Rational Expectations and Asset Prices

We are now going to switch gear and leave the IS-MP-PC model behind us. One of the things we've focused on is how people formulate expectations about inflation. We put forward one model of how these expectations were formulated, an adaptive expectations model in which people formulated their expectations by looking at past values for a series. Over the next few weeks, we will look at an alternative approach that macroeconomists call "rational expectations". This approach is widely used in macroeconomics and we will cover its application to models of asset prices, consumption and exchange rates.

Rational Expectations and Macroeconomics

Almost all economic transactions rely crucially on the fact that the economy is not a "one-period game." In the language of macroeconomists, most economic decisions have an *intertemporal* element to them. Consider some obvious examples:

- We accept cash in return for goods and services because we know that, in the future, this cash can be turned into goods and services for ourselves.
- You don't empty out your bank account today and go on a big splurge because you're still going to be around tomorrow and will have consumption needs then.
- Conversely, sometimes you spend more than you're earning because you can get a bank loan in anticipation of earning more in the future, and paying the loan off then.
- Similarly, firms will spend money on capital goods like trucks or computers largely in anticipation of the benefits they will bring in the future.

Another key aspect of economic transactions is that they generally involve some level of *uncertainty*, so we don't always know what's going to happen in the future. Take two of the examples just given. While it is true that one can accept cash in anticipation of turning it into goods and services in the future, uncertainty about inflation means that we can't be sure of the exact quantity of these goods and services. Similarly, one can borrow in anticipation of higher income at a later stage, but few people can be completely certain of their future incomes.

For these reasons, people have to make most economic decisions based on their subjective *expectations* of important future variables. In valuing cash, we must formulate an expectation of future values of inflation; in taking out a bank loan, we must have some expectation of our future income. These expectations will almost certainly turn out to have been incorrect to some degree, but one still has to formulate them before making these decisions.

So, a key issue in macroeconomic theory is how people formulate expectations of economic variables in the presence of uncertainty. Prior to the 1970s, this aspect of macro theory was largely *ad hoc*. Different researchers took different approaches, but generally it was assumed that agents used some simple extrapolative rule whereby the expected future value of a variable was close to some weighted average of its recent past values. However, such models were widely criticised in the 1970s by economists such as Robert Lucas and Thomas Sargent. Lucas and Sargent instead promoted the use of an alternative approach which they called "rational expectations." This approach had been introduced in an important 1961 paper by John Muth.

The idea that agents expectations are somehow "rational" has various possible interpretations. However, when Muth's concept of rational expectations meant two very specific things:

- They use publicly available information in an efficient manner. Thus, they do not make systematic mistakes when formulating expectations.
- They understand the structure of the model economy and base their expectations of variables on this knowledge.

To many economists, this is a natural baseline assumption: We usually assume agents behave in an optimal fashion, so why would we assume that the agents don't understand the structure of the economy, and formulate expectations in some sub-optimal fashion. That said, rational expectations models generally produce quite strong predictions, and these can be tested. Ultimately, any assessment of a rational expectations model must be based upon its ability to fit the relevant macro data.

How We Will Describe Expectations

We will start with some terminology to explain how we will represent expectations. Suppose our model economy has a uncertainty so that people do not know what is going to happen in the future. Then we will write $E_t Z_{t+2}$ to mean the expected value the agents in the economy have at time t for what Z is going to be at time $t+2$. In other words, we assume people have a distribution of potential outcomes for Z_{t+2} and $E_t Z_{t+2}$ is mean of this distribution.

It is important to realise that E_t is not a number that is multiplying Z_{t+2}. Instead, it is a qualifier explaining that we are dealing with people's prior expectations of a Z_{t+2} rather than the actual realised value of Z_{t+2} itself.

Throughout these notes, we will use some basic properties of the expected value of distributions. Specifically, we will use the fact that expected values of distributions is what is known as a *linear operator*. What is meant by that is that

$$E_t\left(\alpha X_{t+k} + \beta Y_{t+k}\right) = \alpha E_t X_{t+k} + \beta E_t Y_{t+k} \tag{1}$$

Some examples of this are the following. The expected value of five times a series equals five times the expected value of the series

$$E_t\left(5X_{t+k}\right) = 5E_t\left(X_{t+k}\right) \tag{2}$$

And the expected value of the sum of two series equals the sum of the two expected values.

$$E_t\left(X_{t+k} + Y_{t+k}\right) = E_t X_{t+k} + E_t Y_{t+k} \tag{3}$$

We will use these properties a lot, so I won't be stopping all the time to explain that they are being used.

Asset Prices

The first class of rational expectations models that we will look relate to the determination of asset prices. Asset prices are an increasingly important topic in macroeconomics. Movements in asset prices affect the wealth of consumers and businesses and have an important influence on spending decisions. In addition, while most of the global recessions that preceded the year 2000 were due to boom and bust cycles involving inflation getting too high and central banks slowing the economy to constrain it, the most recent two global recessions—the "dot com" recession of 2000/01 and the "great recession" of 2008/09—were triggered by big declines in asset prices following earlier large increases. A framework for discussing these movements is thus a necessary part of any training in macroeconomics.

In these notes, we will start by considering the case of an asset that can be purchased today for price P_t and which yields a *dividend* of D_t. While this terminology obviously fits with the asset being a share of equity in a firm and D_t being the dividend payment, it could also be a house and D_t could be the net return from renting this house out, i.e. the rent minus any costs incurred in maintenance or managment fees. If this asset is sold tomorrow for price P_{t+1}, then it generates a rate of return on this investment of

$$r_{t+1} = \frac{D_t + \Delta P_{t+1}}{P_t} \tag{4}$$

This rate of return has two components, the first reflects the dividend received during the period the asset was held, and the second reflects the *capital gain* (or loss) due to the price of the asset changing from period t to period $t+1$. This can also be written in terms of the so-called *gross return* which is just one plus the rate of return.

$$1 + r_{t+1} = \frac{D_t + P_{t+1}}{P_t} \tag{5}$$

A useful re-arrangement of this equation that we will repeatedly work with is the following:

$$P_t = \frac{D_t}{1 + r_{t+1}} + \frac{P_{t+1}}{1 + r_{t+1}} \tag{6}$$

Asset Prices with Rational Expectations and Constant Expected Returns

We will now consider a *rational expectations* approach to the determination of asset prices. Rational expectations means investors understand equation (6) and that all expectations of future variables must be consistent with it. This implies that

$$E_t P_t = E_t \left[\frac{D_t}{1 + r_{t+1}} + \frac{P_{t+1}}{1 + r_{t+1}} \right] \tag{7}$$

where E_t means the expectation of a variable formulated at time t. The stock price at time t is observable to the agent so $E_t P_t = P_t$, implying

$$P_t = E_t \left[\frac{D_t}{1 + r_{t+1}} + \frac{P_{t+1}}{1 + r_{t+1}} \right] \tag{8}$$

A second assumption that we will make for the moment is that the expected return on assets equals some constant value for all future periods, unrelated to the dividend process.

$$E_t r_{t+k} = r \quad k = 1, 2, 3, ... \tag{9}$$

One way to think of this is that there is a "required return", determined perhaps by the rate of return available on some other asset, which this asset must deliver. With this assumption in hand and assuming that everyone knows the value of D_t, equation (8) can be re-written as

$$P_t = \frac{D_t}{1+r} + \frac{E_t P_{t+1}}{1+r} \tag{10}$$

The Repeated Substitution Method

Equation (10) is a specific example of what is known as a *first-order stochastic difference equation.*[1] Because such equations occur commonly in macroeconomics, it will be useful to write down the general approach to solving these equations, rather than just focusing only on our current asset price example. In general, this type of equation can be written as

$$y_t = a x_t + b E_t y_{t+1} \tag{11}$$

Its solution is derived using a technique called *repeated substitution.* This works as follows. Equation (11) holds in all periods, so under the assumption of rational expectations, the agents in the economy understand the equation and formulate their expectation in a way that is consistent with it:

$$E_t y_{t+1} = a E_t x_{t+1} + b E_t E_{t+1} y_{t+2} \tag{12}$$

[1] Stochastic means random or incorporating uncertainty. It applies to this equation because agents do not actually know P_{t+1} but instead formulate expectations of it.

Note that this last term ($E_t E_{t+1} y_{t+2}$) should simplify to $E_t y_{t+2}$: It would not be rational if you expected that next period you would have a higher or lower expectation for y_{t+2} because it implies you already have some extra information and are not using it. This is known as the *Law of Iterated Expectations*. Using this, we get

$$E_t y_{t+1} = a E_t x_{t+1} + b E_t y_{t+2} \tag{13}$$

Substituting this into the previous equation, we get

$$y_t = a x_t + ab E_t x_{t+1} + b^2 E_t y_{t+2} \tag{14}$$

Repeating this method by substituting in for $E_t y_{t+2}$, and then $E_t y_{t+3}$ and so on, we get a general solution of the form

$$y_t = a x_t + ab E_t x_{t+1} + ab^2 E_t x_{t+2} + ... + ab^{N-1} E_t x_{t+N-1} + b^N E_t y_{t+N} \tag{15}$$

which can be written in more compact form as

$$y_t = a \sum_{k=0}^{N-1} b^k E_t x_{t+k} + b^N E_t y_{t+N} \tag{16}$$

For those of you unfamiliar with the summation sign terminology, summation signs work like this

$$\sum_{k=0}^{2} z_k = z_0 + z_1 + z_2 \tag{17}$$

$$\sum_{k=0}^{3} z_k = z_0 + z_1 + z_2 + z_3 \tag{18}$$

$$\sum_{k=0}^{4} z_k = z_0 + z_1 + z_2 + z_3 + z_4 \tag{19}$$

and so on.

