Learning in Financial Markets

Lubos Pastor and Pietro Veronesi

Booth School of Business, University of Chicago, Chicago,
Illinois 60637; email: lubos.pastor@chicagobooth.edu,
pietro.veronesi@chicagobooth.edu

Key Words

Bayesian, uncertainty, bubble, volatility, predictability

Abstract

We survey the recent literature on learning in financial markets.
Our main theme is that many financial market phenomena that
appear puzzling at first sight are easier to understand once we
recognize that parameters in financial models are uncertain and
subject to learning. We discuss phenomena related to the volatility
and predictability of asset returns, stock price bubbles, portfolio
choice, mutual fund flows, trading volume, and firm profitability,
among others.

INTRODUCTION

Parameter uncertainty is ubiquitous in finance. Agents are uncertain about many of the parameters characterizing financial markets, and they learn about these parameters by observing data. This learning is facilitated by the existence of vast quantities of financial data, but it is also hampered by the large amount of randomness pervading financial markets.

This article reviews selected recent work on learning in finance. The overarching theme is that learning helps us better understand a variety of phenomena observed in financial markets. Many facts that appear baffling at first sight seem less puzzling once we recognize that parameters are uncertain and subject to learning. We ask questions such as, Why are stock returns so volatile? Why are they predictable? Why do investors trade so much? Why do stocks of young firms exhibit high valuations and high volatility? Why are technological revolutions accompanied by stock price "bubbles"? Why do fund flows react strongly to fund performance? Why do firms become less profitable after they go public? We show that learning helps us answer all of these questions, as well as many others.

Our quest for the answers is guided by the principle of parsimony. We always seek the simplest explanation, one that makes as few assumptions as possible. For example, a single-agent model is more parsimonious than a multiagent model, symmetric information is simpler than asymmetric information, and rationality has fewer degrees of freedom than irrationality. If a fact can be explained in a rational single-agent model, then it can surely be explained in more complicated models as well. Of course, many facts cannot be explained with few assumptions. Nonetheless, the world appears a lot more parsimonious once parameter uncertainty is acknowledged.

BAYESIAN UPDATING

Bayes' rule: theorem showing how prior beliefs are revised into posterior beliefs after receiving new information

The cornerstone of learning is Bayes' rule, which describes how rational agents update their beliefs after receiving new information. To illustrate the updating process, consider the following example of an agent who is uncertain about the parameter θ. Before observing any signals, the agent's prior beliefs about θ are normally distributed with mean θ_0 and variance σ_0^2. The agent observes T independent signals about θ, $s_t = \theta + \varepsilon_t$, where each ε_t is normal with zero mean and known variance σ^2. According to Bayes' rule, the agent's posterior (i.e., revised) beliefs about θ are normally distributed with mean $\tilde{\theta}_T$ and variance $\tilde{\sigma}_T^2$, where

$$\tilde{\theta}_T = \theta_0 \frac{\frac{1}{\sigma_0^2}}{\frac{1}{\sigma_0^2} + \frac{T}{\sigma^2}} + \bar{s} \frac{\frac{T}{\sigma^2}}{\frac{1}{\sigma_0^2} + \frac{T}{\sigma^2}}, \tag{1}$$

$$\tilde{\sigma}_T^2 = \frac{1}{\frac{1}{\sigma_0^2} + \frac{T}{\sigma^2}}, \tag{2}$$

and $\bar{s}$ is the average signal value, $\bar{s} = (1/T)\sum_{t=1}^{T} s_t$. The posterior mean $\tilde{\theta}_T$ is a precision-weighted average of the prior mean and the average signal. Unlike $\tilde{\theta}_T$, the posterior variance $\tilde{\sigma}_T^2$ does not depend on the realizations of the signals. This variance, which we also refer to as uncertainty about θ, decreases as the number of signals T increases

(learning reduces uncertainty). The posterior variance is always smaller than the prior variance, $\tilde{\sigma}_T^2 < \sigma_0^2$.

