THE JOURNAL OF FINANCE • VOL. LVII, NO. 3 • JUNE 2002

Learning, Asset-Pricing Tests, and Market Efficiency

JONATHAN LEWELLEN and JAY SHANKEN*

ABSTRACT

This paper studies the asset-pricing implications of parameter uncertainty. We show that, when investors must learn about expected cash flows, empirical tests can find patterns in the data that differ from those perceived by rational investors. Returns might appear predictable to an econometrician, or appear to deviate from the Capital Asset Pricing Model, but investors can neither perceive nor exploit this predictability. Returns may also appear excessively volatile even though prices react efficiently to cash-flow news. We conclude that parameter uncertainty can be important for characterizing and testing market efficiency.

THERE IS MUCH EVIDENCE THAT STOCK RETURNS are predictable. At the aggregate level, variables such as interest rates, financial ratios, and the default premium appear to forecast stock returns (e.g., Fama and French (1989) and Lewellen (2001)). Further, LeRoy and Porter (1981) and Shiller (1981) argue that price volatility cannot be explained solely by changes in dividends, providing indirect evidence that stock returns are predictable. At the firm level, Fama and French (1992, 1996) and Jegadeesh and Titman (1993) show that size, book-to-market, and past returns together explain much of the cross-sectional variation in average returns. There seems little doubt that expected returns vary both cross-sectionally and over time.

The interpretation of predictability is more contentious. The empirical results are potentially consistent with either market efficiency or mispricing. In general terms, market efficiency implies that prices fully reflect available information. To formalize this idea for empirical testing, Fama (1976) distinguishes between the probability distribution of returns perceived by "the market," based on whatever information investors view as relevant, and the true distribution of returns conditional on all information. The market is said to be (informationally) efficient if these distributions are the same. It follows that, in an efficient market, investors should be aware of any cross-

* Lewellen is from the MIT Sloan School of Management and Shanken is from the Simon Graduate School of Business Administration at the University of Rochester, and NBER. We are grateful to Greg Bauer, John Campbell, John Cochrane, Kent Daniel, Eugene Fama, Gur Huberman, Ravi Jagannathan, John Long, Bill Schwert, Rob Stambaugh, Avanidhar Subrahmanyam, Jiang Wang, Zhenyu Wang, two anonymous referees, and especially René Stulz (the editor) for helpful comments and suggestions. We also thank workshop participants at Columbia, Cornell, Emory, Maryland, Michigan, Rochester, Washington in St. Louis, Yale, NBER, and the American, European, and Northern Finance Association meetings.

sectional or time variation in expected returns—predictability simply reflects changes in the risk premium. Thus, researchers must judge whether predictability is consistent with rational behavior or whether it is better explained by irrational mispricing.

In this paper, we argue that there is a third potential source of predictability: *parameter uncertainty*. When investors have imperfect information about expected returns or cash flows, they must learn about the unknown process using whatever information is available, which can be formally modeled using Bayesian analysis. Parameter uncertainty necessarily affects prices at a given point in time, through its impact on investors' beliefs, as well as the evolution of prices over time as investors learn more about the economy. We show that this learning process can be a source of predictability in a way that differs from other models with rational investors.

Research on parameter uncertainty typically focuses on the subjective distribution perceived by investors (which is relevant for portfolio decisions). Our paper emphasizes, instead, the empirical properties of returns that arise endogenously in equilibrium. We show that these properties can differ substantially from those perceived by rational investors. Empirical tests will tend to find patterns in returns—predictability, excess volatility—that seem to violate market efficiency even when all investors are rational. For example, stock returns might appear predictable to an econometrician, or appear to deviate from the Capital Asset Pricing Model (CAPM), but rational investors can neither perceive nor exploit this predictability. Researchers typically, or perhaps always, assume that any patterns found by empirical research could be exploited by rational investors (ignoring transaction costs and data snooping). Our results show that this does not have to be true.

A simple example illustrates the basic idea. Suppose that dividends are i.i.d. over time with unknown mean δ and known variance σ^2. By construction, then, dividends are serially uncorrelated and have constant volatility. Any empirical test using dividends will reflect these true properties. Now consider the process perceived by rational investors. From an investor's perspective, the mean of the dividend process is random, represented by a posterior belief about δ. Realized dividends provide information about future dividends, the mean wanders around over time, and the perceived volatility declines as the investor learns. The empirical properties clearly differ from those perceived by investors. We show, for essentially the same reason, that asset-pricing tests can find patterns in returns that are neither part of the subjective distribution (as assumed by market efficiency) nor caused by irrationality (as assumed by behavioral finance).

We develop these ideas more formally in a simple equilibrium model. Investors have incomplete information about an exogenous dividend process, which they learn about from observed cash flows. We initially assume that all parameters are constant over time, but later allow the dividend process to change (in which case, investors never fully learn the true distribution). Investors are assumed to be rational and use all available information when making decisions. As long as estimates of expected cash flows diverge from the true values, asset prices deviate from their values in the absence of estimation risk. However, prices tend to move toward "fundamental" value over time as investors update their beliefs. Through this learning process, parameter uncertainty affects the predictability, volatility, and cross-sectional distribution of returns.

In our benchmark model with perfect information, returns are unpredictable using past information. When investors learn about the cash-flow process, returns become predictable both cross-sectionally and over time (from the perspective of an econometrician, but not investors). For example, if investors begin with no information about the mean of the dividend process, prices appear to react too strongly to realized dividends. Returns become negatively related to past dividends and prices. In a fairly general sense, it appears that this phenomenon is inherent in a model with parameter uncertainty because investors' "mistakes" eventually reverse as they learn more about the economy. However, predictability can take the form of either reversals or continuations (or neither), depending on investors' prior beliefs and the cash-flow process.

Predictability in the model is fundamentally different from that in other models with rational investors. Predictability arises solely from the learning process, not from changes in investment opportunities. More importantly, predictability is not perceived by investors. It exists in the true data-generating process, and can be detected using standard empirical tests, yet investors perceive constant expected returns. That cannot be true in models with perfect information; predictability must be associated with changes in the perceived risk premium. The difference is illustrated most easily when investors are risk neutral. In our model, excess returns can be predictable under the true data-generating process even when investors are risk neutral. An econometrician, using observed returns, should expect to find predictability. In contrast, excess returns must be unpredictable if investors have perfect information.

We emphasize that predictability is *caused* by the learning process; it is not assumed as part of the model. The source of the predictability is straightforward. If the market's best guess about expected dividends is, say, higher than the true mean, the stock price will be inflated above its fundamental value. Since future dividends have a lower mean than the market's estimate, investors will, on average, perceive a negative surprise over the subsequent period. An econometrician, looking back, will find that relatively high prices predict relatively low future returns. This story resembles the standard mispricing argument, but reversals in our model are driven by completely rational behavior on the part of investors. In fact, investors know that returns are negatively autocorrelated but cannot take advantage of it. They would want to exploit this pattern by investing more aggressively when the price is too low, but of course they cannot know when this is the case.

In short, our primary message is that parameter uncertainty drives a wedge between the distribution perceived by investors and the distribution estimated by empirical tests. We report numerical simulations, roughly calibrated to U.S. data, which suggest that the effects can be economically large. However, we hesitate to draw strong conclusions because the simulations do not match all the characteristics of the U.S. market, and our point is not to argue that parameter uncertainty necessarily explains specific asset-pricing anomalies. Rather, we emphasize that many tests of market efficiency cannot distinguish between a market with learning and an irrational market. We believe that a world with parameter uncertainty is the appropriate benchmark for evaluating apparent anomalies.

Our results extend a growing literature on learning and parameter uncertainty. Our cross-sectional results clarify the single-period models of Bawa, Brown, and Klein (1979) and Coles and Loewenstein (1988). They argue that estimation risk should not affect tests of the CAPM, contrary to the findings of the current paper. In the continuous-time literature, Merton (1971) and Williams (1977) show that parameter uncertainty creates a new state variable representing investors' beliefs. The hedging demand associated with this state variable can cause deviations from the CAPM (see also Detemple (1986), Dothan and Feldman (1986), and Gennotte (1986)). Our results are different because investors attempt to hold mean-variance efficient portfolios; it is their mistakes, not their hedging demands, that induce deviations from the CAPM.

The time-series implications of our model expand on the observations of Stulz (1987) and Lewis (1989). They point out that prices can appear to react inefficiently to information simply because investors learn about the economy (see also Veronesi (1999)). Wang (1993) and Brennan and Xia (1998), like the current paper, show that learning about an unobservable state variable can increase return volatility. Timmermann (1993, 1996) shows that parameter uncertainty might induce both predictability and excess volatility. We extend his work by analyzing an equilibrium model with fully rational (Bayesian) investors, and we discuss the implications for market efficiency and the cross section of expected returns. Finally, Kandel and Stambaugh (1996), Barberis (2000), Pastor (2000), and Xia (2001) all discuss portfolio choice with parameter uncertainty, but they do not analyze the equilibrium implications of learning.

The paper is organized as follows. Section I describes the model and Section II derives equilibrium. Sections III and IV analyze time-series and cross-sectional tests of predictability, respectively. Section V generalizes the model to allow for time-varying parameters, and Section VI presents numerical simulations. Section VII concludes.

I. The Model

We present a simple overlapping-generations model of the capital market. Many features of the model are borrowed from DeLong et al. (1990), who study how noise traders affect prices. Like DeLong et al., we allow investors' beliefs to diverge from the true distribution. In our model, investors are rational and use all available information when making decisions.

A. Assets

There exists a riskless asset which pays real dividend r in every period, $t = 1,\ldots,\infty$. Following DeLong et al. (1990), the riskless asset is assumed to have perfectly elastic supply: It can be converted into, or created from, one unit of the consumption good in any period. As a result, its price in real terms must equal one and the riskless rate of return equals r.