The Dividend-Discount Model

Comparing equations (10) and (11), we can see that our asset price equation is a specific case of the first-order stochastic difference equation with

$$\begin{aligned} y_t &= P_t && (20) \\ x_t &= D_t && (21) \\ a &= \frac{1}{1+r} && (22) \\ b &= \frac{1}{1+r} && (23) \end{aligned}$$

This implies that the asset price can be expressed as follows

$$P_t = \sum_{k=0}^{N-1} \left(\frac{1}{1+r}\right)^{k+1} E_t D_{t+k} + \left(\frac{1}{1+r}\right)^{N} E_t P_{t+N} \tag{24}$$

Another assumption usually made is that this final term tends to zero as N gets big:

$$\lim_{N\to\infty} \left(\frac{1}{1+r}\right)^{N} E_t P_{t+N} = 0 \tag{25}$$

What is the logic behind this assumption? One explanation is that if it did not hold then we could set all future values of D_t equal to zero, and the asset price would still be positive. But an asset that never pays out should be inherently worthless, so this condition rules this possibility out. With this imposed, our solution becomes

$$P_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t D_{t+k} \tag{26}$$

This equation, which states that asset prices should equal a discounted present-value sum of expected future dividends, is known as the *dividend-discount model.*

Explaining the Solution Without Equations

The repeated substitution solution is really important to understand so let me try to explain it without equations. Suppose I told you that the right way to price a stock was as follows.

> Today's stock price should equal today's dividend plus half of tomorrow's expected stock price.

Now suppose it's Monday. Then that means the right formula should be

> Monday's stock price should equal Monday's dividend plus half of Tuesday's expected stock price.

It also means the following applies to Tuesday's stock price

> Tuesday's stock price should equal Tuesday's dividend plus half of Wednesday's expected stock price.

If people had rational expectations, then Monday's stock prices would equal

> Monday's dividend plus half of Tuesday's expected dividend plus one-quarter of Wednesday's expected stock price

And being consistent about it—factoring in what Wednesday's stock price should be—you'd get the price being equal to

> Monday's dividend plus half of Tuesday's expected dividend plus one-quarter of Wednesday's expected dividend plus one-eigth of Thursday's expected dividend and so on.

This is the idea being captured in equation (26).

Constant Expected Dividend Growth: The Gordon Growth Model

A useful special case that is often used as a benchmark for thinking about stock prices is the case in which dividend payments are expected to grow at a constant rate such that

$$E_t D_{t+k} = (1+g)^k D_t \tag{27}$$

In this case, the dividend-discount model predicts that the stock price should be given by

$$P_t = \frac{D_t}{1+r} \sum_{k=0}^{\infty} \left(\frac{1+g}{1+r} \right)^k \tag{28}$$

Now, remember the old multiplier formula, which states that as long as $0 < c < 1$, then

$$1 + c + c^2 + c^3 + ... = \sum_{k=0}^{\infty} c^k = \frac{1}{1-c} \tag{29}$$

This geometric series formula gets used *a lot* in modern macroeconomics, not just in examples involving the multiplier. Here we can use it as long as $\frac{1+g}{1+r} < 1$, i.e. as long as r (the expected return on the stock market) is greater than g (the growth rate of dividends). We will assume this holds. Thus, we have

$$\begin{aligned} P_t &= \frac{D_t}{1+r} \frac{1}{1 - \frac{1+g}{1+r}} && (30) \\ &= \frac{D_t}{1+r} \frac{1+r}{1+r-(1+g)} && (31) \\ &= \frac{D_t}{r-g} && (32) \end{aligned}$$

When dividend growth is expected to be constant, prices are a multiple of current dividend payments, where that multiple depends positively on the expected future growth rate of dividends and negatively on the expected future rate of return on stocks. This formula is often called the *Gordon growth model*, after the economist that popularized it.[2] It is often used as a benchmark for assessing whether an asset is above or below the "fair" value implied by rational expectations. Valuations are often expressed in terms of dividend-price ratios, and the Gordon formula says this should be

$$\frac{D_t}{P_t} = r - g \tag{33}$$

Allowing for Variations in Dividend Growth

A more flexible way to formulate expectations about future dividends is to assume that dividends fluctuate around a steady-growth trend. An example of this is the following

$$D_t = c(1+g)^t + u_t \tag{34}$$

$$u_t = \rho u_{t-1} + \epsilon_t \tag{35}$$

These equations state that dividends are the sum of two processes: The first grows at rate g each period. The second, u_t, measures a cyclical component of dividends, and this follows what is known as a first-order autoregressive process (AR(1) for short). Here ϵ_t is a zero-mean random "shock" term. Over large samples, we would expect u_t to have an average value of zero, but deviations from zero will be more persistent the higher is the value of the parameter ρ.

We will now derive the dividend-discount model's predictions for stock prices when dividends follow this process. The model predicts that

$$P_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t \left(c(1+g)^{t+k} + u_{t+k}\right) \tag{36}$$

[2] The formula appeared in Myron Gordon's 1962 book *The Investment, Financing and Valuation of the Corporation.*

Let's split this sum into two. First the trend component,

$$
\begin{aligned}
\sum_{k=0}^{\infty}\left(\frac{1}{1+r}\right)^{k+1} E_t\left(c(1+g)^{t+k}\right) &= \frac{c(1+g)^t}{1+r}\sum_{k=0}^{\infty}\left(\frac{1+g}{1+r}\right)^k && (37)\\
&= \frac{c(1+g)^t}{1+r}\frac{1}{1-\frac{1+g}{1+r}} && (38)\\
&= \frac{c(1+g)^t}{1+r}\frac{1+r}{1+r-(1+g)} && (39)\\
&= \frac{c(1+g)^t}{r-g} && (40)
\end{aligned}
$$

Second, the cyclical component. Because $E(\epsilon_{t+k}) = 0$, we have

$$
\begin{aligned}
E_t u_{t+1} &= E_t(\rho u_t + \epsilon_{t+1}) = \rho u_t && (41)\\
E_t u_{t+2} &= E_t(\rho u_{t+1} + \epsilon_{t+2}) = \rho^2 u_t && (42)\\
E_t u_{t+k} &= E_t(\rho u_{t+k-1} + \epsilon_{t+k}) = \rho^k u_t && (43)
\end{aligned}
$$

So, this second sum can be written as

$$
\begin{aligned}
\sum_{k=0}^{\infty}\left(\frac{1}{1+r}\right)^{k+1} E_t u_{t+k} &= \frac{u_t}{1+r}\sum_{k=0}^{\infty}\left(\frac{\rho}{1+r}\right)^k && (44)\\
&= \frac{u_t}{1+r}\frac{1}{1-\frac{\rho}{1+r}} && (45)\\
&= \frac{u_t}{1+r}\frac{1+r}{1+r-\rho} && (46)\\
&= \frac{u_t}{1+r-\rho} && (47)
\end{aligned}
$$

Putting these two sums together, the stock price at time t is

$$
P_t = \frac{c(1+g)^t}{r-g} + \frac{u_t}{1+r-\rho} \tag{48}
$$

In this case, stock prices don't just grow at a constant rate. Instead they depend positively on the cyclical component of dividends, u_t, and the more persistent are these cyclical deviations (the higher ρ is), the larger is their effect on stock prices. To give a concrete example, suppose $r = 0.1$. When $\rho = 0.9$ the coefficient on u_t is

$$\frac{1}{1+r-\rho} = \frac{1}{1.1-0.9} = 5 \tag{49}$$

But if $\rho = 0.6$, then the coefficient falls to

$$\frac{1}{1+r-\rho} = \frac{1}{1.1-0.6} = 2 \tag{50}$$

Note also that when taking averages over long periods of time, the u components of dividends and prices will average to zero. Thus, over longer averages the Gordon growth model would be approximately correct, even though the dividend-price ratio isn't always constant. Instead, prices would tend to be temporarily high relative to dividends during periods when dividends are expected to grow at above-average rates for a while, and would be temporarily low when dividend growth is expected to be below average for a while. This is why the Gordon formula is normally seen as a guide to long-run average valuations rather than a prediction as to what the market should be right now.