Bayesian updating can also be formulated recursively. Denoting $\Delta \tilde{\theta}_t = \tilde{\theta}_t - \tilde{\theta}_{t-1}$, we have

$$\Delta \tilde{\theta}_t = m_t \left(s_t - \tilde{\theta}_{t-1} \right) \text{ with } m_t = \frac{1}{1 + \sigma^2 / \tilde{\sigma}_{t-1}^2}. \tag{3}$$

Intuitively, observing a higher-than-expected signal, $s_t > \tilde{\theta}_{t-1}$, leads the agent to revise the expectation upward, $\tilde{\theta}_t > \tilde{\theta}_{t-1}$, and vice versa. This revision is large when the multiplier m_t is large, which happens when the ratio of uncertainty $\tilde{\sigma}_{t-1}^2$ to signal variance σ^2 is large.

If time is viewed as continuous rather than discrete, the signal takes the differential form, $ds_t = \theta dt + \sigma dW_t$, where dW_t denotes a Brownian motion. The updating formula is then

$$d\tilde{\theta}_t = m_t \left(ds_t - \tilde{\theta}_t dt \right) \text{ with } m_t = \frac{\tilde{\sigma}_t^2}{\sigma^2}, \tag{4}$$

which is analogous to Equation 3. Note that m_t in Equation 3 can also be written as $\tilde{\sigma}_t^2 / \sigma^2$. Even in continuous time, uncertainty $\tilde{\sigma}_t^2$ declines over time according to the same Equation 2.

Uncertainty: variance of an unknown parameter given the available information

Brownian motion: continuous-time stochastic process with independent normally distributed increments

P/D: price-to-dividend ratio

STOCK VALUATION

When the discount rate r and dividend growth g are constant, the stock price is given by

$$P = \frac{D}{r - g}, \tag{5}$$

where D is the next period's dividend. This well-known Gordon growth formula holds not only when dividend growth is constant but also when it follows the process

$$\frac{dD_t}{D_t} = g\, dt + \sigma\, dW_t, \tag{6}$$

in which case g represents average dividend growth. See the Appendix for proof.

Interesting things happen when g in Equation 6 is unknown. Pastor & Veronesi (2003, 2006) argue that uncertainty about g increases the stock price. The Appendix shows that for any probability density $f(g)$ such that $r > g$ with probability one,

$$P = E\left\{ \frac{D}{r - g} \right\} > \frac{D}{r - E\{g\}}, \tag{7}$$

where E{.} denotes an expectation with respect to $f(g)$. The inequality in Equation 7 follows from Jensen's inequality because $1/(r - g)$ is convex in g. For the same reason, the price-to-dividend ratio (P/D) increases with the dispersion of $f(g)$. Intuitively, uncertainty about g makes the distribution of future dividends right-skewed, thereby increasing expected future dividends. Loosely speaking, a firm with some probability of failing (a very low g) and some probability of becoming the next Google (a very high g) is very valuable. When r is endogenously determined in equilibrium with a power-utility representative agent, uncertainty about g may increase or decrease r, but its overall effect

M/B: market-to-book ratio

Implied uncertainty: level of uncertainty that equates model-implied valuations with observed valuations

Stock-price bubble: phenomenon occurring when stock prices exhibit a substantial increase followed by a sharp fall without being accompanied by an obvious similar pattern in the fundamentals

on *P/D* is positive (Pastor & Veronesi 2006). Instead of focusing on *P/D*, which does not exist for non-dividend-paying firms, Pastor & Veronesi focus on the market-to-book ratio (*M/B*). This ratio increases with uncertainty about the firm's average profitability, which can be interpreted as uncertainty about the average growth rate of book value.