The capital market also consists of N risky securities. These assets each have one unit outstanding and pay real dividend d_t, an $N \times 1$ vector, in period t. We initially assume that dividends are i.i.d. over time and have a multivariate normal distribution (MVN):

$$d_t \sim \text{MVN}[\delta, \Sigma], \tag{1}$$

where δ is the mean vector and Σ is a nonsingular covariance matrix. Notice that the parameters are assumed to be constant. This implies that estimation risk vanishes as t goes to infinity. In reality, the economy evolves over time, so parameter uncertainty is unlikely to disappear even after a long history of data. Section V captures this idea by generalizing the model to include unobservable shocks to the parameters. The model with constant parameters is much easier to analyze and, we believe, conveys the intuition more effectively.

The i.i.d. assumption is not intended to be realistic, but dramatically simplifies the exposition. Again, we later relax this assumption and allow dividends to follow a geometric random walk. We have also explored a model with autocorrelated dividends, and the qualitative results appear to be similar. Throughout the paper, investors are assumed to know the form of the distribution function (i.i.d. and normal), but may not know its parameters.

B. Investors

Individuals live for two periods, with overlapping generations. Following DeLong et al. (1990), there is no first-period consumption, no labor supply decision, and no bequest. Individuals decide only how to invest their exogenously given wealth. We assume that the representative investor has constant absolute risk aversion:

$$U(w) = -\exp(-2\gamma w), \tag{2}$$

where w is second-period wealth and $\gamma > 0$ is the risk-aversion parameter. This setup ignores the intertemporal nature of the investment decision, which limits the way learning can affect equilibrium. Merton (1971) and Williams (1977) show, for example, that learning creates a state variable representing investors' current beliefs. Portfolio decisions will contain a corresponding hedging component if investors live for many periods. Because investors in our model are short-lived, a similar hedging demand does not arise.

The investor maximizes expected utility, where the expectation is based on the subjective belief about next-period wealth. Let p_t be the vector of risky-asset prices and x_t the vector of shares held in the portfolio. If prices and dividends are normally distributed, the investor will choose

$$x_t^* = \frac{1}{2\gamma}[\mathrm{var}_t^s(p_{t+1} + d_{t+1})]^{-1}[E_t^s(p_{t+1} + d_{t+1}) - (1 + r)p_t], \tag{3}$$

where E_t^s and var_t^s denote the subjective expectation and variance at t.[1] (We always label subjective moments with an s superscript.) The first term in brackets is the payoff covariance matrix and the second term is the expected payoff in excess of the riskless return. Investors attempt to hold mean-variance efficient portfolios, so x_t^* is the tangency portfolio under the subjective distribution.

Equilibrium in the economy, which treats current and future prices as endogenous, must satisfy (3). In addition, demand for the risky assets must equal the supply in every period. Setting $x_t^* = \iota$, where ι is an $N \times 1$ vector of ones, and solving for price yields

$$p_t = \frac{1}{1 + r}[E_t^s(p_{t+1} + d_{t+1}) - 2\gamma\,\mathrm{var}_t^s(p_{t+1} + d_{t+1})\iota]. \tag{4}$$

This equation gives the equilibrium current price in terms of next period's price, which in turn will be endogenously determined.

II. Equilibrium

We derive equilibrium with and without parameter uncertainty. We assume in both cases that investors anticipate how prices react to new information. In other words, equilibrium satisfies the rational-expectations property that the pricing function perceived by investors equals the true pricing function (Lucas (1978)). This condition does not imply, however, that investors' subjective belief about the *distribution* of returns equals the true *distribution*. The two distributions are the same only with perfect information.

A. Perfect Information

Suppose, initially, that investors know the dividend process. Since dividends are i.i.d., a natural equilibrium to look for is one in which prices are constant, or $p_t = p$. This implies that $E_t(p_{t+1} + d_{t+1}) = p + \delta$ and $\mathrm{var}_t(p_{t+1} + d_{t+1}) = \Sigma$. Substituting into (4) and solving for price yields

[1] Prices are normally distributed in equilibrium because they are a linear function of dividends (see equation (8)). Some of our simulations later assume that investors are risk neutral and dividends are log-normally distributed.

$$p = \frac{1}{r}\,\delta - \frac{2\gamma}{r}\,\Sigma\iota. \tag{5}$$

Price equals expected dividends discounted at the riskless rate minus an adjustment for risk. An asset's systematic risk, $\Sigma\iota$, is important, rather than its total variance. The market portfolio has value

$$p_M = \iota' p = \frac{1}{r}\,\delta_M - \frac{2\gamma}{r}\,\sigma_M^2, \tag{6}$$

where the market's expected dividend and variance are $\delta_M = \iota'\delta$ and $\sigma_M^2 = \iota'\Sigma\iota$. Equation (6) also gives the price of a single risky asset when $N = 1$ (compare (5) and (6)). We will focus on the market portfolio when we discuss the time-series implications of learning.

B. Parameter Uncertainty

We now relax the assumption that investors know the dividend process. Suppose, in particular, that investors begin with a diffuse prior over δ. Although this prior permits δ to be negative, it is the standard representation of "knowing little" about the mean and simplifies the algebra. We later consider informative priors, which can have important effects on price behavior. Our initial results should, therefore, be viewed as illustrative, but not completely representative. For simplicity, we continue to assume that investors know the covariance matrix of dividends. Previous research finds that uncertainty about the covariance matrix is relatively unimportant (e.g., Coles, Loewenstein, and Suay (1995)), and we doubt that it would affect our basic conclusions.

Investors update their beliefs using Bayes rule, incorporating the information in observed dividends. With a diffuse prior, the posterior distribution of δ at time t is $\mathrm{MVN}[\bar{d}_t, (1/t)\Sigma]$, where $\bar{d}_t$ is the vector of average dividends observed up to time t. The subjective, or in Bayesian terms "predictive," distribution of dividends is

$$d_{t+1} \sim^{s} \mathrm{MVN}\left[\bar{d}_t, \frac{t+1}{t}\,\Sigma\right]. \tag{7}$$

An investor's best guess about the mean is just the average realized dividend. The covariance matrix reflects both the true variance, Σ, and uncertainty about the mean, Σ/t.

Rational expectations requires not only that investors are Bayesian, but also that they anticipate how prices react to new information. Assuming rationality, Appendix A shows that the equilibrium price is

$$p_t = \frac{1}{r}\,\bar{d}_t - 2\gamma f(t)\Sigma\iota, \tag{8}$$

where $f(t)$ is a deterministic function of time. Price is similar to the model with perfect information. The mean of the predictive distribution, $\bar{d}_t$, replaces δ in the first term and $f(t)$ replaces $1/r$ in the second term. The function $f(t)$ decreases as t gets larger and approaches $1/r$ in the limit, implying that price eventually converges to the price with perfect information. The market portfolio has value

$$p_{M,t} = \frac{1}{r}\bar{d}_{M,t} - 2\gamma f(t)\sigma_M^2, \tag{9}$$

where $\bar{d}_{M,t} = \iota'\bar{d}_t$ is the average dividend on the market portfolio from $t = 1$ to t. Again, it is straightforward to show that the general model collapses to (9) when $N = 1$.

Several colleagues have noted that the pricing function in equation (8) could also be generated by a model with nonstationary dividends and no estimation risk. In particular, suppose that investors have perfect information and the *true* mean of the dividend process evolves over time as a function of average dividends (that is, $\delta_{t+1} = \bar{d}_t$). In this case, the pricing function would be identical to the price in our model. Notice, however, that prices should evolve quite differently in the two models. With a changing dividend process, gross returns would be positively related to lagged dividends and prices would exhibit no tendency to revert to a long-run mean. The opposite is true in our model, as we show in the next section.

III. Asset-Pricing Tests: Predicting Returns

Equilibrium is determined by investors' beliefs. Empirical tests depend, instead, on the true data-generating process. In this section, we study the time-series behavior of prices and returns, highlighting the difference between the properties perceived by investors and the properties of empirical tests. The next section considers the cross-sectional distribution of returns. In both cases, the exact nature of predictability is somewhat specific to the current model, but our conclusions about market efficiency are more general.

We now consider a model with one risky asset, interpreted as the market portfolio. The stock price is given by equation (9); we drop the subscript M throughout the section for convenience. Also, to focus on the main ideas, we assume in this section that investors are risk neutral. To see why, notice that the change in price from t to $t + 1$ equals

$$p_{t+1} - p_t = \frac{1}{r}(\bar{d}_{t+1} - \bar{d}_t) + 2\gamma[f(t) - f(t+1)]. \tag{10}$$

The price change contains two components. The first term is random and reflects changes in beliefs about expected dividends. The second term is deterministic and arises because estimation risk declines steadily over time.

When we talk about predictability, the deterministic portion serves only to add an additional, nonrandom component to the equations. It does not affect any of the results and drops out if investors are risk neutral.

A. Investor Perceptions versus the True Data-Generating Process

Before we discuss predictability, it is useful to emphasize the difference between beliefs and the true data-generating process. The standard definition of market efficiency says the two should be the same. That definition is fine if investors have perfect information, but it breaks down with parameter uncertainty. The reason is simple: From an investors' perspective, the underlying parameters are random and fluctuate over time as new information arrives. For an empirical test, however, the data-generating process depends on whatever the actual, but unknown, parameters really are. The true process does not have random parameters. As a consequence, the empirical properties of returns can differ significantly from the properties perceived by investors.

The difference can be seen clearly in the pricing process. Under the subjective distribution, prices follow a martingale:

$$E_t^s[p_{t+1} - p_t] = 0. \tag{11}$$

Prices react to changes in (subjective) expected dividends, and it must be the case that investors cannot predict these changes. In contrast, prices do not wander completely randomly under the true data-generating process; they must eventually converge to "fundamental" value. If current beliefs about expected dividends are greater than the true mean, prices are also temporarily inflated and will drop in the future. If current beliefs are below the true mean, prices are temporarily depressed and will eventually rise. Thus, under the true distribution,

$$E_t[p_{t+1} - p_t] = \frac{1}{r(t+1)}(\delta - \bar{d}_t). \tag{12}$$

Fundamentally, the difference between (11) and (12) is that dividends are truly drawn from a distribution with a constant mean. The subjective distribution, on the other hand, views the mean as a random variable. This difference drives all of the following results on predictability, excess volatility, and tests of market efficiency.