Unpredictability of Stock Returns

The dividend-discount model has some very specific predictions for how stock prices should change over time. It implies that the change in prices from period t to period $t+1$ should be

$$P_{t+1} - P_t = \sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t D_{t+k+1} - \sum_{k=0}^{\infty} \left(\frac{1}{1+r}\right)^{k+1} E_t D_{t+k} \tag{51}$$

Taking away the summation signs and writing this out in long form, it looks like this

$$\begin{aligned} P_{t+1} - P_t \quad = \quad & \left[\left(\frac{1}{1+r}\right) D_{t+1} + \left(\frac{1}{1+r}\right)^2 E_{t+1} D_{t+2} + \left(\frac{1}{1+r}\right)^3 E_{t+1} D_{t+3} + ...\right] \\ & - \left[\left(\frac{1}{1+r}\right) D_t + \left(\frac{1}{1+r}\right)^2 E_t D_{t+1} + \left(\frac{1}{1+r}\right)^3 E_t D_{t+2} + ...\right] \end{aligned} \tag{52}$$

We can re-arrange this equation in a useful way by grouping together each of the two terms that involve D_{t+1}, D_{t+2}, D_{t+3} and so on. (There is only one term involving D_t.) This can be written as follows

$$\begin{aligned} P_{t+1} - P_t \quad = \quad & -\left(\frac{1}{1+r}\right) D_t + \left[\left(\frac{1}{1+r}\right) D_{t+1} - \left(\frac{1}{1+r}\right)^2 E_t D_{t+1}\right] \\ & + \left[\left(\frac{1}{1+r}\right)^2 E_{t+1} D_{t+2} - \left(\frac{1}{1+r}\right)^3 E_t D_{t+2}\right] + \\ & + \left[\left(\frac{1}{1+r}\right)^3 E_{t+1} D_{t+3} - \left(\frac{1}{1+r}\right)^4 E_t D_{t+3}\right] + ... \end{aligned} \tag{53}$$

This equation explains three reasons why prices change from period P_t to period P_{t+1}.

- P_{t+1} differs from P_t because it does not take into account D_t – this dividend has been paid now and has no influence any longer on the price at time $t+1$. This is the first term on the right-hand side above.

- P_{t+1} applies a smaller discount rate to future dividends because have moved forward one period in time, e.g. it discounts D_{t+1} by $\left(\frac{1}{1+r}\right)$ instead of $\left(\frac{1}{1+r}\right)^2$.

- People formulate new expectations for the future path of dividends e.g. $E_t D_{t+2}$ is gone and has been replaced by $E_{t+1} D_{t+2}$

In general, the first few items above should not be too important. A single dividend payment being made shouldn't have too much impact on a stock's price and the discount rate shouldn't change too much over a single period (e.g. if r is relatively small, then $\left(\frac{1}{1+r}\right)$ and $\left(\frac{1}{1+r}\right)^2$ shouldn't be too different.) This means that changing expectations about future dividends should be the main factor driving changes in stock prices.

In fact, it turns out there is a very specific result linking the behaviour of stock prices with changing expectations. Ultimately, it is not stock prices, *per se*, that investors are interested in. Rather, they are interested in the combined return incorporating both price changes and dividend payments, as described by equation (4). It turns out that movements in stock returns are entirely driven by changes in dividend expectations.

With a number of lines of algebra (described in an appendix) equation (53) can be re-expressed as

$$P_{t+1} - P_t = -D_t + rP_t + \sum_{k=1}^{\infty} \left[\left(\frac{1}{1+r} \right)^k (E_{t+1}D_{t+k} - E_t D_{t+k}) \right] \tag{54}$$

Recalling the definition of the one-period return on a stock from equation (1), this return can be written as

$$r_{t+1} = \frac{D_t + \Delta P_{t+1}}{P_t} = r + \frac{\sum_{k=1}^{\infty} \left(\frac{1}{1+r} \right)^k (E_{t+1}D_{t+k} - E_t D_{t+k})}{P_t} \tag{55}$$

This is a very important result. It tells us that, if the dividend-discount model is correct, then the rate of return on stocks depends on how people change their minds about what they expect to happen to dividends in the future: The $E_{t+1}D_{t+k} - E_t D_{t+k}$ terms on the right-hand side of equation (55) describe the difference between what people expected at time $t+1$ for the dividend at time $t+k$ and what they expected for this same dividend payment at time t.

Importantly, if we assume that people formulate rational expectations, then the return on stocks should be unpredictable. This is because, if we could tell in advance how people were going to change their expectations of future events, then that would mean people have not been using information in an efficient manner. So, with rational expectations, the term in the summation sign in equation (55) must be zero on average and must reflect "news" that could not have been forecasted at time t. So the only thing determining changes in stock returns in the innately unforecastable process of people incorporating completely new information.

One small warning about this result. It is often mis-understood as a prediction that *stock* *prices* (rather than stock returns) should be unpredictable. This is not the case. The series that should be unpredictable is the total stock return including the dividend payment. Indeed, the model predicts that a high dividend payments at time t lowers stock prices at time $t+1$. Consider for example a firm that promises to make a huge dividend payment next month but says they won't make any payments after this for a long time. In that case, we would expect the price of the stock to fall after the dividend is payment. This shows that, even with rational expectations, stock prices movements can sometimes be predictable. Because dividend payments are only made on an occasional basis, this prediction can be tested and various studies have indeed found so-called "ex-dividend" effects whereby a stock price falls after a dividend is paid.

Evidence on Predictability and "Efficient Markets"

The theoretical result that stock returns should be unpredictable was tested in a series of empirical papers in the 1960s and 1970s, most notably by University of Chicago professor Eugene Fama and his co-authors. The website contains a link to Fama's famous 1970 paper "Efficient Capital Markets: A Review of Theory and Empirical Work". This literature came to a clear conclusion that stock returns did seem to be essentially unpredictable. The idea that you could not make easy money by "timing the market" entered public discussion with Burton Malkiel's famous 1973 book, *A Random Walk Down Wall Street* being particularly influential. A "random walk" is a series whose changes cannot be forecasted and rational expectations implies that changes in the cumulative return on a stock is unforecastable.

The work of Fama and his co-authors was very important in establishing key facts about how financial markets work. One downside to this research, though, was the introduction of a terminology that proved confusing. Fama's 1970 paper describes financial market as being "efficient" if they "fully reflect all available information." In general, the researchers contributing to this literature concluded financial markets were efficient because stock returns were difficult to forecast. However, this turned out to be a bit of a leap. It is certainly true that if stock prices incorporate all available information in the rational manner described above, then returns should be hard to forecast. But the converse doesn't necessarily apply: Showing that it was difficult to forecasting stock returns turned out to not be the same thing as proving that stock markets were efficient.

Robert Shiller on Excess Volatility

The idea that financial markets were basically efficient was widely accepted in the economics profession by the late 1970s. Then, a Yale economist in his mid-thirties, Robert Shiller, dropped something of a bombshell on the finance profession. Shiller showed that the dividend-discount model beloved of finance academics completely failed to match the observed *volatility* of stock prices.[3] Specifically, stock prices were much more volatile than could be justified by the model.

To understand Shiller's basic point, we need to take a step back and think about some basic concepts relating to the formulation of expectations. First note that the *ex post* outcome for any variable can be expressed as the sum of its *ex ante* value expected by somebody and the unexpected component (i.e. the amount by which that person's expectation was wrong).

[3] "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review*, June 1981

This can be written in a formula as

$$X_t = E_{t-1}X_t + \epsilon_t \tag{56}$$

From statistics, we know that the variance of the sum of two variables equals the sum of their two variances plus twice their covariance. This means that the variance of X_t can be described by

$$\text{Var}\left(X_t\right) = \text{Var}\left(E_{t-1}X_t\right) + \text{Var}\left(\epsilon_t\right) + 2\text{Cov}\left(E_{t-1}X_t, \epsilon_t\right) \tag{57}$$

Now note that this last covariance term—between the "surprise" element ϵ_t and the ex-ante expectation $E_{t-1}X_t$—should equal zero if expectations are fully rational. If there was a correlation—for instance, so that a low value of the expectation tended to imply a high value for the error—then this would means that you could systematically construct a better forecast once you had seen the forecast that was provided. For example, if a low forecasted value tended to imply a positive error then you could construct a better forecast by going for a higher figure. But this contradicts the idea that investors have rational expectations and thus use all information efficiently.

So, if expectations are rational, then we have

$$\text{Var}\left(X_t\right) = \text{Var}\left(E_{t-1}X_t\right) + \text{Var}\left(\epsilon_t\right) \tag{58}$$

The variance of the observed series must equal the variance of the *ex ante* expectation plus the variance of the unexpected component. Provided there is uncertainty, so there is some unexpected component, then we must have

$$\text{Var}\left(X_t\right) > \text{Var}\left(E_{t-1}X_t\right) \tag{59}$$

In other words, the variance of the *ex post* outcome should be higher than the variance of *ex ante* rational expectation.

This reasoning has implications for the predicted volatility of stock prices. Equation (26) says that stock prices are an *ex ante* expectation of a discount sum of future dividends. Shiller's observation was that rational expectations should imply that the variance of stock prices be less than the variance of the present value of subsequent dividend movements:

$$\mathrm{Var}(P_t) < \mathrm{Var}\left[\sum_{k=0}^{\infty}\left(\frac{1}{1+r}\right)^{k+1} D_{t+k}\right] \tag{60}$$

A check on this calculation, using a wide range of possible values for r, reveals that this inequality does not hold: Stocks are actually much more volatile than suggested by realized movements in dividends.[4]

Figure 1 on the next page reproduces the famous graph from Shiller's 1981 paper showing actual stock prices (the sold line) moving around much more over time than his "discounted outcome of dividends" series (the dashed line). Figure 2 shows an updated version of his chart, presented by Shiller as part of his 2014 Nobel acceptance talk.