Because uncertainty declines over time due to learning (see Equation 2), the Pastor & Veronesi (2003) model predicts that *M/B* declines over a typical firm's lifetime, so that younger firms should have higher *M/B*'s than otherwise identical older firms. This prediction is confirmed in U.S. stock data: The median *M/B* falls monotonically from 2.25 for 1-year-old firms to 1.25 for 10-year-old firms, and the cross-sectional relation between firm age and *M/B* is reliably negative. The model also implies that the effect of age on *M/B* should be stronger for younger firms and non-dividend-paying firms. Furthermore, *M/B* should decrease with expected return and increase with both the level and the volatility of profitability. All of these predictions are confirmed empirically.

Stock Price "Bubbles"

Pastor & Veronesi (2006) extend their 2003 model and calibrate it to match the observed stock valuations at the peak of the Nasdaq "bubble." They argue that stocks were not necessarily overvalued in the late 1990s because uncertainty about *g* was unusually high. The higher the uncertainty about *g*, the higher the stock price in Equation 7. The authors compute the level of uncertainty that allows their model to match the Nasdaq valuation at its peak in March 2000. This uncertainty, which they call implied uncertainty for its similarity to implied volatility in option pricing, seems plausible because it matches not only the level but also the volatility of Nasdaq stock prices. In the late 1990s, these prices were not only high but also highly volatile, and both facts are consistent with high uncertainty about *g*. (We show later that uncertainty about *g* increases return volatility.) Moreover, cross-sectionally, stocks with high *M/B*'s also had highly volatile returns, suggesting that these stocks had highly uncertain future growth rates. In general, the authors argue that the level and volatility of stock prices are positively linked through firm-specific uncertainty about *g*.

The same learning model also seems capable of explaining the bursting of the Nasdaq bubble. Nasdaq's profitability plummeted in 2000 and 2001. As a result, investors revised their expectations of Nasdaq's future profitability downward, pushing prices down. Because the investors' prior uncertainty was large, their expectation revision was also large (see Equation 3). Starting with prior beliefs that match Nasdaq's level and volatility in March 2000, the model predicts a postpeak Nasdaq price decline that is comparable to that observed in the data.

The Nasdaq bubble, which developed during the Internet boom, is an example of a more general pattern. Technological revolutions tend to be accompanied by bubbles in the stock prices of innovative firms. This evidence is typically attributed to market irrationality, but Pastor & Veronesi (2009) argue that it is also consistent with a rational general equilibrium model of learning. They argue that new technologies are characterized by high uncertainty about their future productivity, and that the time-varying nature of this uncertainty can produce the observed bubbles. In their model, a representative agent is learning about a new technology's productivity. If the agent learns that the technology is sufficiently productive, he adopts it on a large scale, creating a technological revolution. Most new technologies do not cause revolutions, but those that do exert two opposing effects

on stock prices: a positive cash flow effect and a negative discount rate effect. On the one hand, the new technology must surprise the agent with high realized productivity (otherwise he would not adopt it), and this positive cash-flow news pushes stock prices up. On the other hand, the risk associated with the new technology gradually changes from idiosyncratic to systematic, thereby pushing up discount rates and thus depressing stock prices. The cash flow effect prevails initially, but the discount rate effect prevails eventually, producing an apparent bubble in stock prices. Importantly, these bubbles are observable only in hindsight—they are unexpected by investors in real time but we observe them ex post when we focus only on technologies that eventually lead to technological revolutions.

The Pastor-Veronesi model makes numerous additional predictions, which are supported by the evidence from 1830–1861 and 1992–2005 when the railroad and Internet technologies spread in the United States. A key prediction is that the market beta of innovative firms—a measure of systematic risk—should increase during technological revolutions. Indeed, the beta of the technology-loaded Nasdaq index doubled between 1997 and 2002, and the beta of railroad stocks increased sharply in the 1850s. Because stories based on irrationality do not predict increases in systematic risk during revolutions, this evidence suggests that rational learning about new technologies is useful in explaining the bubble-like patterns in stock prices.

Technological revolutions exhibit not only stock price bubbles but also apparent over-investment. This fact is also consistent with rational learning, as shown by Johnson (2007). Johnson develops an equilibrium model of investment in a new industry whose production function has an unknown return to scale. The model implies that the most efficient way to learn about returns to scale is by overinvestment relative to the full-information case. This overinvestment is accompanied by high stock prices and low expected returns.