To illustrate the ideas, Figure 1 depicts a sample price path for the risky asset. The figure assumes that investors are risk neutral and the riskless rate is 0.05. Dividends have mean 0.05 and standard deviation 0.10, taken to be similar to the dividend yield and volatility of dividends on the market portfolio. With perfect information, the price of the risky asset would equal one (fundamental value). The price with estimation risk depends on realized dividends, which we randomly draw from a normal distribution. The figure

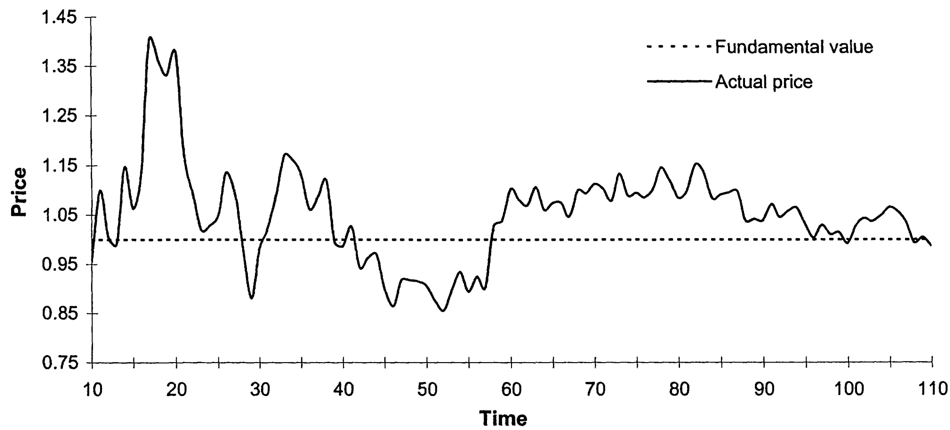

Figure 1. Equilibrium price of the risky asset. This figure illustrates a sample price path for the risky asset when the dividend process is known (fundamental value in the figure; see equation (6) in the text) and when investors must estimate expected dividends (actual price; see equation (9)). The riskless rate is 0.05, dividends have true mean 0.05 and standard deviation 0.10, and investors are risk neutral. Without estimation risk, the price of the risky asset is one. With estimation risk, the price depends on past average dividends, which we randomly select from a normal distribution.

shows that the price of the risky asset tends to revert towards fundamental value. Empirically, prices will not appear to be a random walk—they eventually converge to one—even though price changes *are* completely unpredictable by investors. Notice, also, that we do not need to know the underlying parameters to observe mean reversion in prices. All that is important is that prices converge to *some* value.

B. Predicting Returns

As observed above, prices follow a martingale under the subjective distribution. Realized returns at $t + 1$ are given by[2]

$$R_{t+1} = d_{t+1} + \frac{1}{r(t+1)}(d_{t+1} - \bar{d}_t). \qquad (13)$$

The first term equals the realized dividend and the second term equals the change in price. The price change reflects new information about expected dividends, given by the investor's dividend surprise, $d_{t+1} - \bar{d}_t$. Although the price change is not predictable by investors, it is predictable under the true

[2] For simplicity, we examine the predictability of *gross* returns rather than *rates of* return. The analysis with rates of return is more difficult because it involves ratios, but the qualitative results are similar.

data-generating process. Given the true mean of the dividend process, the expected return is

$$E_t[R_{t+1}] = \delta + \frac{1}{r(t+1)}(\delta - \bar{d}_t). \tag{14}$$

Price revisions, and hence returns, are negatively related to past cash flows. Prices depend on the mean dividend perceived by investors, $\bar{d}_t$; when past dividends have been high, future expected returns are low.

Equation (14) implies that prices, dividends, and past returns will all capture time variation in expected returns. For example, the autocovariance of returns is given by[3]

$$\mathrm{cov}[R_t, R_{t+1}] = -\frac{1}{rt(t+1)}\sigma^2. \tag{15}$$

Because of the learning process, past mistakes eventually reverse and returns are negatively autocorrelated. A researcher who ignores estimation risk, and observes that business conditions do not change, would come to the incorrect conclusion that investors overreact: higher returns today predict lower returns in the future. Similarly,

$$\mathrm{cov}[d_t, R_{t+1}] = -\frac{1}{rt(t+1)}\sigma^2. \tag{16}$$

A high dividend today predicts lower future returns, which would suggest that investors naively extrapolate recent dividend performance into the future (e.g., Lakonishok, Shleifer, and Vishny (1994)). However, investors are completely rational in our model and the predictability is driven entirely by parameter uncertainty.

These results imply that an econometrician will tend to find mean reversion in stock prices; past prices and dividends are both negatively correlated with future returns under the true data-generating process. We later present detailed simulations, allowing for changing parameters and nonstationary dividends, but the basic idea is illustrated by Figure 1. Estimation error eventually reverses, so a researcher looking at the data will, on average, find mean reversion in returns. We have also estimated, using simulations, the correlation between excess rates of return and lagged average dividends. For $t = 10$ to 80 (i.e., 70 "years" of data), the correlation with perfect information is -0.136 (this is negative because of small-sample bias). It drops to -0.259 when investors begin with diffuse priors. The current model is not

[3] This covariance is time dependent because estimation risk declines over time. We will break the strong connection between time and predictability in Section V when we allow the parameters to change.

realistic, so we do not take the estimates very seriously, but they do suggest that parameter uncertainty can have a large impact on empirical tests. Investors ignore the negative relation between returns and past dividends because it provides no useful information about *future* expected returns (a point we will return to shortly).

Researchers generally assume that lead-lag correlations, like those above, imply either time variation in the (perceived) risk premium or irrationality. Parameter uncertainty provides a third explanation. Unlike other models with rational investors, investors do not perceive that the risk premium varies over time. To see this clearly, consider the unexpected portion of returns, $UR_{t+1} \equiv R_{t+1} - E_t^s[R_{t+1}]$. Unexpected returns simply equal excess returns since investors are risk neutral. From (13):

$$UR_{t+1} = \left[1 + \frac{1}{r(t+1)}\right](d_{t+1} - \bar{d}_t). \tag{17}$$

It follows that

$$E_t[UR_{t+1}] = \left[1 + \frac{1}{r(t+1)}\right](\delta - \bar{d}_t). \tag{18}$$

Like total returns, the unexpected portion of returns is correlated with past prices and dividends. It is precisely this result that differentiates our model from other models with rational investors. Indeed, a standard result in finance, and one that underlies all tests of market efficiency, is that forecast errors should be uncorrelated with past information if investors are rational. Equation (18) shows that this does not have to be true.

The apparent predictability of forecast errors is consistent with rational expectations because it is based on the unknown, true data distribution. Investors never know whether past dividends have been above or below the true mean. However, as investors learn more about expected cash flows, they correct prior mistakes and prices converge to fundamental value. A researcher, looking back, will observe a negative relation between prices and unexpected returns (as illustrated by Figure 1) that seems to violate market efficiency.

C. Interpretation

The last paragraph notes that a researcher, *looking back*, can observe the predictability of returns. Predictability in the true data-generating process shows up in realized returns, and standard empirical tests (like predictive regressions) can detect this predictability. Notice, however, that an econometrician, like investors in the model, cannot actually forecast *future* price movements. Predictive regressions implicitly use information that investors do not have when portfolio decisions are made (since they use an entire time series of data). Looking forward, regressions estimated from past data do not help to forecast future returns.[4]

The distinction between "in-sample" and "out-of-sample" predictability is interesting. When researchers say that returns are predictable, we typically mean that returns are correlated with ex ante observable variables. Our model generates this sort of predictability, yet investors cannot actually forecast returns. We are unaware of any other model that makes this prediction. These results provide strong motivation for designing tests of out-of-sample predictability (e.g., Bossaerts and Hillion (1999) and Goyal and Welch (1999)). Although such tests are sometimes reported, the usual motivation is to see whether investors could have learned about predictability early enough to take advantage of it. Our results are different: The model shows why predictability arises in the first place (learning creates the predictability), and investors are aware of the predictability but cannot take advantage of it.

More importantly, the simple dichotomy between in-sample and out-of-sample predictability is misleading. The learning process has subtle effects on asset-pricing tests that go beyond the simple observation that predictive regressions use information not available to investors. Because this point is important, we provide three additional examples to illustrate how parameter uncertainty can affect asset-pricing tests:

1. With parameter uncertainty, we can devise an *implementable* trading strategy that appears to generate excess returns. The strategy earns abnormal profits in a frequentist sense, but not from the Bayesian perspective of investors. The strategy is simple: Hold the riskless asset if average dividends have been above some arbitrary cutoff value K and, otherwise, hold the stock. The excess return from the strategy at $t + 1$ is

$$\pi_{t+1} = \lambda_t (R_{t+1} - r p_t), \tag{19}$$

where λ_t equals one if $\bar{d}_t < K$ and zero otherwise. Taking unconditional expectations, the expected profit is

$$E[\pi_{t+1}] = c \cdot \text{prob}(\lambda_t = 1) \cdot E[\delta - \bar{d}_t | \bar{d}_t < K], \tag{20}$$

where $c = 1 + 1/r(t + 1)$. The last term is just the expected value of a truncated normal distribution. It is positive for any K since $E[\bar{d}_t | \bar{d}_t < K] < \delta$ (e.g., Greene (2000, p. 899)). Thus, we have found a

[4] To see why, consider our result that lagged dividends forecast unexpected returns. We showed that $\text{cov}(\bar{d}_t, UR_{t+1})$ is negative, or equivalently, $E[(\bar{d}_t - \delta) \times UR_{t+1}] < 0$. To use this correlation to forecast returns, an investor needs to know whether $\bar{d}_t$ is greater than or less than δ. Since our best guess is that $\bar{d}_t = \delta$, we cannot actually predict whether returns will be high or low.

simple market-timing strategy that earns abnormal profits. An econometrician who implements the strategy in real time should expect to find (in a frequentist sense) excess returns, but a Bayesian investor perceives zero profits.