[4]While technically, the infinite sum of dividends can't be calculated because we don't have data going past the present, Shiller filled in all terms after the end of his sample based on plausible assumptions, and the results are not sensitive to these assumptions.

Figure 1: Shiller's 1981 Chart Illustrating Excess Volatility

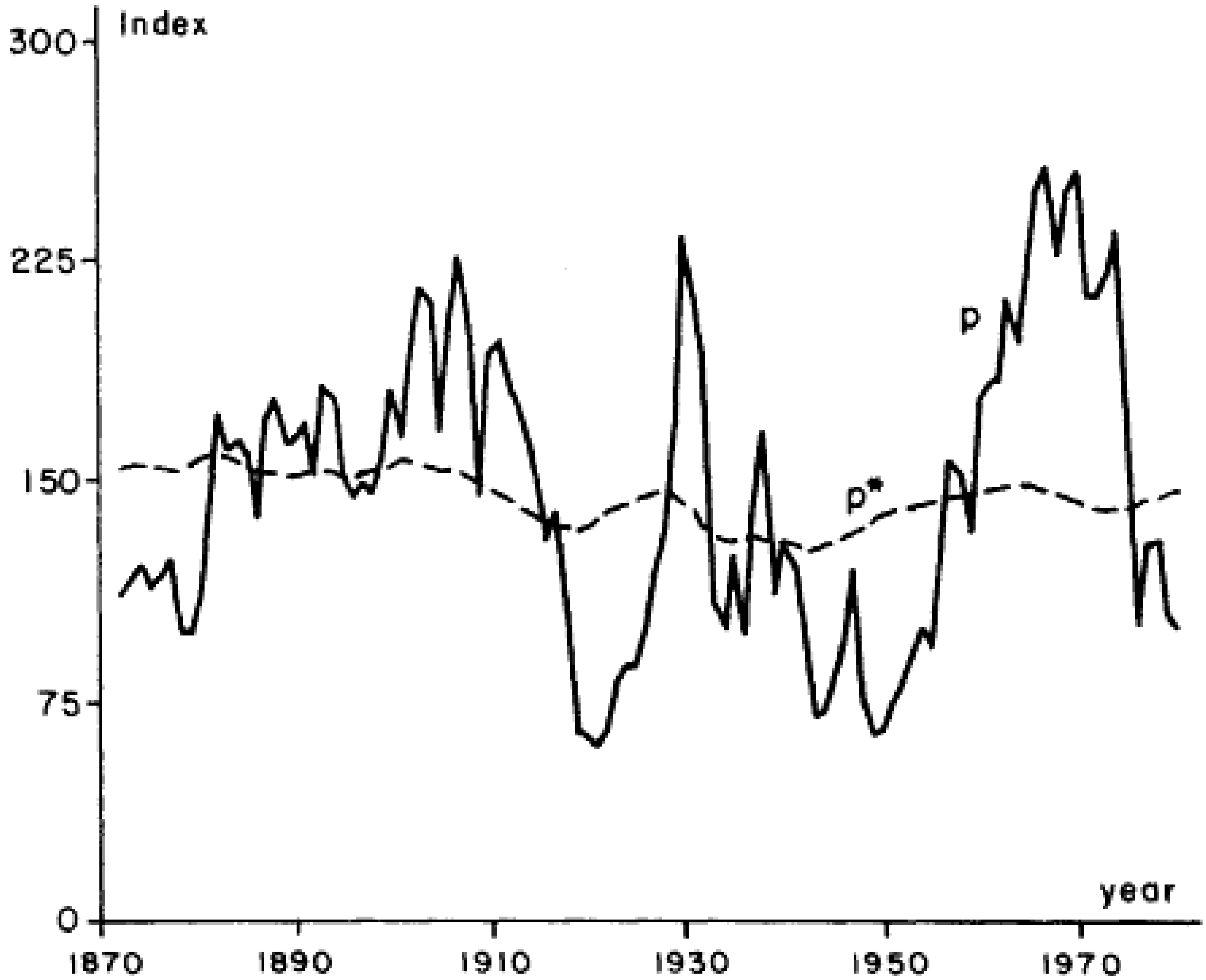
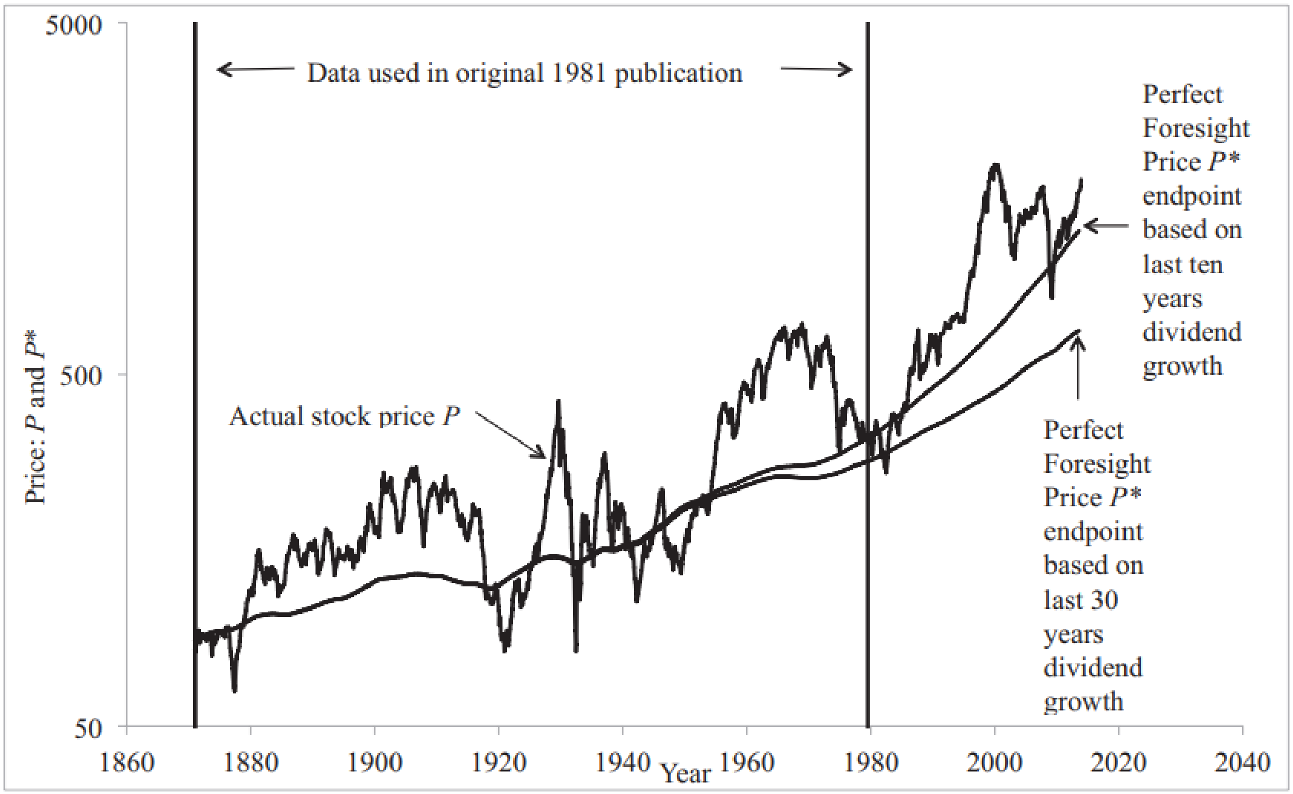


FIGURE 1

Note: Real Standard and Poor's Composite Stock Price Index (solid line p) and *ex post* rational price (dotted line p^*), 1871–1979, both detrended by dividing a long-run exponential growth factor. The variable p^* is the present value of actual subsequent real detrended dividends, subject to an assumption about the present value in 1979 of dividends thereafter. Data are from Data Set 1, Appendix.

Figure 2: Shiller's Updated 2014 Chart



Longer-Run Predictability

We saw earlier that the dividend-discount model predicts that when the ratio of dividends to prices is low, this suggests that investors are confident about future dividend growth. Thus, a low dividend-price ratio should help to predict higher future dividend growth. Shiller's volatility research pointed out, however, that there appears to be a lot of movements in stock prices that never turn out to be fully justified by later changes dividends. In fact, later research went a good bit further. For example, Campbell and Shiller (2001) show that over longer periods, dividend-price ratios are of essentially no use at all in forecasting future dividend growth.[5] In fact, a high ratio of prices to dividends, instead of forecasting high growth in dividends, tends to forecast lower future returns on the stock market albeit with a relatively low R-squared. See Figure 3.