Other models that link stock price bubbles to learning include Scheinkman & Xiong (2003) and Hong et al. (2006, 2008). Unlike the models discussed above, these models feature heterogeneous beliefs, and they produce bubbles with the help of additional assumptions such as short-sale constraints and behavioral biases. Battalio & Schultz (2006) argue that short-sale constraints were not responsible for the Nasdaq bubble. Li & Xue (2008) argue that this bubble can be explained by uncertainty about a possible structural break in the economy's productivity. Finally, Donaldson & Kamstra (1996) argue against a bubble in the 1920s based on a neural network model of dividend expectations.

STOCK RETURN VOLATILITY

The volatility of stock returns exhibits interesting empirical features. For example, it is high relative to the volatility of the underlying dividends, and it varies over time in a persistent fashion. Learning helps us understand these facts.

The Level of Volatility

When the discount rate r and the average dividend growth g are both constant and known, the stock price is given by Equation 5, and return volatility equals the volatility of dividend growth. In reality, however, the postwar volatility of market returns has averaged 17% per year, whereas the dividend growth volatility has averaged only 5%. To reconcile this difference, it helps to view g in Equation 6 as unknown (Timmermann 1993). Agents

learn about g by observing realized dividends. Unexpectedly high dividends increase the stock price not only through current dividends but also by raising expectations of future dividends. This double kick to the stock price increases return volatility compared to the case in which g is known.

To formalize Timmermann's intuition, let r be constant and known, let g have a truncated normal distribution that assigns zero probability to $g \geq r$, and let $\tilde{g}_t$ and $\tilde{\sigma}_t^2$ denote the mean and variance of g as perceived at time t. Extending Timmermann's work, we show in the Appendix that the standard deviation of returns is approximately equal to

$$\text{Return Volatility} \approx \text{Dividend Growth Volatility} \times \left[1 + \left(\frac{\partial \log(P/D)_t}{\partial \tilde{g}_t}\right) m_t\right], \quad (8)$$

where dividend-growth volatility stands for σ in Equation 6, $\partial \log(P/D)_t / \partial \tilde{g}_t > 0$, and $m_t > 0$ is given in Equation 4. This formula shows that return volatility exceeds the volatility of dividend growth. The difference can be substantial. For example, let $\sigma = 5\%$, $r = 10\%$, $\tilde{g}_t = 3\%$, and $\tilde{\sigma}_t = 2\%$. Return volatility is then approximately 20%, four times higher than the 5% volatility of dividend growth. Equation 8 also shows that return volatility increases with σ, and it also increases with uncertainty $\tilde{\sigma}_t^2$, through m_t. If $\tilde{\sigma}_t^2 \to 0$, then $m_t \to 0$, and return volatility converges to σ. Finally, return volatility increases with the sensitivity of log P/D to $\tilde{g}_t$. This sensitivity is higher when the discount rate is lower because distant future dividends then matter more for today's stock price.

The key implications of the simple model used above carry over to more sophisticated models. For example, Brennan & Xia (2001a) consider a general equilibrium model with a representative agent who learns about time-varying g_t. They obtain results similar to ours in a model successfully calibrated to aggregate consumption and dividend data.

Uncertainty $\tilde{\sigma}_t^2$ declines over time as investors learn about g (see Equation 2), so return volatility should decline over time as well (see Equation 8). One might therefore expect the stocks of younger firms to have more volatile returns than the stocks of older firms. Indeed, Pastor & Veronesi (2003) find a negative cross-sectional relation between volatility and firm age. The median return volatility of U.S. stocks falls monotonically from 14% per month for 1-year-old firms to 11% per month for 10-year-old firms. The authors' model predicts higher stock volatility for firms with more volatile profitability, firms with more uncertain average profitability, and firms that pay no dividends. These predictions are confirmed empirically.