This apparent contradiction highlights the difference between frequentist empirical tests and Bayesian decision making. The expectation above is based on *repeated sampling* of dividends and prices; it averages across all possible outcomes for realized dividends (given a fixed δ). The profitability of the strategy is driven by the simple observation that, in repeated sampling, if $\bar{d}_t$ is less than K, then on average $\bar{d}_t$ will also be less than the true mean. A Bayesian investor, in contrast, must make decisions *conditional on a given draw* of the dividend process. The investor views δ as a random variable and, conditional on past dividends, the investor's belief about δ is always centered around $\bar{d}_t$; from the investor's perspective, $\bar{d}_t$ is just as likely to be above as below δ. The hypothetical repeated-sampling behavior of prices is irrelevant. Thus, a frequentist empirical test views δ as fixed and dividends as random, while a Bayesian investor reverses these roles (see Berger (1985) for an extensive discussion of these issues).

2. Although we are getting a bit ahead of ourselves, the next section shows that stock returns are cross-sectionally predictable: The slope coefficient in Fama–MacBeth regressions, which estimate the cross-sectional correlation between next period's returns and this period's prices, is expected to be negative. In this regression, the researcher only uses information available to investors. A negative coefficient on lagged price, for example, says that the relative prices of two firms, both of which are observed today, tell us something about their relative returns in the future. There is no sense in which a cross-sectional regression uses forecasting information that investors do not have. The idea that returns are predictable in-sample, but not out-of-sample, breaks down in cross-sectional regressions.
3. We observed above that predictive regressions are expected to find in-sample predictability. Intuitively, the reason is that regressions implicitly measure the independent variable relative to its full-sample mean, thus using information not available to investors. We can fix this "problem" by using a moment condition that does not rely on the full-sample mean. For example, suppose we test whether $E[p_t \times UR_{t+1}] = 0$. This test, similar to many GMM tests of rational expectations models, focuses on the orthogonality between return surprises and lagged information.[5] With perfect information, the moment condition must be zero. However, with estimation risk, $E[p_t \times UR_{t+1}] = \text{cov}(p_t, UR_{t+1}) < 0$, where the equality follows from the fact that UR_{t+1} has unconditional mean zero. This

[5] Note that, in our model, unexpected returns simply equal excess returns because investors are risk neutral. UR_{t+1} does not depend on unknown parameters that need to be estimated, as it would in a typical GMM setting.

test satisfies the requirement that the econometrician only uses past information to predict returns; in contrast to predictive regressions, lagged prices are compared to zero rather than to the full-sample mean. Thus, even in a time-series setting, an out-of-sample test can find predictability (much like the market-timing example).

These examples show that parameter uncertainty can affect empirical tests in surprising ways. Even a test based on an implementable trading strategy is expected to find, from a frequentist perspective, abnormal returns. Researchers always assume that, in the absence of trading costs or sampling error, any patterns found in the data should be exploitable by rational investors. This does not have to be true: The apparent predictability in our model cannot be used to improve investment decisions.

D. Volatility

Return predictability and volatility are closely related (e.g., Campbell (1991)). Given our results above, it is clear that parameter uncertainty will also affect volatility. We briefly comment on a few results here. In the model with perfect information, prices are constant, and the variance of returns equals the variance of dividends, σ^2. With parameter uncertainty, the (true) conditional variance of price is

$$\mathrm{var}_t[p_{t+1}] = \left[\frac{1}{r(t+1)}\right]^2 \sigma^2, \tag{21}$$

and the unconditional variance is

$$\mathrm{var}[p_{t+1}] = \frac{1}{r^2(t+1)}\sigma^2. \tag{22}$$

Parameter uncertainty increases both the conditional and unconditional variances of prices. Similar to our earlier results, parameter uncertainty gives the appearance of overreaction. Again, "excess" volatility simply reflects learning; volatility is high precisely because investors rationally update their beliefs.

In the model, a modest amount of parameter uncertainty will create substantial price volatility. Suppose, for example, that investors are risk neutral, the riskless rate is 0.05, and dividends have mean 0.05 and standard deviation 0.10. (These are the values used in Figure 1.) Using these parameters, the value of the risky asset equals one when δ is known. With parameter uncertainty, the standard deviation of p_t equals $2/\sqrt{t}$. This remains significant as a percentage of fundamental value for rather large t. For example, when t is 100, a two-standard-deviation confidence interval is 80 percent of fundamental value. Yet, after this many periods in the model, the subjective standard deviation of dividends is less than 1 percent greater than the true standard deviation.

Notice that price movements are not explained by subsequent changes in dividends. In fact, prices are *completely* uncorrelated with future dividends; prices are backward looking and appear to overreact to information. As a result, prices violate the variance bounds that have been the focus of much empirical research. For example, Shiller (1981) argues that an immediate consequence of optimal forecasts is that

$$\mathrm{var}(p_t) \leq \mathrm{var}(p_t^*), \tag{23}$$

where p_t^* is the ex post rational price, or the price based on realized, rather than expected, dividends. That is, p_t^* is given by

$$p_t^* = \sum_{k=1}^{\infty} \frac{1}{(1+r)^k} d_{t+k} - 2\gamma f(t)\sigma^2. \tag{24}$$

With perfect information and rational investors, the bound holds because p_t^* equals the actual price plus a random, unpredictable forecast error. But we saw that, with learning, forecast errors can be negatively related to price. In the current model, the variance of p_t^* is

$$\mathrm{var}(p_t^*) = \sum_{k=1}^{\infty} \frac{1}{(1+r)^{2k}} \sigma^2 = \frac{1}{r^2 + 2r} \sigma^2. \tag{25}$$

Comparing this to (22), we see that the volatility bound will be violated for $t \leq 1 + 2/r$. Perhaps more directly, prices violate the basic premise of the volatility-bound literature, that price changes should only reflect changes in true expected dividends. With learning, new information about future dividends does not have to correspond to changes in the true distribution.

The model also provides an interesting perspective on Campbell's (1991) variance decomposition of returns. He shows that return movements can be decomposed, mathematically, into news about dividends and news about expected returns. In our model, all movement in prices is due to dividend news (as perceived by rational investors). The discount rate is constant, so there is no variation in perceived expected returns. Yet, in the true data-generating process, the mean of the dividend process is fixed; the data should provide no evidence that (true) expected dividends change over time. Price volatility will appear, to an econometrician, to be driven solely by innovations in expected returns. Again, asset-pricing tests might not accurately capture the underlying perceptions of investors.

IV. Asset-Pricing Tests: The CAPM

Section III showed that, with parameter uncertainty, empirical tests can find predictability not perceived by investors. We now return to the model with many assets and analyze the cross-sectional behavior of returns. Bawa and Brown (1979) and Coles and Loewenstein (1988) argue that the CAPM should continue to hold with estimation risk. That is true in our model as well, but only for the perceived, not the empirical, distribution of returns.

A. Covariances and Betas

We start with a few results on covariances and market betas. With perfect information, the covariance matrix of returns equals the covariance matrix of dividends, Σ. Introducing parameter uncertainty scales up the covariance matrix. Specifically, with parameter uncertainty, the (true) covariance matrix of returns is

$$\mathrm{var}_t[R_{t+1}] = \left[1 + \frac{1}{r(t+1)}\right]^2 \Sigma. \tag{26}$$

Equation (26) shows that parameter uncertainty increases all variances and covariances proportionally. As a result, estimation risk has no effect on market betas (for gross returns). With and without estimation risk, the vector of betas equals

$$\beta = \frac{1}{\mathrm{var}(R_{M,t})} \mathrm{cov}(R_t, R_{M,t}) = \frac{1}{\iota' \Sigma \iota} \Sigma \iota. \tag{27}$$

It is straightforward to show that the subjective covariance matrix equals the true covariance matrix multiplied by $(t+1)/t$. This implies that perceived and true betas are the same.[6]

B. Expected Returns and the CAPM

With and without parameter uncertainty, investors attempt to hold mean-variance efficient portfolios. It is immediate that the CAPM must hold *under the subjective distribution*. In terms of gross returns, the CAPM says that

$$E_t^s[R_{t+1}] = rp_t + \beta[E_t^s(R_{M,t+1}) - rp_{M,t}]. \tag{28}$$

This equation can be verified by substituting for equilibrium prices and subjective expected returns, derived earlier. The equation does not imply that the CAPM should hold empirically: Empirical tests depend on true, not sub-

[6] These results are an artifact of the diffuse prior. With an informative prior, perceived and true covariance matrices may not be proportional and neither has to be proportional to Σ (see Clarkson, Guedes, and Thompson (1996) and Lewellen and Shanken (2000)).

jective, expected returns. To understand the properties of empirical tests, we consider ex post deviations from the CAPM, given by

$$a_{t+1} \equiv R_{t+1} - rp_t - \beta[R_{M,t+1} - rp_{M,t}]. \tag{29}$$

The vector a_{t+1} is similar to the vector of unexpected returns, but the realized market return enters (29) rather than the expected market return. By examining a_{t+1}, we eliminate predictability that is related to the aggregate market.

Deviations from the CAPM must be unpredictable under the subjective distribution:

$$E_t^s[a_{t+1}] = 0. \tag{30}$$

With perfect information, market efficiency implies that the true expectation is also zero. That restriction, of course, forms the basis for empirical tests of the CAPM. Cross-sectional regressions, like those in Fama and MacBeth (1973), test whether firm characteristics predict cross-sectional variation in $a_{i,t+1}$. The time-series approach of Gibbons, Ross, and Shanken (1989) tests whether the unconditional expectation of a_{t+1} is zero (which follows from the law of iterated expectations). Finally, various asset-pricing tests directly examine the conditional expectation of a_{t+1} (e.g., Harvey (1989) and Shanken (1990)).