This last finding seems to contradict Fama's earlier conclusions that it was difficult to forecast stock returns but these results turn out to be compatible with both those findings and the volatility results. Fama's classic results on predictability focused on explaining short-run stock returns e.g. can we use data from this year to forecast next month's stock returns? However, the form of predictability found by Campbell and Shiller (and indeed a number of earlier studies) related to predicting average returns over multiple years. It turns out an inability to do find short-run predictability is not the same thing as an inability to find longer-run predictability.

To understand this, we need to develop some ideas about forecasting time series. Consider a series that follows the following $AR(1)$ time series process:

$$y_t = \rho y_{t-1} + \epsilon_t \tag{61}$$

[5] NBER Working Paper No. 8221.

where ϵ_t is a random and unpredictable "noise" process with a zero mean. If $\rho = 1$ then the change in the series is

$$y_t - y_{t-1} = \epsilon_t \tag{62}$$

so the series is what is what we described earlier as a random walk process whose changes cannot be predicted. Suppose, however, that ρ was close to but a bit less than one, say $\rho = 0.99$. The change in the series would now be given by

$$y_t - y_{t-1} = -0.01 y_{t-1} + \epsilon_t \tag{63}$$

Now suppose you wanted to assess whether you could forecast the change in the series based on last period's value of the series. You could run a regression of the change in y_t on last period's value of the series. The true coefficient in this relationship is -0.01 with the ϵ_t being the random error. This coefficient of -0.01 is so close to zero that you will probably be unable to reject that the true coefficient is zero unless you have far more data than economists usually have access to.

But what if you were looking at forecasting changes in the series over a longer time-horizon? To understand why this might be different, we can do another repeated substitution trick. The series y_t depends on its lagged value, y_{t-1} and a random shock. But y_{t-1} in turn depended on y_{t-2} and another random shock. And y_{t-2} in turn depended on y_{t-3} and another random shock. And so on. Plugging in all of these substitutions you get the following.

$$\begin{aligned} y_t &= \rho y_{t-1} + \epsilon_t \\ &= \rho^2 y_{t-2} + \epsilon_t + \rho \epsilon_{t-1} \\ &= \rho^3 y_{t-3} + \epsilon_t + \rho \epsilon_{t-1} + \rho^2 \epsilon_{t-2} \\ &= \rho^N y_{t-N} + \sum_{k=0}^{N-1} \rho^k \epsilon_{t-k} \end{aligned} \tag{64}$$

Now suppose you wanted to forecast the change in y_t over N periods with the value of the series from N periods ago. This change can be written as

$$y_t - y_{t-N} = \left(\rho^N - 1\right) y_{t-N} + \sum_{k=0}^{N-1} \rho^k \epsilon_{t-k} \tag{65}$$

Again the change in y_t over this period can be written as a function of a past value of the series and some random noise. The difference in this case is that the coefficient on the lagged value doesn't have to be small anymore even if had a near-random walk series. For example, suppose $\rho = 0.99$ and $N = 50$ so were looking at the change in the series over 50 periods. In this case, the coefficient is $(0.99^{50} - 1) = -0.4$. For this reason, regressions that seek to predict combined returns over longer periods have found statistically significant evidence of predictability even though this evidence cannot be found for predicting returns over shorter periods.

It is very easy to demonstrate this result using any software that can generate random numbers. For example, in an appendix at the back of the notes I provide a short programme written for the econometric package RATS. The code is pretty intuitive and could be repeated for lots of other packages. The programme generates random $AR(1)$ series with $\rho = 0.99$ by starting them off with a value of zero and then drawing random errors to generate full time series. Then regressions for sample sizes of 200 are run to see if changes in the series over one period, twenty periods and fifty periods can be forecasted by the relevant lagged values. This is done 10,000 times and the average t-statistics from these regressions are calculated.

Table 1 shows the results. The average t-statistic for the one-period forecasting regression is -1.24, not high enough to reject the null hypothesis that there is no forecasting power. In contrast, the average t-statistic for the 20-period forecasting regression is -5.69 and the average t-statistic for the 50-period forecasting regression is -8.95, so you can be very confident that there is statistically significant forecasting power over these horizons.

The intuition behind these results is fairly simple. Provided ρ is less than one in absolute value, $AR(1)$ series are what is known as mean-stationary. In other words, they tend to revert back to their average value. In the case of the y_t series here, this average value is zero. The speed at which you can expect them to return to this average value will be slow if ρ is high but they will eventually return. So if you see a high value of y_t, you can't really be that confident that it will fall next period but you can be very confident that it will eventually tend to fall back towards its average value of zero.

Pulling these ideas together to explain the various stock price results, suppose prices were given by

$$P_t = \sum_{k=0}^{\infty} \left[\left(\frac{1}{1+r} \right)^{k+1} E_t D_{t+k} \right] + u_t \tag{66}$$

where

$$u_t = \rho u_{t-1} + \epsilon_t \tag{67}$$

with ρ being close to one and ϵ_t being an unpredictable noise series. This model says that stock prices are determined by two elements. The first is the rational dividend-discount price and the second is a non-fundamental $AR(1)$ element reflecting non-rational market sentiment. The latter could swing up and down over time as various fads and manias affect the market.

In this case, statistical research would generate three results:

1. Short-term stock returns would be very hard to forecast. This is partly because of the rational dividend-discount element but also because changes in the non-fundamental element are hard to forecast over short-horizons.

2. Longer-term stock returns would have a statistically significant forecastable element,

though with a relatively low R-squared. This is because the fundamental element that accounts for much of the variation cannot be forecasted while you can detect a statistically significant forecastable element for the non-fundamental component.

3. Stock prices would be more volatile than predicted by the dividend-discount model, perhaps significantly so. This is because non-fundamental series of the type described here can go through pretty long swings which adds a lot more volatility than the dividend-discount model would predict.

This suggests a possible explanation for the behaviour of stock prices. On average, they appear to be determined by something like the dividend-discount model but they also have a non-fundamental component that sees the market go through temporary (but potentially long) swings in which it moves away from the values predicted by this model.

