the long run by the real growth rate of the economy, which has been in the range of 1.5–3.0 percent. Second, as Table 2 shows, for an S&P level of about 1,000, you simply cannot have an equity risk premium any higher than 2 percent, 2.5 percent, or (at most) 3 percent.
What simplifying assumptions can be made to work with the unconditional data? I have made some relatively innocuous simplifying assumptions. First, that \( b \) should adjust until the cost of capital equals the ROE at the margin. To be very conservative, therefore, I will assume that the ROE equals the cost of capital, or expected returns, in the aggregate. The problem that arises is: What if the retention rate times the cost of capital (that is, the minimal expected return on equity), \( bk \), is greater than GNP growth? The second assumption deals with this possibility: I assume \( bk \) cannot be greater than GNP growth because political forces will come into play that will limit the ROE if earnings start to rise as a fraction of GNP.

The relationship between aggregate earnings and GNP is one of the research questions that I have been unable to find interesting papers on—perhaps because I have not searched well enough—but I want to bring up the subject to this group. It seems to me that if aggregate earnings start to rise, and Robert Shiller mentioned several reasons why it can happen [see the “Current Estimates and Prospects for Change” session], then tax rates can change, antitrust regulation can change (one of Microsoft’s problems probably was that it was making a great deal of money, which is an indication that some type of regulation may be necessary), labor regulation can change, and so forth. And these variables can change ex post as well as ex ante. So, once a company starts making superior returns using a particular technology, the government may step in ex post and limit those returns. The critical research question is how earnings relate to GNP.

The constant-growth model is

\[
P = \frac{D}{k - g}
\]

or

\[
k = \frac{D}{p} + g,
\]

where

\( P \) = price
\( D \) = dividends
\( k \) = cost of capital
\( g \) = growth rate

What I am going to do is just an approximation because I am going to work with aggregate, not per share, data. I am going to assume that total payouts are 1.5 times dividends.\(^1\) Payouts will probably be lower in the future, but if I work with aggregate

\(^1\)This choice is based on recent findings by Jagannathan, Stephens, and Weisbach (2000) that we are seeing significant payouts today.

payouts, then \( g \) should be the growth rate in aggregate potential payouts, which I will characterize as earnings.

One of the implications of the simplifying assumptions I have made, and it relates to the data that Jeremy Siegel just produced [“Historical Results I”], is that the expected returns on stocks should be equal to the earnings-to-price ratio. (In the more complicated equations, you have situations in which the ROE is not exactly equal to expected returns, but for my long-run data, the simplifying assumption that earnings yield equals the expected ROE is fine.) So, with these assumptions,

\[
P = \frac{D}{k - g}
\]

\[
k = \frac{D}{k - bk}
\]

\[
= 1 - (b) \left( \frac{E}{1 - b} \right) (k)
\]

or

\[
k = \frac{E}{P'}.
\]

A further implication is that if \( g \) is constrained to be close to the growth of GNP, then it is reasonable to substitute GNP growth for \( g \) in the constant-growth model. The implication of this conclusion is that the expected return, or cost of capital, in the long run should unconditionally be about 1.5 times the dividend-to-price ratio plus GNP growth:

\[
k = 1.5 \frac{D}{P} + \text{GNP growth}.
\]

With this background, we can now look at some of the data.

**Earnings and GNP**

Figure 1 allows a comparison of dividends/GNP and (after-tax) earnings/GNP for 1950 through July 2001.\(^2\) The data begin in 1950 because Fama believed that the data before then were unreliable. Figure 1 shows that, historically, earnings have declined as a fraction of GNP in this period. My assumption that earnings keep up with GNP works from about 1970 on, but I am looking at the picture in Figure 1 in order to make that conclusion. The ratio of earnings to GNP depends on a lot of things: the productivity of labor, capital, the labor-to-capital ratio, taxes, and (as I said earlier) a host of political forces. Figure 1 shows that earnings have, at best, kept up with GNP.

\(^2\)These data were provided by Eugene Fama, who attributed them to Robert Shiller.
Valuation

Why is the market so high? As an aside, and this concern is not directed toward our topic today of the equity risk premium, but I think it is an interesting question: Why is the market where it is today relative to where it was on September 10 or September 9 or just before the events of September 11, 2001? The market then and now is at about the same level. Almost every economist and analyst has said that the September 11 attacks accelerated a recession, that they changed perceptions of risk, and so forth. It is curious to me that such a situation does not seem to be reflected in market prices.

But in general, why is the market so high? I believe three possible explanations exist. One idea, and I consider it a “rational” theory, is that stocks are simply seen as less risky than in the past. I do not know whether the behavioral theories are rational or not, in the sense that prices are high because of behavioral phenomena that are real and are going to persist. If so, then those phenomena—as identified by Jeremy Siegel and Richard Thaler [see the “Theoretical Foundations” session]—are also rational. In that case, the market is not “too high”; it is not, in a sense, a mistake. It is simply reflecting characteristics of human beings that are not fully explained by economic theories.

Another rational explanation has been given less attention but is the subject of a recent paper by McGrattan and Prescott (2001). It is that the valuation of equities is fundamentally determined by taxation. McGrattan and Prescott argue that the move toward holding equities in nontaxable accounts has led to a drop in the relative tax rate on dividends. Therefore, stock prices should rise relative to the valuation of the underlying capital and expected returns should fall. This effect is a rational tax effect.

Both this theory and the theory that stocks are now seen as less risky say that the market is high because it should be high and that, looking ahead, equities are going to have low expected returns, or low risk premiums—about 2 percent—but that investors have nothing to worry about.

The final explanation, which I attribute to John Campbell and Robert Shiller, focuses on the view that equity prices today are simply a mistake. (I suppose mistakes are a behavioral phenomenon, but presumably, they are not as persistent as an underlying psychological condition.) Now, when people realize they have made a mistake, they attempt to correct the behavior. And those corrections imply a period of negative returns from the U.S. equity market before the risk premium can return to a more normal level.

Closing

To close, I want to repeat that, to me, the fundamental historical piece of data that needs more explanation is the relationship between the aggregate behavior of earnings and GNP—what it has been in the past and what it can reasonably be going forward. This relationship is interesting, and I look forward to hearing what all of you have to say about it. In my view, it is the key to unlocking the mystery of the equity risk premium’s behavior.