With parameter uncertainty, the true expected value of a_{t+1} does not have to be zero. Substituting for prices and returns in (29) and taking expectations yields

$$E_t[a_{t+1}] = -\left[1 + \frac{1}{r(t+1)}\right][(\bar{d}_t - \delta) - (\bar{d}_{M,t} - \delta_M)\beta]. \tag{31}$$

Deviations from the CAPM are negatively related to past prices and dividends. In particular, for any asset:

$$\mathrm{cov}[p_{i,t}, a_{i,t+1}] = -\left[1 + \frac{1}{r(t+1)}\right]\frac{1}{rt}[\mathrm{var}(d_i) - \beta_i \mathrm{cov}(d_i, d_M)], \tag{32}$$

where the last term in brackets is the residual variance when the asset's dividend is regressed on the market dividend. The empirical properties of $a_{i,t+1}$ are similar to the behavior of unexpected returns. Again, we see that parameter uncertainty induces price reversals and apparent overreaction. When investors' best guess about a stock's expected dividend is above the true mean (after adjusting for marketwide mispricing), price is inflated above its fundamental value and expected returns are lower than predicted by the CAPM. These properties imply that empirical tests will tend to reject the

CAPM even though it works perfectly from the perspective of rational investors. The apparent deviations from the CAPM cannot be used to improve portfolio decisions.

Equation (32) is essentially a time-series relation. Predictability arises because investors do not know whether past dividends have been above or below the true mean. At any point in time, however, investors *do* observe whether a security's dividends are above or below the cross-sectional average. Our initial guess, then, was that deviations from the CAPM would not be *cross-sectionally* related to lagged prices: If cross-sectional variation in $a_{i,t+1}$ is related to the observable quantity $p_{i,t}$, it would seem that investors could use this information to earn abnormal returns. Surprisingly, this intuition is wrong. In sample, the cross-sectional relation between $a_{i,t+1}$ and $p_{i,t}$ is

$$\mathrm{cov}_{t+1}^{cs}[p_{i,t}, a_{i,t+1}] = \frac{1}{N}\sum_i (a_{i,t+1} - \bar{a}_{t+1}^{cs})(p_{i,t} - \bar{p}_t^{cs}). \tag{33}$$

Taking unconditional expectations yields

$$E[\mathrm{cov}_{t+1}^{cs}(p_{i,t}, a_{i,t+1})] = \frac{1}{N}\sum_i \mathrm{cov}(a_{i,t+1}, p_{i,t}) < 0, \tag{34}$$

which is negative because every covariance term is negative. In the presence of estimation risk, lagged dividends and prices explain cross-sectional variation in expected returns (after controlling for beta). Investors understand the negative cross-sectional relation, but they cannot use this information to be better off.

We find this result paradoxical. To gain some intuition, consider the decision-making process of a rational investor. Implicitly, the expectation in (34) integrates over all possible price paths from time 1 to $t + 1$. However, at time t, the *conditional* cross-sectional relation can be either positive or negative, depending on the difference between $\bar{d}_t$ and δ. In other words, conditional on observing $\bar{d}_t$, the cross-sectional covariance depends on the true value of δ. Investors understand this dependence, and their beliefs about δ determine their investment choices. They integrate over the posterior distribution of δ to make portfolio decisions. The resulting belief about a_{t+1} will always have mean zero. The point is simply that investors do not ignore the relation between dividends and returns, but their best forecast of a_{t+1} at any point in time is always zero.

Alternatively, we can think about this in terms of an individual asset. Suppose that an asset has a relatively high price compared with other stocks. Does this imply that the asset is overvalued relative to its fundamental value? The answer depends, of course, on the actual value of δ_i, which is unknown. Integrating over the posterior belief about δ_i, an investors' best guess at all times is that the asset is fairly priced. Yet, in hypothetical repeated sampling, the asset with the highest price will, on average, be overvalued. This puzzle again highlights the distinction between the conditional nature of Bayesian decision making (conditional on the observed prices) and the frequentist perspective of classical statistics. For a Bayesian investor, hypothetical repeated sampling is irrelevant to the portfolio decision, which must be made after observing only a single realization of prices.

To illustrate the cross-sectional results, we simulate prices and returns in the model. Similar to the example in Section III, we assume that investors are risk neutral and the riskless rate is 0.05. In addition, all risky assets, with $N = 15$, have true expected dividends equal to 0.05. Hence, all prices equal one in the absence of estimation risk. When δ is unknown, security prices depend on realized dividends, which we randomly generate from a MVN distribution. To provide a reasonable covariance matrix, we estimate the return covariance matrix for 15 industry portfolios formed from all stocks on the Center for Research in Security Prices (CRSP) database.

Both the time-series and cross-sectional behavior of returns reveal the price-reversal effect of estimation risk. Take, first, one randomly selected simulation. For $t = 10$ through 110, the correlation between total return and lagged price is negative for every security, with a mean correlation of -0.21. Deviations from the CAPM also appear predictable based on lagged prices. In time series, the average correlation between $a_{i,t+1}$ and $p_{i,t}$ is -0.16, and 14 out of the 15 correlations are negative. Cross-sectionally, the relation between $a_{i,t+1}$ and $p_{i,t}$ is significantly negative in Fama–MacBeth style regressions, with a t-statistic of -3.97. On average, an increase in price from one standard deviation below to one standard deviation above the cross-sectional mean leads to a -0.042 change in $a_{i,t+1}$. Since prices are generally close to one, this would imply that Jensen's alpha, based on rates of return, decreases by approximately 4.2%. This example is typical. Across 2,500 simulations, Fama–MacBeth regressions produce an average t-statistic of -3.75 with a standard deviation of 0.94. Although investors attempt to hold mean-variance efficient portfolios and use all available information when making decisions, expected returns differ substantially from the predictions of the CAPM.

V. Parameter Uncertainty in the Long Run

The analysis above shows that parameter uncertainty can, in principle, substantially affect asset-pricing tests. The model makes a number of unrealistic assumptions—diffuse priors, i.i.d. dividends, and constant parameters—which make it difficult to judge the empirical significance of estimation risk. We now relax these assumptions to make the model a bit more realistic. The next section presents simulations to suggest the practical importance of parameter uncertainty.

A. Informative Priors

Our goal is a model in which parameter uncertainty remains in the long run. This requires that the parameters change over time, so investors never learn the true dividend process. Since investors should have some information about changes in the parameters, the model inherently gives rise to informative priors. Therefore, as a first step, we discuss the role of informative priors in the basic model.

Diffuse priors have two important effects on the model. First, investors' beliefs are determined entirely by past realized dividends. With informative priors, investors would put some weight on their initial information and less weight on the data. Second, investors' beliefs about one asset depend only on the realized dividends of that asset. They do not depend at all on the realized payoffs of other securities. With an informative prior, dividends on assets with relatively high amounts of prior information can be useful in valuing other assets. We focus on the first issue in this section; Lewellen and Shanken (2000) discuss the second issue in detail.

Consider the model with one risky asset. Assume that the variance of the dividend process, σ^2, remains known, and suppose that investors begin with some information about the mean. In particular, prior beliefs are centered around some δ^* and have variance σ^2/h, where h is a measure of prior information. Writing the variance in this form is simply for notational convenience; a variance equal to σ^2/h means that prior information is equivalent to a sample of h dividends. With this prior, the investor's belief about dividends at time t is

$$d_{t+1} \sim^s N\left[\frac{h}{t+h}\,\delta^* + \frac{t}{t+h}\,\bar{d}_t, \frac{t+h+1}{t+h}\,\sigma^2\right]. \tag{35}$$

Investors shrink their best guess about expected dividends toward their prior mean, and the variance reflects both the volatility of dividends, σ^2, and uncertainty about the mean, $\sigma^2/(t+h)$. It is clear that the prior mean exerts a permanent, yet diminishing, influence on beliefs. If the prior mean deviates from δ, investors' beliefs are biased away from the true value. As before, beliefs eventually converge to the true distribution.

Equilibrium takes nearly the same form as the original model. Price now reflects prior beliefs as well as the information in realized dividends. Denote the mean of the subjective distribution as m_t. At time t, the price of the risky asset equals

$$\begin{aligned} p_t &= \frac{1}{r}\,m_t - 2\gamma f(t+h)\sigma^2 \\ &= \frac{1}{r}\,\frac{t}{t+h}\,\bar{d}_t + \frac{1}{r}\,\frac{h}{t+h}\,\delta^* - 2\gamma f(t+h)\sigma^2, \end{aligned} \tag{36}$$

where $f(t)$ is defined in Appendix A. With informative priors, price contains a new term corresponding to the initial belief about expected dividends. The time-series properties of prices will be determined by m_t. Prior information anchors price to the investor's initial guess, but does not have a stochastic effect on prices. As a result, in the model with fixed parameters, informative priors have little effect on our earlier conclusions. Returns continue to be negatively related to past prices and dividends. For example,

$$\text{cov}[p_t, R_{t+1}] = -\frac{t}{r^2(t+h)^2(t+h+1)}\sigma^2, \tag{37}$$

which is negative but smaller than the expression with diffuse priors ($h = 0$). This result is quite intuitive: Prior information basically adds h periods to the model before time zero.

We should add an important caveat: Informative priors can play a larger role if the parameters change over time. Investors may appear to react slowly to changes in the dividend process (see the next section). In addition, even in the current model, forecast errors are all expected to have the same sign because of the permanent influence of the prior mean. Although the influence is nonstochastic and does not affect serial correlation in returns, it could create the appearance of underreaction in some contexts. For example, Lewis (1989) argues that a similar phenomenon accounts for the persistent forecast errors observed in the foreign exchange market in the 1980s.

Informative priors can also play a more important role with many assets. Lewellen and Shanken (2000) show that, with informative priors, the cross-sectional relation between returns and lagged prices can be either positive or negative. Investors can appear to update their beliefs too slowly because they place less weight on the data and more on their prior beliefs. In addition, if investors have more information about some securities than others, empirical estimates of beta may not fully capture the risk perceived by investors.