Figure 3: Campbell and Shiller's 2001 Chart

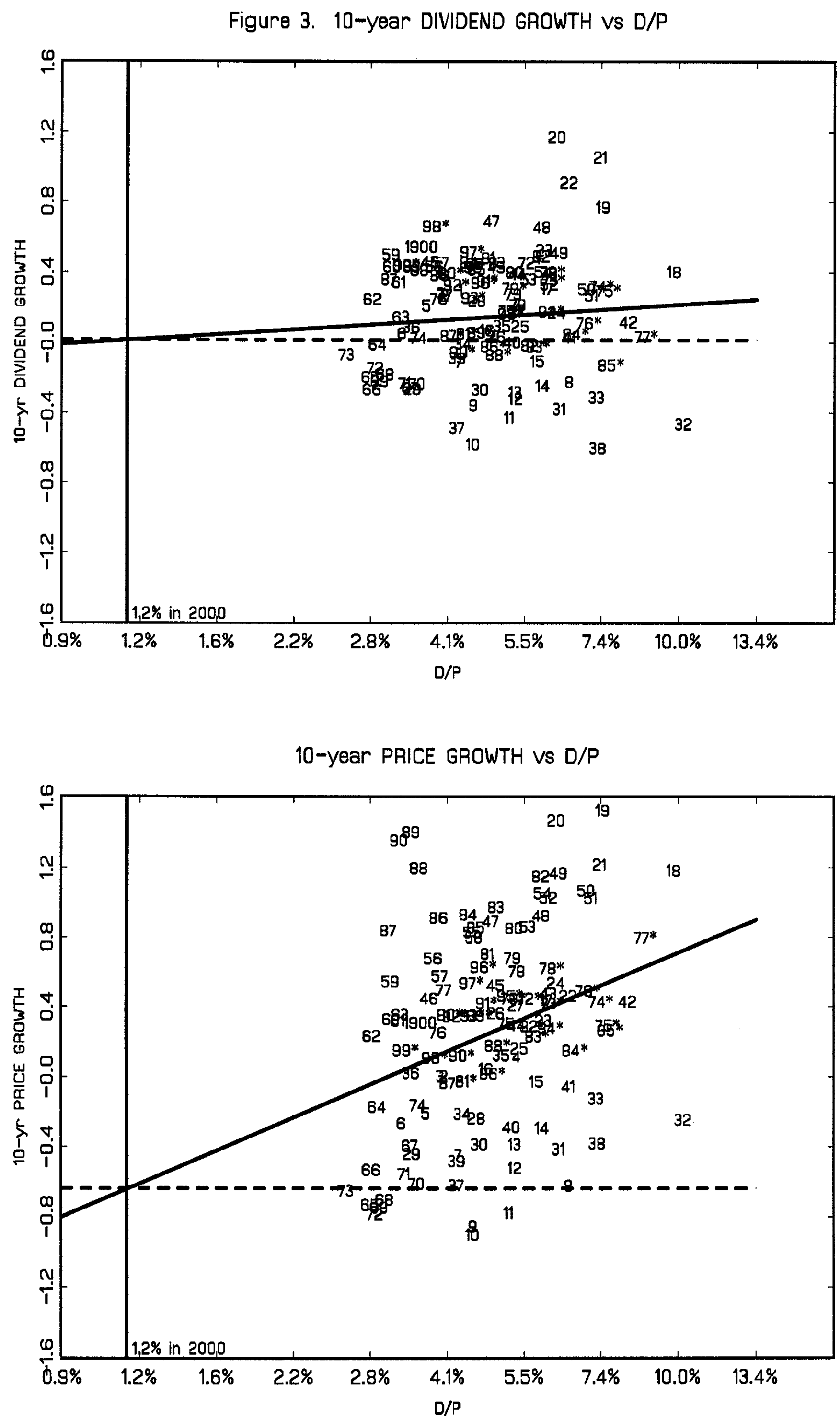


Table 1: Illustrating Long-Run Predictability

```
Statistics on Series TSTATS_1LAG
Observations              10000
Sample Mean           -1.244998      Variance             0.463179
Standard Error         0.680573      of Sample Mean       0.006806
t-Statistic (Mean=0) -182.933896     Signif Level         0.000000
Skewness              -0.123331      Signif Level (Sk=0)  0.000000
Kurtosis (excess)      0.512791      Signif Level (Ku=0)  0.000000
Jarque-Bera          134.915363      Signif Level (JB=0)  0.000000

Statistics on Series TSTATS_20LAG
Observations              10000
Sample Mean           -5.693510      Variance             11.267177
Standard Error         3.356662      of Sample Mean       0.033567
t-Statistic (Mean=0) -169.618235     Signif Level         0.000000
Skewness              -0.277797      Signif Level (Sk=0)  0.000000
Kurtosis (excess)      0.778124      Signif Level (Ku=0)  0.000000
Jarque-Bera          380.901249      Signif Level (JB=0)  0.000000

Statistics on Series TSTATS_50LAG
Observations              10000
Sample Mean           -8.951905      Variance             31.554735
Standard Error         5.617360      of Sample Mean       0.056174
t-Statistic (Mean=0) -159.361419     Signif Level         0.000000
Skewness              -0.338830      Signif Level (Sk=0)  0.000000
Kurtosis (excess)      0.926286      Signif Level (Ku=0)  0.000000
Jarque-Bera          548.845975      Signif Level (JB=0)  0.000000
```

Example: The Prognosis for U.S. Stock Prices

These results would suggest that it may be possible to detect whether a stock market is over-valued or under-valued and thus to forecast its future path. Let's consider an example and look at the current state of the U.S. stock market as measured by the S&P 500 which is a broad measure of large-capitalisation stocks.

Figure 4 shows that the U.S. market has been on a tear over the past decade and has moved well past previous historical highs. The last two times the market expanded rapidly ended in big crashes. Will this end the same way?

Figure 5 shows the ratio of dividends to prices for the S&P 500 index of US stocks over the period since the second world war. The measure of dividends used in the numerator is based on the average value of dividends over a twelve month period to smooth out volatile month-to-month movements. The chart shows that the dividend-price ratio for this index has been very low by historical standards in recent years. In other words, prices are very high relative to dividends, which is normally a bad sign.

Still, comparisons with the long-run historical average might be a bad idea. The average value of this ratio over the period since 1945 is 3.2 percent. The only point since the mid-1990s that the ratio has exceeded that historical average was a brief period in early 2009 due to the plunge in stock prices after the Lehman Brothers bankruptcy and the emergence of the worst recession since the Great Depression.

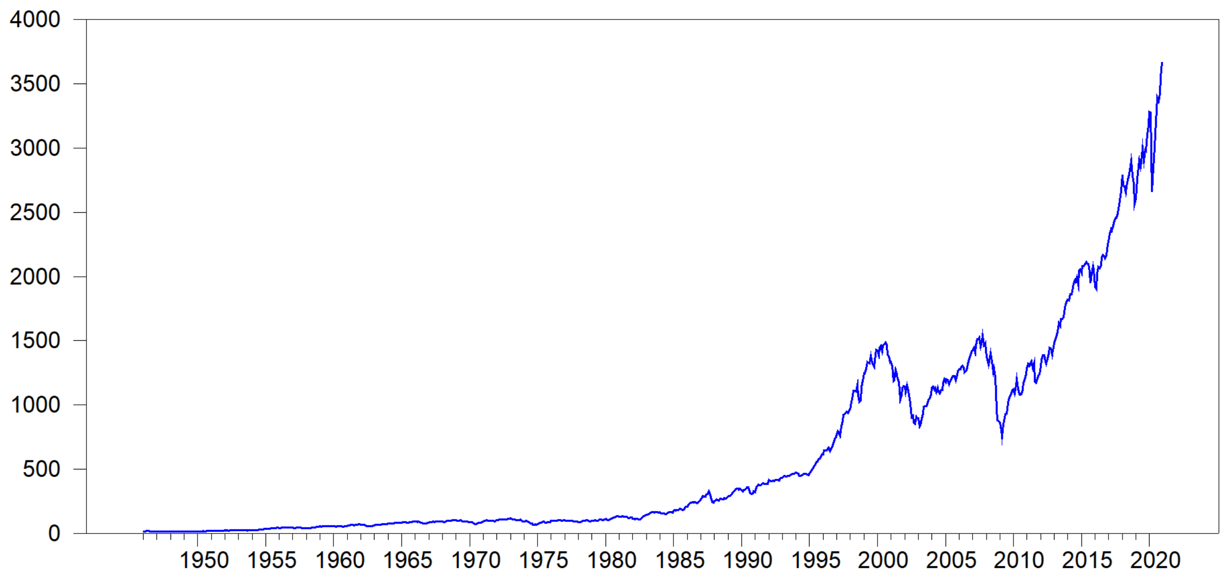
One reason for this change is that many firms have moved away from paying dividends. In more recent years, it is particularly clear that dividends have been low relative to how much companies can afford to pay. The ratio of dividends to earnings for S&P 500 firms is about half its historical level. There are two reasons for this. There has been a long-run trend of moving away from paying dividends and towards using earnings to fund share repurchases. This is a way to return money to shareholders and increase the value of the remaining shares (each of them can get a higher share of future dividend payments) without the shareholders explicitly receiving dividend income at present, which would be taxable at a higher rate than capital gains.

Because of these factors, many analysts instead look at the ratio of total corporate earnings to prices. Figure 6 compares this to the dividend-price ratio. This ratio of average earnings over the previous twelve months to prices was about 3.4 percent in September 2020. This is well below the historical average of about 6.7 percent so this series suggests that stocks are perhaps a bit over-valued relative to historical norms.

A final consideration, however, is the discount rate being used to value stocks. We know from the Gordon growth model that one reason stock prices might be high relative to dividends (or earnings) is that the expected rate of return r may be low. Assuming the required rate of return on stocks reflects some premium over safe investments such government bonds, this could provide another explanation for high stock price valuations. Figure 7 shows that real interest rates on US Treasury bonds are at historically low levels (this series is the yield on ten-year Treasury bonds minus inflation over the previous year). If these rates are being used as a benchmark for calculating the required rate of return on stocks, then one might expect stock prices to be high relative to dividends.

This shows that figuring out whether stocks are under- or over-valued is rarely as easy as examining one specific metric.

Figure 4: The S&P 500 Index of U.S. Stock Prices



Time-Varying Expected Returns

I have suggested an alternative explanation of the facts to the dividend-discount model — one in which stock prices are also determined by a temporary but volatile non-fundamental component. However, the last observation in the previous section — about discount rates — suggests another way to "mend" the dividend-discount model and perhaps explain the extra volatility that affects stock prices: Change the model to allow for variations in expected returns. Consider the finding that a high value of the dividend-price ratio predicts poor future stock returns. Shiller suggests that this is due to temporary irrational factors gradually disappearing. But another possibility is that that the high value of this ratio is rationally anticipating low future returns.

We can reformulate the dividend-discount model with time-varying returns as follows. Let