Time Variation in Volatility

Stock return volatility varies dramatically over time—it has been as low as 10% per year in the mid-1990s and as high as 70% in October 2008. Moreover, volatility is persistent, as there are extended periods of sustained high or low volatility. Learning helps us understand the variation in volatility. The models of Timmermann (1993) and Pastor & Veronesi (2003) cannot generate increases in volatility because they feature a constant g, and uncertainty about a constant g declines deterministically to zero (in Equation 2, $\tilde{\sigma}_T^2 \to 0$ as $T \to \infty$). Even when g varies over time in a smooth manner, as in Brennan & Xia (2001a), the posterior uncertainty about g converges deterministically to a constant. However, if g follows a process with unobservable regime shifts, then uncertainty about g can fluctuate stochastically, and return volatility can rise. For example, if a dividend growth realization

is far from the current estimate of g, the probability of a regime shift in g increases. The posterior uncertainty about g then increases because after a regime shift, past data become less useful for forecasting. The higher uncertainty pushes up return volatility through a mechanism similar to that in Equation 8: Investors' expectations react more swiftly to news when uncertainty is higher. Moreover, volatility is persistent because perceptions of regime shifts change slowly.

David (1997) develops a model with unobservable regime shifts in the average productivities of linear technologies, which are subject to learning by a representative agent. Learning induces time-varying allocations to these technologies, resulting in persistent stochastic variation in return volatility. Veronesi (1999) uses similar means to show that even if dividends display low constant volatility, stock returns may possess high volatility with persistent variation. He also shows that learning about a regime-shifting g generates stock price overreaction to bad news in good times. Such news increases uncertainty about g, which might have shifted from a high-g to a low-g regime. This increase in uncertainty increases not only volatility but also the equilibrium discount rate, thereby amplifying the stock price drop. In a similar setting, Veronesi (2004) shows that a small probability of a long recession can induce volatility to cluster at high levels during recessions. Johnson (2001) shows that learning about the degree of persistence of fundamental shocks generates time-varying return volatility, as well as a novel relation between volatility and momentum. David & Veronesi (2002) employ unobservable regime shifts to explain the dynamics of option-implied volatility and skewness spreads. David & Veronesi (2008) develop a structural model for volatility forecasting that exploits learning-induced relations between volatility and price multiples. This model improves upon regression-based volatility forecasts.

RETURN PREDICTABILITY

Stock returns are somewhat predictable. When the aggregate P/D is low, future stock market returns tend to be high. Timmermann (1993, 1996) explains that such predictability can arise due to learning about g. When the current estimate of g, $\tilde{g}_t$, is below the true value of g, investors are pessimistic about future dividends, so P/D is low. The future returns are likely to be high, however, because $\tilde{g}_t$ is likely to be revised upward. As a result, low P/D forecasts high future returns.

This learning-induced predictability is observable only in hindsight, as explained by Lewellen & Shanken (2002). Returns appear predictable to econometricians analyzing historical data, but real-time investors cannot exploit this predictability. Learning drives a wedge between the distribution perceived by investors and the true distribution estimated by empirical tests. Lewellen & Shanken show that learning can also induce cross-sectional predictability. For example, econometricians may observe violations of the capital asset pricing model (CAPM) even if all real-time investors believe this model holds. Coles & Loewenstein (1988) argue that the CAPM should hold even with estimation risk, but that is true only for the perceived, not the empirical, distribution of returns.

Learning can also generate risk-driven predictability that is detectable by real-time investors. Veronesi (1999, 2000) shows how learning induces time-varying expected returns that are correlated with P/D. Massa & Simonov (2005) and Ozoguz (2009) argue that uncertainty is a priced risk factor in the cross section of stock returns.