B. Renewing Parameter Uncertainty

Our basic model has the unattractive feature that parameter uncertainty steadily diminishes. We assumed that the dividend process is fixed, so investors never lose information. In this section, we extend the model to include unobservable shocks to the true parameters. We consider a model with one risky asset because the section seems most applicable to aggregate returns. At the microeconomic level, firms continually appear and disappear from the stock market. It is not clear that the long-run implications of estimation risk are relevant for individual stocks.

There are many ways to prevent parameter uncertainty from vanishing in the limit. We choose a particularly simple form of "renewal" to illustrate the ideas. The model remains the same with one exception: We now assume that the true mean of the dividend process changes at known, fixed intervals. Specifically, every K periods, the mean is redrawn from a normal distribution with mean δ^* and variance σ_s^2. Thus, the model is essentially a sequence of short "regimes" that look like our basic model. We have analyzed alternative models in which (1) the length of the intervals is random rather than fixed, and (2) the true mean of the dividend process follows a persistent process. The qualitative conclusions from these models appear to be similar.

After an infinite number of periods, it is clear that investors should learn the distribution from which the short-run mean is drawn. Therefore, in the limit, investors' priors at the beginning of each regime would be $N[\delta^*, \sigma_s^2]$. Although we analyze these priors as a special case, we do not think that it is either the most realistic or interesting scenario because it represents an extreme amount of learning. Instead, we consider the more general beliefs $N[\delta^*, \sigma_h^2]$, which have the same mean as the actual distribution but not necessarily the same variance. Permitting the variances to be different can be justified on several grounds.

First, we are trying to capture the idea that the economy moves though periods of high and low growth. These periods might cover many years, so learning about the switching process—and its variance—is likely to be slow. Second, we have made the artificial assumption that the mean is repeatedly drawn from the same distribution. The economy undoubtedly moves through periods of relative stability and periods of rapid change, and the variance of shocks to expected dividends is likely to change over time. If investors cannot observe changes in volatility, their current estimate of volatility will not be perfect. Finally, alternative assumptions about the evolution of the true mean do not necessarily have the property that the prior variance ever converges to the true variance.[7] We abstract from these issues, and take the more expeditious approach of simply permitting the prior variance to be different from σ_s^2.

The pricing function is similar to the price in the basic model. Assume that investors are risk neutral and, for notational convenience, let $\sigma_s^2 = \sigma^2/s$ and $\sigma_h^2 = \sigma^2/h$. Realized dividends during the current interval provide no information about payoffs after the end of the interval; beliefs about those payoffs always have mean δ^*. Therefore, the price at the beginning of every regime equals δ^*/r, the value of expected dividends in perpetuity. After t periods in the current regime, the investor's predictive belief about short-run dividends has mean

$$m_t = \frac{h}{t+h}\,\delta^* + \frac{t}{t+h}\,\bar{d}_t, \tag{38}$$

where $\bar{d}_t$ is the average dividend observed from the start of the regime. Price equals

$$p_t = AF_{K-t}\, m_t + \frac{1}{(1+r)^{K-t}}\,\frac{\delta^*}{r}, \tag{39}$$

[7] For example, suppose the dividend mean δ_t follows a random walk, dividends have conditional variance σ^2, and the shocks to δ_t are uncorrelated with dividends and have variance σ_s^2. In the long run, investors' beliefs about δ_t will be $N[m_t, \sigma_h^2]$, where σ_h^2 is time invariant and $\sigma_h^2 > \sigma_s^2$.

where AF_{K-t} is an annuity factor for $K - t$ periods. The properties of prices and returns once again depend on the behavior of m_t. It is straightforward to show that excess, or unexpected, returns are given by

$$UR_{t+1} = \left[1 + \frac{AF_{K-t-1}}{t+h+1}\right](d_{t+1} - m_t). \tag{40}$$

The term in parentheses is just the unexpected dividend. It has both an immediate effect on returns (the "1" in brackets) and an indirect effect on prices (with the multiplier $AF_{K-t-1}/(t + h + 1)$).

The analysis of predictability is more complicated than in our basic model. In particular, now that the short-run mean is random, we must distinguish between expectations that are conditional on the current mean and expectations that treat the parameter as random. A combination of the two seems to be relevant for empirical tests (see below). After t periods into the current regime, the unexpected return has true mean

$$E_t[UR_{t+1}] = \left[1 + \frac{AF_{K-t-1}}{t+h+1}\right](\delta_k - m_t), \tag{41}$$

where δ_k is the current draw of the short-run mean. As in our basic model, the true unexpected return is negatively related to past dividends and prices. Consequently, *taking the value of* δ_k *as given*, the covariance between excess returns and lagged prices equals

$$\mathrm{cov}[p_t, UR_{t+1}] = -AF_{K-t}\left(1 + \frac{AF_{K-t-1}}{t+h+1}\right)\mathrm{var}(m_t), \tag{42}$$

which is negative. We refer to this expression as the *conditional covariance* because it regards the short-run mean as fixed (although it does not depend on the value of δ_k). The equation is very similar to our previous result with informative priors, except that the covariance is attenuated because price fluctuations are less pronounced (the price always returns to δ^*/r at the end of the regime). Therefore, in one sense, the effects of parameter uncertainty remain the same even in the long run: The true and subjective distributions are different, leading to price reversals.

Unfortunately, things are not quite so simple. The *unconditional covariance* between prices and unexpected returns—which regards the short-run mean as random—does not have to be negative. The unconditional covariance equals

$$\mathrm{cov}[p_t, UR_{t+1}] = AF_{K-t}\left(1 + \frac{AF_{K-t-1}}{t+h+1}\right)\frac{t}{(t+h)s}\left[1 - \frac{t+s}{t+h}\right]\sigma^2. \tag{43}$$

The sign of the covariance depends on the relative magnitudes of s and h. Recall that σ^2/s is the true variance of δ_k while σ^2/h is the prior variance. The unconditional covariance is negative when the prior variance is greater than the true ($h < s$), but positive when the prior variance is less ($h > s$). When investors believe that the shocks to expected dividends are volatile, they are relatively sensitive to realized dividends, and price reversals show up both conditionally and unconditionally. On the other hand, if the short-run mean is more variable than investors believe, they tend to be surprised by the large movements in expected dividends and require many observations to update their beliefs; returns exhibit continuations, or momentum. The cutoff value occurs when investors have exactly the right beliefs about the variance of δ_k, or when $s = h$. In this case, the unconditional covariance between excess returns and lagged prices is exactly zero.

Thus, we have two results on predictability in the renewal model: (1) the conditional covariance is always negative, regardless of the relative magnitudes of s and h, and (2) the unconditional covariance depends on whether h is less than or greater than s. The fact that the conditional covariance is negative implies immediately that excess returns are predictable, but it is not obvious to us whether the conditional or unconditional covariance is more relevant for standard empirical tests.[8] An empirical test depends on the observed sample, and implicitly conditions on the sample value (or values) of the mean parameter δ_k. This observation suggests that the conditional covariance might be most relevant. Indeed, take a particularly simple case in which the observed sample covers only one regime. Regardless of the value of δ_k, the covariance between unexpected returns and prices is expected to be negative; the correlation in this case corresponds directly to the conditional covariance. If, however, a sample covers multiple regimes, the empiricist implicitly conditions on several values of δ_k. Our formula for the conditional covariance no longer represents the population counterpart of the estimate. To muddy the waters further, if the empiricist suspects that a change in regime occurs and adds a dummy variable to the regression, or focuses on subperiod estimates, then the sample covariance will correspond once again to the conditional covariance. In short, we find it difficult to know a priori whether the conditional or unconditional covariance exerts a stronger influence on empirical tests. We turn to simulations to better understand these issues.[9]

[8] Additional explanation might be useful. A predictive regression for returns that includes regime dummies would estimate the conditional covariance, and can therefore detect the price reversals. However, it is not common to include regime dummies in predictive regressions, nor is it easy to identify regime changes.

[9] As an aside, parameter uncertainty affects more than just predictability. Volatility in the model jumps at the beginning of each regime and then slowly decays as investors learn about the short-run mean. We find this feature of the model quite appealing. It generates persistent volatility in combination with high volatility following economic shocks (even though dividend volatility has not changed). Both properties seem to capture important patterns in the data (e.g., Schwert (1989)). These observations deserve a fuller treatment, but we focus on predictability because of space limitations. See also Veronesi (1999).

VI. Simulations

The simulations explore predictability in the renewal model. To make the model more realistic, we assume that dividends follow a geometric random walk with time-varying growth:

$$\ln d_{t+1} = g_k + \ln d_t + \epsilon_{t+1}, \tag{44}$$

where $\epsilon_{t+1} \sim N[0, \sigma^2]$ and g_k is randomly drawn every K periods from a normal distribution with mean g^* and variance σ^2/s. Investors, who are risk neutral, must estimate g_k from observed dividends (Appendix B describes the Bayesian inference problem). The simulations normalize the initial dividend to equal one, the discount rate equals 0.12, $\sigma = 0.10$, and the long-run growth rate g^* equals 0.03. These parameters are chosen to be reasonably close to actual values, interpreting a period as one year. In comparison, from 1926 to 1997, dividends grew at an annual rate of 4.4 percent with a standard deviation of 12.0 percent. The CRSP value-weighted index had an annual return of 12.5 percent.

The simulations estimate predictive regressions of returns on lagged dividend yield (DY). In the model, DY and lagged growth rates both capture information about investors' beliefs (in the same way that dividends and prices do in the basic model). When past dividend growth has been high relative to g_k, subsequent growth is likely to be disappointing and true expected returns are, accordingly, low. We use DY as the predictive variable because it is more common and it provides the most timely information about expected growth. In fact, with a constant discount rate, changes in expected growth are the only source of variation in DY. Reversals show up as a positive slope in the regressions since DY moves inversely with growth rates. We report regressions using roughly 75 years of data, similar to a typical study, and for several combinations of the parameters s, h, and K. These parameters determine the true variance in short-run growth rates, the variance of investors' priors, and the length of a regime, respectively.