$$R_t = 1 + r_t \tag{68}$$

Figure 5: Dividend-Price Ratio for S&P 500

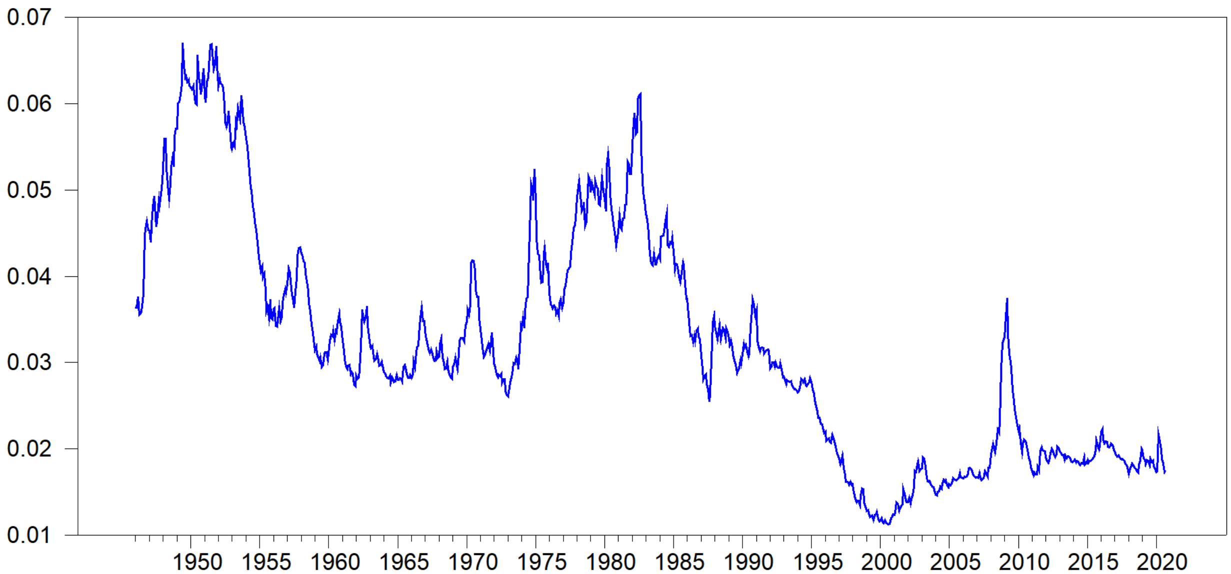


Figure 6: Dividend-Price and Earnings-Price Ratios for S&P 500

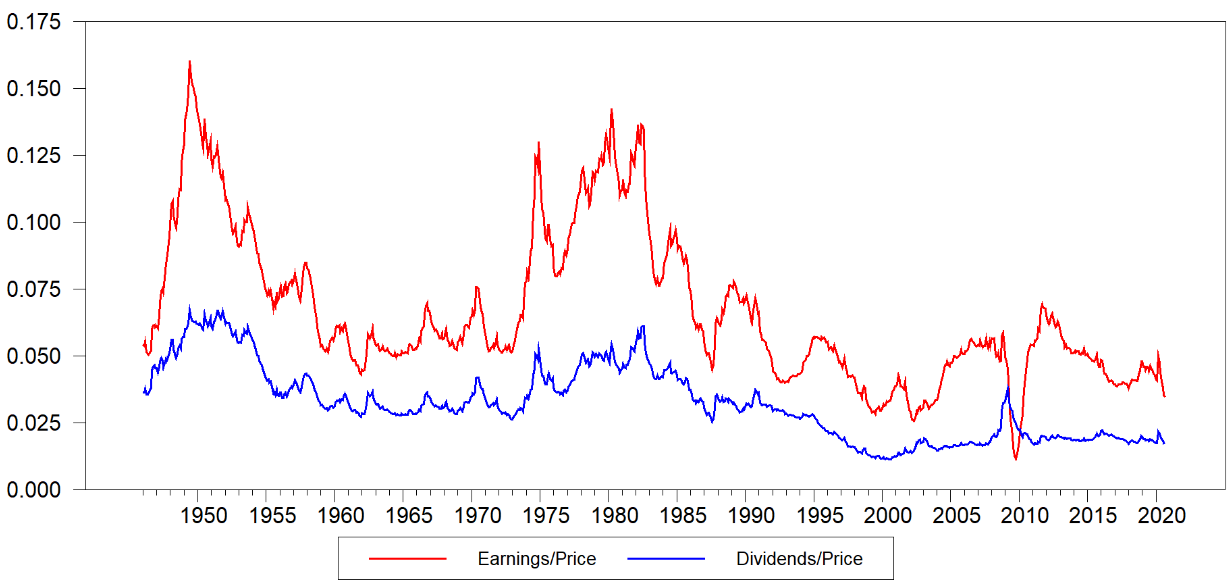
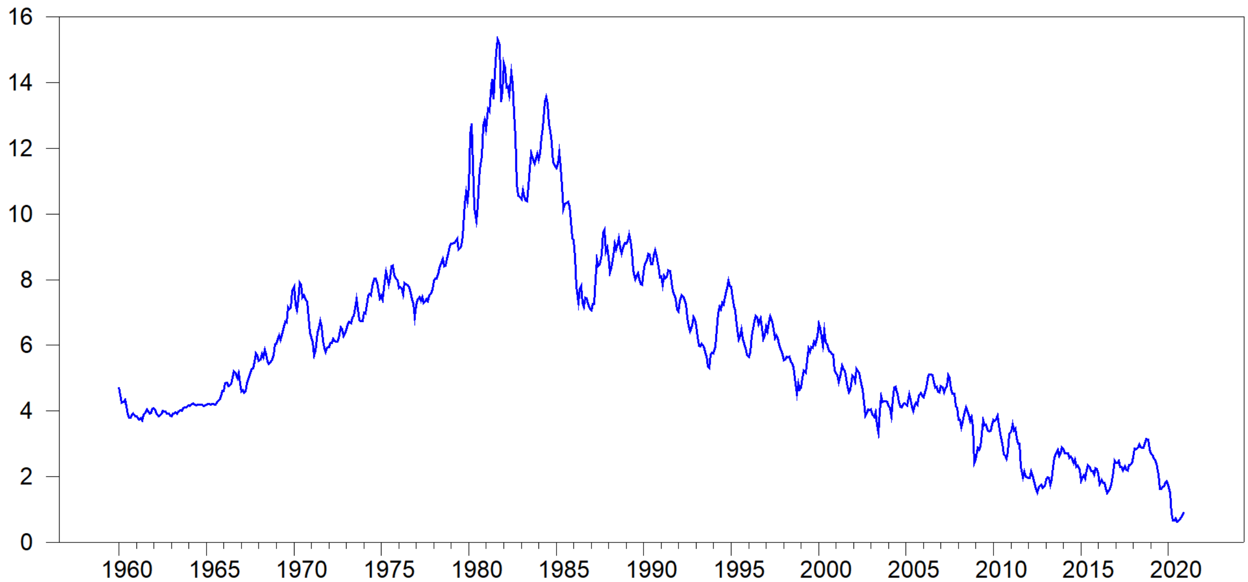


Figure 7: Real 10-Year Treasury Bond Rate



Start again from the first-order difference equation for stock prices

$$P_t = \frac{D_t}{R_{t+1}} + \frac{P_{t+1}}{R_{t+1}} \tag{69}$$

where R_{t+1} is the return on stocks in period $t+1$. Moving the time-subscripts forward one period, this implies

$$P_{t+1} = \frac{D_{t+1}}{R_{t+2}} + \frac{P_{t+2}}{R_{t+2}} \tag{70}$$

Substitute this into the original price equation to get

$$\begin{aligned} P_t &= \frac{D_t}{R_{t+1}} + \frac{1}{R_{t+1}} \left(\frac{D_{t+1}}{R_{t+2}} + \frac{P_{t+2}}{R_{t+2}} \right) \\ &= \frac{D_t}{R_{t+1}} + \frac{D_{t+1}}{R_{t+1}R_{t+2}} + \frac{P_{t+2}}{R_{t+1}R_{t+2}} \end{aligned} \tag{71}$$

Applying the same trick to substitute for P_{t+2} we get

$$P_t = \frac{D_t}{R_{t+1}} + \frac{D_{t+1}}{R_{t+1}R_{t+2}} + \frac{D_{t+2}}{R_{t+1}R_{t+2}R_{t+3}} + \frac{P_{t+3}}{R_{t+1}R_{t+2}R_{t+3}} \tag{72}$$

The general formula is

$$P_t = \sum_{k=0}^{N-1} \left(\frac{D_{t+k}}{\prod_{m=1}^{k+1} R_{t+m}} \right) + \frac{P_{t+N}}{\prod_{m=1}^{N} R_{t+m}} \tag{73}$$

where $\prod_{n=1}^{h} x_i$ means the product of x_1, x_2 .... x_h. Again setting the limit of the $t+N$ term to zero and taking expectations, we get a version of the dividend-discount model augmented to account for variations in the expected rate of return.

$$P_t = \sum_{k=0}^{\infty} E_t \left(\frac{D_{t+k}}{\prod_{m=1}^{k+1} R_{t+m}} \right) \tag{74}$$

This equation gives a potential explanation for the failure of news about dividends to explain stock price fluctuations. Stock prices depend positively on expected future dividends. But they also depend negatively on the R_{t+k} values which measure the expected future return on stocks. So perhaps news about future stock returns explains movements in stock prices: When investors learn that future returns are going to be lower, this raises current stock prices.

The website provides a link to a 1991 paper by Eugene Fama which provide his updated overview of the literature on the predictability of stock returns. By this point, Fama accepted the evidence on long-horizon predictability and had contributed to this literature. However, Fama and French (1988) put forward predictable time-variation in expected returns as the likely explanation for this result.[6] This explanation has also been promoted by leading modern finance economists such as John Campbell and John Cochrane and is currently the leading hypothesis for reconciling the evidence on stock prices movements with rational expectations.[7]

[6] Eugene Fama and Kenneth French (1988). "Dividend Yields and Expected Stock Returns" *Journal of Financial Economics*, Volume 22, pages 3-25.

[7] Campbell's 1991 paper, "A Variance Decomposition for Stock Returns," *Economic Journal*, provides a nice framework for understanding this theory.

What About Interest Rates?