Table I reports the average slope coefficient and average t-statistic from 2,500 simulations. An important complication arises because the slope coefficient is biased upward in small samples (see Stambaugh (1999)). The bias is caused by contemporaneous correlation between returns and DY. To correct for this bias, we estimate auxiliary regressions of *true* unexpected returns on lagged DY. These regressions use exactly the same dividend and price series, but we subtract out from each return its true conditional expectation. The difference between the slopes in the raw and auxiliary regressions shows how time variation in expected returns (created by parameter uncertainty) affects the predictive slope. This is exactly what we are trying to measure. The difference is reported as a bias-adjusted slope in Table I. We have also corrected for bias using Stambaugh's (1999) approximation, with similar results.

Table I shows that learning can have a large effect on empirical tests. Even when investors know both the mean and variance of the growth

Table I

Predictability in Steady State

We simulate the renewal model 2,500 times and estimate predictive regressions of excess returns on lagged dividend yield. Dividends are assumed to follow a geometric random walk with time-varying expected growth, where the short-run growth rate g_k is randomly drawn every K periods from $N[g^*, \sigma^2/s]$. Investors are risk neutral and have beliefs about g_k at the beginning of each regime equal to $N[g^*, \sigma^2/h]$. In the simulations, $r = 0.12$, $\sigma = 0.10$, and $g^* = 0.03$. The table reports, for various combinations of s, h, and K, the average slope coefficient, bias-adjusted slope coefficient, and t-statistic from the regressions. The bias correction is described in the text.

		Estimation Risk			Perfect Information		
		h			h		
	s	16	25	49	16	25	49
2 regimes (K = 38)							
Slope	16	2.91	3.05	3.68	0.40	0.39	0.50
	25	3.31	3.57	4.41	0.37	0.31	0.52
	49	3.92	4.16	5.62	0.44	0.50	0.41
Bias-adj slope	16	**0.74**	**0.36**	**−0.37**	0.00	0.00	0.00
	25	**0.95**	**0.70**	**0.09**	0.00	0.00	0.00
	49	**1.28**	**1.04**	**0.68**	0.00	0.00	0.00
T-statistic	16	1.02	0.80	0.53	0.25	0.25	0.28
	25	1.15	0.98	0.71	0.23	0.19	0.18
	49	1.34	1.14	0.94	0.11	0.14	0.13
4 regimes (K = 19)							
Slope	16	2.56	2.44	1.97	0.37	0.30	0.34
	25	3.34	3.29	3.61	0.32	0.32	0.28
	49	4.01	4.21	5.25	0.35	0.35	0.27
Bias-adj slope	16	**0.42**	**−0.31**	**−2.04**	0.00	0.00	0.00
	25	**1.01**	**0.50**	**−0.90**	0.00	0.00	0.00
	49	**1.64**	**1.30**	**0.44**	0.00	0.00	0.00
T-statistic	16	0.62	0.38	0.05	0.24	0.20	0.22
	25	0.86	0.59	0.32	0.18	0.18	0.16
	49	1.05	0.80	0.56	0.14	0.14	0.11
6 regimes (K = 13)							
Slope	16	2.59	1.94	1.41	0.27	0.32	0.24
	25	3.65	3.44	3.65	0.31	0.28	0.25
	49	4.38	4.87	5.80	0.29	0.30	0.26
Bias-adj slope	16	**0.42**	**−0.77**	**−3.65**	0.00	0.00	0.00
	25	**1.28**	**0.39**	**−1.60**	0.00	0.00	0.00
	49	**2.02**	**1.56**	**0.36**	0.00	0.00	0.00
T-statistic	16	0.48	0.19	−0.05	0.17	0.20	0.16
	25	0.42	0.45	0.22	0.16	0.15	0.13
	49	0.90	0.71	0.44	0.11	0.10	0.10

rate ($h = s$), the coefficient on DY is positive. With two regimes, the slope coefficient ranges from 0.68 when $h = s = 49$ (g_k has a standard deviation of 1.4 percent) to 0.74 when $h = s = 16$ (g_k has a standard deviation of

2.5 percent). With four regimes, the slope coefficient ranges from 0.42 to 0.50, and with six regimes, the slope ranges from 0.36 to 0.42. Reversals become much more pronounced when investors' prior variance is higher than actual variance. For example, if the prior variance is 2.5 percent ($h = 16$) but the true variation in g_k is only 1.4 percent ($s = 49$), the predictive slope varies from 1.28 with two regimes to 2.02 with six regimes. Cases in which $s > h$, so the subjective variance is greater than the true, are of particular interest because they show roughly how prices behave before we reach steady state (even if investors know σ_s^2, the subjective variance of dividends is always greater than the true after a finite number of periods). We suspect that the evolutionary process is as relevant for empirical tests as the long-run equilibrium.

The table also shows that continuations—a negative slope coefficient in the DY regressions—are possible if $h > s$. These cases represent an economy that is changing more rapidly than investors realize. In fact, the average slope coefficient can be quite negative, up to -3.65 when there are six regimes, if investors substantially understate the true volatility of g_k. Investors require many observations until their beliefs catch up with the actual changes, which creates persistence in expected returns. As mentioned earlier, Lewis (1989) argues that this sort of phenomenon characterized the foreign exchange market in the 1980s.

To add some perspective, the historical slope coefficient for the period 1941 to 1997 is 3.93 (t-statistic of 2.27), before adjusting for bias, when the CRSP value-weighted return is regressed on lagged DY. The simulations suggest that parameter uncertainty could account for a nontrivial portion of the predictability. We hesitate to draw firm conclusions because the simulations do not (and probably cannot) capture all of the relevant properties of actual dividends and returns. It is beyond the scope of the current paper to understand which parameters best characterize the stock market.

Finally, we note that adding a regime dummy variable to the regressions produces an estimate of the conditional covariance. In results not reported, the average bias-adjusted slope coefficient is approximately 1.47 with two regimes, 2.00 with four regimes, and 2.55 with six regimes. These values are not sensitive to the values of h and s, presumably because h and s affect the covariance in the numerator and the variance in the denominator by similar magnitudes. Although we believe these issues deserve a more complete treatment, we simply note here that the simulations confirm, in substance, our earlier results. Even in the long run, parameter uncertainty can be a source of predictability in empirical tests. Investors can neither perceive, nor take advantage of, this predictability.

VII. Summary and Conclusions

Financial economists generally assume that, unlike themselves, investors know the means, variances, and covariances of the cash-flow process. Practitioners do not have this luxury. To apply the elegant framework of modern portfolio theory, they must estimate the process using whatever information is available. However, as Black (1986) so memorably observes, the world is a noisy place; our observations are necessarily imprecise. The literature on estimation risk formalizes this problem, but it seems to have had little impact on mainstream thinking about asset pricing and market efficiency. We believe that this is due, in part, to its focus on the subjective beliefs of investors rather than the empirical properties of returns.

We show that learning can significantly affect asset-pricing tests. Prices in our model satisfy commonly accepted notions of market efficiency and rational expectations: Investors use all available information when making decisions and, in equilibrium, the perceived pricing function equals the true pricing function. In spite of this, the empirical properties of returns differ significantly from the properties perceived by investors. Excess returns can appear to be predictable even though investors perceive a constant risk premium; prices can appear to be too volatile even though all investors are rational; and the CAPM can fail to describe returns even though investors attempt to hold mean-variance efficient portfolios. Put simply, empirical tests can find patterns in returns that rational investors can neither perceive nor exploit.

It is important to note that predictability is not due to some spurious estimation problem. Rather, it is a feature of the true data-generating process. This means that parameter uncertainty can affect empirical tests in surprising ways. We find, for example, that an *implementable* market-timing strategy might generate abnormal profits. An econometrician replicating the strategy in real time (using past data) is expected to find, in a frequentist sense, risk-adjusted profits. Again, however, a rational investor does not gain anything from following the strategy. The investors' perceived profit is zero. (A similar phenomenon explains why investors cannot use cross-sectional predictability to beat the market.) This puzzle highlights the distinction between the repeated-sampling perspective of empirical tests and the conditional perspective of investment decisions (conditional on a given realization of dividends). It also shows how difficult it might be to construct valid asset-pricing tests in the presence of parameter uncertainty.

The fact that parameter uncertainty might explain observed asset-pricing anomalies does not, of course, mean that it does. Our simulations suggest that learning might be important, but empirical tests are necessary to draw strong conclusions (see, also, Brav and Heaton (2001)). To assess market efficiency, the researcher may, in effect, need to mimic the Bayesian-updating process of rational investors to determine whether the patterns observed in the data could have been exploited. This is not an easy task: It would necessarily require some judgment about the learning process and what constitutes a "reasonable" prior. This observation is reminiscent of Fama's (1970) critique of asset-pricing tests. He emphasizes that empirical tests always entail a joint hypothesis of market efficiency and a model of perceived expected returns. Our study suggests that empirical tests may also require an assumption about prior beliefs. The role of prior beliefs and learning is typically ignored, but it might be critical for understanding anomalies.

It is tempting to compare parameter uncertainty to data mining. In recent years, researchers have become increasingly sensitive to the possibility that, with the intensive scrutiny of data, statistically "significant" return patterns can emerge even when returns are random (e.g., Lo and MacKinlay (1990)). Thus, we might observe patterns that do not exist in the true underlying process. Our analysis of parameter uncertainty suggests a complementary concern. With hindsight, we can discern patterns that existed in the true return process but could not have been exploited by rational investors. Similar to data snooping, these patterns would not be relevant for investment decisions. Unlike data snooping, the patterns can persist in the future because they are part of the true data-generating process. This conclusion provides an alternative perspective on empirical anomalies. For example, Fama (1998) argues that various long-horizon return anomalies in the literature are chance results, consistent with market efficiency. Our work reinforces his point. Reversals and continuations might be expected in an efficient market with parameter uncertainty, not only as a random outcome of the data, but as a feature of the actual process.

Appendix A: Equilibrium in the Basic Model

Beliefs about the dividend process are given by equation (7) in the text. Rational expectations requires that investors anticipate how prices are determined in the future. This is imposed by recursively substituting for p_{t+k} in (4), yielding

$$p_t = \frac{1}{r}\bar{d}_t - 2\gamma\left[\sum_{k=1}^{\infty}\frac{1}{(1+r)^k}E_t^s \mathrm{var}_{t+k-1}^s(p_{t+k}+d_{t+k})\right]\cdot\iota. \tag{A1}$$

Equation (A1) assumes that $\lim_{k\to\infty}E_t[p_{t+k}]/(1+r)^k = 0$, which is satisfied in equilibrium. Price is a function of expected dividends and the expected conditional variance of gross returns. The conditional variance is likely to change over time since parameter uncertainty vanishes in the limit. However, if price is a linear function of $\bar{d}_t$, the conditional variance should be deterministic. Price volatility is driven entirely by the first term and the subjective variance is

$$\mathrm{var}_t^s(p_{t+1}+d_{t+1}) = \left[1+\frac{1}{r(t+1)}\right]^2\left(\frac{t+1}{t}\right)\Sigma. \tag{A2}$$

Substituting into (A1) yields the equilibrium pricing function:

$$p_t = \frac{1}{r}\bar{d}_t - 2\gamma f(t)\Sigma\iota, \tag{A3}$$

where

$$f(t) = \sum_{k=1}^{\infty} \frac{1}{(1+r)^k}\left[1 + \frac{1}{r(t+k)}\right]^2\left(\frac{t+k}{t+k-1}\right). \tag{A4}$$

The function $f(t)$ decreases as t gets larger and converges to $1/r$ in the limit. Price is a linear function of $\bar{d}_t$, so the investor's wealth is MVN (which we assumed to derive the optimal portfolio choice) and conditional volatility is deterministic (which we assumed above).

Appendix B: Bayesian Inference Problem for the Simulations

The simulations assume that dividends follow a geometric random walk with time-varying growth:

$$\ln d_{t+1} = g_k + \ln d_t + \epsilon_{t+1}, \tag{B1}$$

where $\epsilon_{t+1} \sim N[0, \sigma^2]$. The growth rate g_k is randomly drawn every K periods from a normal distribution with mean g^* and variance σ^2/s. At the beginning of a regime, investors' prior beliefs about g_k are $N[g^*, \sigma^2/h]$. After t periods in a regime ($t \le K$), investors beliefs about g_k are $N[c_t, \sigma^2_{c,t}]$, where

$$c_t = \frac{h}{t+h} g^* + \frac{t}{t+h}\frac{1}{t}\Sigma_{i=1}^{t} \Delta \ln d_i, \tag{B2}$$

$$\sigma^2_{c,t} = \frac{1}{t+h}\sigma^t. \tag{B3}$$

The predictive belief about log dividends next period is normally distributed with mean $c_t + \ln d_t$ and variance $[(t+h+1)/(t+h)]\sigma^2$. Actual dividends are log-normally distributed. Converting the expectations about log dividends into actual dividends, and extending the results to any dividend in the next q periods, where $t + q \le K$ (that is, dividends in the current regime), the predictive distribution of dividends is log-normal with mean

$$E_t^s[d_{t+q}] = d_t \exp\left[c_t + \frac{1}{2}q\sigma^2 + \frac{1}{2}q^2\sigma^2_{c,t}\right]. \tag{B4}$$

This equation recognizes that changes in log dividends are correlated with changes in beliefs about the growth rate. In other words, investors recognize that their beliefs, both the mean and the variance, will evolve over time. After the end of the current regime, investors expect dividends to grow once again at the rate g^*, and the variance of the growth rate is σ^2/h. Therefore, to derive beliefs about long-run dividends requires two steps: (1) Take the expectation conditional on the realized dividend at the end of the current regime, d_k, and (2) take the expectation conditional only on the current dividend, d_t. Details available on request.

REFERENCES

Barberis, Nicholas, 2000, Investing for the long run when returns are predictable, *Journal of Finance* 55, 225–264.

Bawa, Vijay, and Stephen Brown, 1979, Capital market equilibrium: Does estimation risk really matter?, in Vijay Bawa, Stephen Brown, and Roger Klein, eds.: *Estimation Risk and Optimal Portfolio Choice* (North-Holland, Amsterdam).

Bawa, Vijay, Stephen Brown, and Roger Klein, 1979, *Estimation Risk and Optimal Portfolio Choice* (North-Holland, Amsterdam).

Berger, James, 1985, *Statistical Decision Theory and Bayesian Analysis* (Springer-Verlag, New York).

Black, Fischer, 1986, Noise, *Journal of Finance* 41, 529–543.

Bossaerts, Peter, and Pierre Hillion, 1999, Implementing statistical criteria to select return forecasting models: What do we learn? *Review of Financial Studies* 12, 405–428.

Brav, Alon, and John B. Heaton, 2001, Competing theories of financial anomalies, Working paper, Duke University.

Brennan, Michael, and Yihong Xia, 1998, Stock price volatility, learning, and the equity premium, Working paper, University of California at Los Angeles.

Campbell, John, 1991, A variance decomposition for stock returns, *The Economic Journal* 101, 157–179.

Clarkson, Peter, Jose Guedes, and Rex Thompson, 1996, On the diversification, observability, and measurement of estimation risk, *Journal of Financial and Quantitative Analysis* 31, 69–84.

Coles, Jeffrey, and Uri Loewenstein, 1988, Equilibrium pricing and portfolio composition in the presence of uncertain parameters, *Journal of Financial Economics* 22, 279–303.

Coles, Jeffrey, Uri Loewenstein, and Jose Suay, 1995, On equilibrium pricing under parameter uncertainty, *Journal of Financial and Quantitative Analysis* 30, 347–364.

DeLong, J. Bradford, Andrei Shleifer, Lawrence Summers, and Robert Waldmann, 1990, Noise trader risk in financial markets, *Journal of Political Economy* 98, 703–738.

Detemple, Jerome, 1986, Asset pricing in a production economy with incomplete information, *Journal of Finance* 41, 383–391.

Dothan, Michael, and David Feldman, 1986, Equilibrium interest rates and multiperiod bonds in a partially observable economy, *Journal of Finance* 41, 369–382.

Fama, Eugene, 1970, Efficient capital markets: A review of theory and empirical work, *Journal of Finance* 25, 383–417.

Fama, Eugene, 1976, *Foundations of Finance* (Basic Books, New York).

Fama, Eugene, 1998, Market efficiency, long-term returns, and behavioral finance, *Journal of Financial Economics* 49, 283–306.

Fama, Eugene, and Kenneth French, 1989, Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics* 25, 23–49.

Fama, Eugene, and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

Fama, Eugene, and Kenneth French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55–84.

Fama, Eugene, and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Gennotte, Gerard, 1986, Optimal portfolio choice under incomplete information, *Journal of Finance* 41, 733–746.

Gibbons, Michael, Stephen Ross, and Jay Shanken, 1989, A test of the efficiency of a given portfolio, *Econometrica* 57, 1121–1152.

Goyal, Amit, and Ivo Welch, 1999, Predicting the equity premium, Working paper, Yale School of Management.

Greene, William, 2000, *Econometric Analysis*, 4th edition (Prentice Hall, Upper Saddle River, NJ).

Harvey, Campbell, 1989, Time-varying conditional covariances in tests of asset pricing models, *Journal of Financial Economics* 24, 289–317.

Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65–91.

Kandel, Shmuel, and Robert Stambaugh, 1996, On the predictability of stock returns: An asset allocation perspective, *Journal of Finance* 51, 385–424.

Lakonishok, Josef, Andrei Shleifer, and Robert Vishny, 1994, Contrarian investment, extrapolation, and risk, *Journal of Finance* 49, 1541–1578.

LeRoy, Stephen, and Richard Porter, 1981, The present value relation: Tests based on implied variance bounds, *Econometrica* 49, 555–574.

Lewellen, Jonathan, 2001, Predicting returns with financial ratios, Working paper, MIT Sloan School of Management.

Lewellen, Jonathan, and Jay Shanken, 2000, Market efficiency, estimation risk, and the predictability of returns, Working paper, MIT and the University of Rochester.

Lewis, Karen, 1989, Changing beliefs and systematic rational forecast errors with evidence from foreign exchange, *American Economic Review* 79, 621–636.

Lo, Andrew, and A. Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431–467.

Lucas, Robert, 1978, Asset prices in an exchange economy, *Econometrica* 46, 1429–1446.

Merton, Robert, 1971, Optimum consumption and portfolio rules in a continuous-time model, *Journal of Economic Theory* 3, 373–413.

Pastor, Lubos, 2000, Portfolio selection and asset pricing models, *Journal of Finance* 55, 179–223.

Schwert, G. William, 1989, Why does stock market volatility change over time? *Journal of Finance* 44, 1115–1153.

Shanken, Jay, 1990, Intertemporal asset pricing: An empirical investigation, *Journal of Econometrics* 45, 99–120.

Shiller, Robert, 1981, Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review* 7, 421–436.

Stambaugh, Robert, 1999, Predictive regressions, *Journal of Financial Economics* 54, 375–421.

Stulz, René, 1987, An equilibrium model of exchange rate determination and asset pricing with nontraded goods and imperfect information, *Journal of Political Economy* 95, 1024–1040.

Timmermann, Allan, 1993, How learning in financial markets generates excess volatility and predictability in stock prices, *Quarterly Journal of Economics* 108, 1135–1145.

Timmermann, Allan, 1996, Excess volatility and predictability of stock prices in autoregressive dividend models with learning, *Review of Economic Studies* 63, 523–557.

Veronesi, Pietro, 1999, Stock market overreaction to bad news in good times: A rational expectations equilibrium model, *Review of Financial Studies* 12, 975–1007.

Wang, Jiang, 1993, A model of intertemporal asset prices under asymmetric information, *Review of Economic Studies*, 60, 249–282.

Williams, Joseph, 1977, Capital asset prices with heterogeneous beliefs, *Journal of Financial Economics* 5, 219–239.

Xia, Yihong, 2001, Learning about predictability: The effects of parameter uncertainty on dynamic asset allocation, *Journal of Finance* 56, 205–246.