Changing interest rates on bonds are the most obvious source of changes in expected returns on stocks. Up to now, we only briefly discussed what determines the rate of return that investors require to invest in the stock market, but it is usually assumed that there is an arbitrage equation linking stock and bond returns, so that

$$E_t r_{t+1} = E_t i_{t+1} + \pi \tag{75}$$

In other words, next period's expected return on the market needs to equal next period's expected interest rate on bonds, i_{t+1}, plus a risk premium, π, which we will assume is constant.

Are interest rates the culprit accounting for the volatility of stock prices? They are certainly a plausible candidate. Stock market participants spend a lot of time monitoring the Fed and the ECB and news interpreted as implying higher interest rates in the future certainly tends to provoke declines in stock prices. Perhaps surprisingly, then, Campbell and Shiller (1988) showed that this type of equation still doesn't help that much in explaining stock market fluctuations.[8] Their methodology involved plugging in forecasts for future interest rates and dividend growth into the right-hand-side of (74) and checking how close the resulting series is to the actual dividend-price ratio. They concluded that expected fluctuations in interest rates contribute little to explaining the volatility in stock prices. A study co-authored by Federal Reserve Chairman Ben Bernanke examining the link between monetary policy and the stock market came to the same conclusions.[9]

[8] "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, Autumn 1988.

[9] See "What Explains the Stock Market's Reaction to Federal Reserve Policy?", by Ben Bernanke and Kenneth Kuttner, *Journal of Finance*, June 2005.

Time-Varying Risk Premia or Behavioural Finance?

So, changes in interest rates do not appear to explain the volatility of stock market fluctuations. The final possible explanation for how the dividend-discount model may be consistent with the data is that changes in expected returns do account for the bulk of stock market movements, but that the principal source of these changes comes, not from interest rates, but from changes in the *risk premium* that determines the excess return that stocks must generate relative to bonds: The π in equation (75) must be changing over time. According to this explanation, asset price booms are often driven by investors being willing to take risks and receive a relatively low compensation for them (when investors are "risk-on" in the commonly-used market terminology) while busts often happen when investors start to demand higher risk premia (when they are "risk-off").

One problem with this conclusion is that it implies that, most of the time, when stocks are increasing it is because investors are anticipating lower stock returns at a later date. However, the evidence that we have on this seems to point in the other direction. For example, surveys have shown that even at the peak of the most recent bull market, average investors still anticipate high future returns on the market.

If one rejects the idea that, together, news about dividends and news about future returns explain all of the changes in stock prices, then one is forced to reject the rational expectations dividend-discount model as a complete model of the stock market. What is missing from this model? Many believe that the model fails to take into account of various human behavioural traits that lead people to act in a manner inconsistent with pure rational expectations. Economists like Shiller point to the various asset price "bubbles" of the past twenty years — such as the dot-com boom and bust and the rise and fall in house prices in countries like the U.S. and Ireland, as clear evidence that investors go through periods of "irrational exhuberance" which sees asset prices become completely detached from the fundamental values suggested by reasonable applications of the dividend-discount model.

Indeed, the inability to reconcile aggregate stock price movements with rational expectations is not the only well-known failure of modern financial economics. For instance, there are many studies documenting the failure of rational optimisation-based models to explain various cross-sectional patterns in asset returns, e.g. why the average return on stocks exceeds that on bonds by so much, or discrepancies in the long-run performance of small- and large-capitalisation stocks. Eugene Fama is the author of a number of famous papers with Kenneth French that have demonstrated these discrepancies though he interprets these results as most likely due to a rational pricing of the risk associated with certain kinds of assets.

For many, the answers to these questions lie in abandoning the pure rational expectations, optimising approach. Indeed, the field of *behavioural finance* is booming, with various researchers proposing all sorts of different non-optimising models of what determines asset prices. That said, at present, there is no clear front-runner "alternative" behavioural-finance model of the determination of aggregate stock prices.

Things to Understand from these Notes

Here's a brief summary of the things that you need to understand from these notes.

1. The meaning of rational expectations, as used by economists.
2. The repeated substitution method for solving first-order stochastic difference equations.
3. How to derive the dividend-discount formula.
4. How to derive the Gordon growth formula and the variant with cyclical dividends.
5. The dividend-discount model's predictions on predictability.
6. Fama's evidence and the meaning of efficient markerts.
7. The logic behind Robert Shiller's test of the dividend-discount model and his findings.
8. How the dividend-price ratio forecasts longer-horizon stock returns.
9. Why stock returns may be predictable over a longer horizon but not a shorter one.
10. How to incorporate time-varying expected returns, interest rates and risk premia.
11. The state of debate about rational expectations and asset pricing.

Appendix 1: Proof of Equation (54)

We start by repeating equation (53):

$$\begin{aligned} P_{t+1} - P_t = \; &= -\left(\frac{1}{1+r}\right) D_t + \left[\left(\frac{1}{1+r}\right) D_{t+1} - \left(\frac{1}{1+r}\right)^2 E_t D_{t+1}\right] \\ &+ \left[\left(\frac{1}{1+r}\right)^2 E_{t+1} D_{t+2} - \left(\frac{1}{1+r}\right)^3 E_t D_{t+2}\right] + \\ &+ \left[\left(\frac{1}{1+r}\right)^3 E_{t+1} D_{t+3} - \left(\frac{1}{1+r}\right)^4 E_t D_{t+3}\right] + ... \end{aligned}$$

This shows that the change in stock prices is determined by a term relating to this period's dividend's "dropping out" and then a whole bunch of terms that involve period $t+1$ and period t expectations of future dividends. To be able to pull all the terms for each D_{t+k} together, we both add and subtract a set of terms of the form $\left(\frac{1}{1+r}\right)^k E_t D_{t+k}$. The equation then looks like this

$$\begin{aligned} P_{t+1} - P_t = \; &= \left(\frac{1}{1+r}\right) [D_{t+1} - E_t D_{t+1}] \\ &+ \left(\frac{1}{1+r}\right)^2 [E_{t+1} D_{t+2} - E_t D_{t+2}] \\ &+ \left(\frac{1}{1+r}\right)^3 [E_{t+1} D_{t+3} - E_t D_{t+3}] + ... \\ &- \left(\frac{1}{1+r}\right) D_t \\ &+ \left(1 - \frac{1}{1+r}\right)\left(\frac{1}{1+r}\right) E_t D_{t+1} \\ &+ \left(1 - \frac{1}{1+r}\right)\left(\frac{1}{1+r}\right)^2 E_t D_{t+2} \\ &+ \left(1 - \frac{1}{1+r}\right)\left(\frac{1}{1+r}\right)^3 E_t D_{t+3} + ... \end{aligned} \tag{76}$$

The sequence summarised on the first three lines of equation (76) can be described using a summation sign as

$$\sum_{k=1}^{\infty} \left[\left(\frac{1}{1+r}\right)^k (E_{t+1} D_{t+k} - E_t D_{t+k})\right] \tag{77}$$

This is an infinite discounted sum of changes to people's expectations about future dividends.

The sequence summarised on the last three lines of equation (76) can be simplified to be

$$\left(\frac{r}{1+r}\right)(1+r)\left(P_t - \left(\frac{1}{1+r}\right)D_t\right) = rP_t - \left(\frac{r}{1+r}\right)D_t \tag{78}$$

Using these two simplifications, equation (76) can be re-written as

$$\begin{aligned} P_{t+1} - P_t &= -\left(\frac{r}{1+r}\right)D_t - \left(\frac{1}{1+r}\right)D_t + rP_t + \sum_{k=1}^{\infty}\left[\left(\frac{1}{1+r}\right)^k (E_{t+1}D_{t+k} - E_t D_{t+k})\right] \\ &= -D_t + rP_t + \sum_{k=1}^{\infty}\left[\left(\frac{1}{1+r}\right)^k (E_{t+1}D_{t+k} - E_t D_{t+k})\right] \end{aligned} \tag{79}$$

which is the equation we were looking for: So the return on stocks can be written as

$$r_{t+1} = \frac{D_t + \Delta P_{t+1}}{P_t} = r + \frac{\sum_{k=1}^{\infty}\left(\frac{1}{1+r}\right)^k (E_{t+1}D_{t+k} - E_t D_{t+k})}{P_t} \tag{80}$$

Appendix 2: Programme For Return Predictability Results

Below is the text of a programme to generate the return predictability results reported in Table 1. The programme is written for the econometric package RATS but a programme of this sort could be written for any package that has a random number generator.

```
allocate  10000

set y = 0
set tstats_1lag = 0
set tstats_20lag = 0
set tstats_50lag = 0

do k = 1,10000

set y 2 300 = 0.99*y{1} + %ran(1)
set dy   = y - y{1}
set dy20 = y - y{20}
set dy50 = y - y{50}

linreg(noprint) dy 101 300
# y{1}
comp tstats_1lag(k) = %tstats(1)

linreg(noprint) dy20 101 300
# y{20}
comp tstats_20lag(k) = %tstats(1)

linreg(noprint) dy50 101 300
# y{50}
comp tstats_50lag(k) = %tstats(1)

end do k

stats tstats_1lag
stats tstats_20lag
stats tstats_50lag
```