Croce et al. (2006) show that learning helps explain the cross-sectional value effect. In their model, consumption growth has a small but persistent long-run component (see Bansal & Yaron 2004), as well as a transitory short-run component. Stocks that are more exposed to the long-run component command higher risk premia. Even though value stocks tend to have shorter-duration cash flows, they can exhibit more long-run risk and therefore higher risk premia than growth stocks. When the long-run component of consumption is unobservable, its optimal forecasts covary with short-run consumption shocks. Learning induces positive correlation between the long-run and short-run consumption risks.

Another cross-sectional puzzle that can be understood via learning is the negative relation between stock returns and the dispersion of analysts' earnings forecasts, as documented by Diether et al. (2002). The authors interpret their result as evidence of market frictions that preclude investors with pessimistic views from shorting stocks, which are then temporarily overvalued. Johnson (2004) delivers the same result in a frictionless rational learning model. He interprets dispersion as a proxy for uncertainty about asset value. After adding leverage to a model similar to Pastor & Veronesi (2003), Johnson shows that expected stock return decreases with this uncertainty. Equity is a call option on the levered firm's assets. More idiosyncratic uncertainty raises the option value, which lowers the stock's exposure to priced risk, thereby reducing the expected return. The model also predicts that the negative relation found by Diether et al. should be stronger for firms with more leverage. Johnson finds empirical support for this prediction.

Uncertainty about the value of a firm's assets also helps us understand credit spreads on corporate bonds. In structural models of corporate bond valuation à la Merton (1974), the firm's value follows an observable diffusion process. These models imply counterfactually small credit spreads for short-term bonds because they imply that the default probability over a short period is small. Uncertainty about firm value increases short-term credit spreads, as Duffie & Lando (2001) explain, because investors are uncertain about the nearness of current assets to the default-triggering level. Supporting this explanation, Yu (2005) finds empirically that firms with more accounting disclosure (and as a result less uncertainty about firm value) tend to have lower credit spreads, especially on short-term bonds. Finally, David (2008a) uses learning about unobservable regime shifts in the fundamentals to explain why the observed credit spreads are higher than spreads produced by Merton-like models calibrated to the observed default frequencies.

Empirical Bayesian studies that analyze return predictability include Kandel & Stambaugh (1996), Stambaugh (1999), Avramov (2002), Cremers (2002), Avramov & Chordia (2006), Pastor & Stambaugh (2009), Wachter & Warusawitharana (2009), and others.

THE EQUITY PREMIUM

Learning can help us understand the equity premium puzzle. Uncertainty about average dividend growth g can increase or decrease the equity premium. In Veronesi (2000), this uncertainty decreases the equity premium. Veronesi considers an endowment economy with a power-utility representative agent whose elasticity of intertemporal substitution (EIS) is below one. The agent consumes aggregate dividends. Bad news about dividends decreases not only current consumption but also expected future consumption, as the agent revises $\tilde{g}_t$ downward. The agent's desire to smooth consumption leads him to save more today and demand more stock, which cushions the decline in the stock price.

Therefore, learning about g decreases the covariance between stock returns and consumption growth compared to the case with known g. As a result, the equity premium is lower as well.

The opposite is obtained under different preferences. When EIS exceeds one, downward revisions in $\tilde{g}_t$ lead the agent to save less and demand less stock, resulting in a positive relation between uncertainty and the equity premium (e.g., Brandt et al. 2004 and Ai 2007). The relation is positive also when the agent has exponential utility (Veronesi 1999) and when dividends and consumption follow separate processes with correlated unobservable drift rates (Li 2005). Finally, if the agent learns about average consumption growth, the expected consumption growth varies over time. Such variation increases the equity premium under Epstein-Zin preferences (e.g., Bansal & Yaron 2004).

The equity premium is also affected by uncertainty about the volatility of consumption growth. Weitzman (2007) considers an endowment economy with unknown consumption volatility. He shows that the posterior distribution of consumption growth is fat-tailed, which induces a power-utility representative agent to demand a substantially higher equity premium compared to the case of known volatility. Lettau et al. (2008) also assume that consumption volatility is unobservable, but they allow it to jump between two states. They find empirically that the posterior probability of the low-volatility state increased in the 1990s, helping justify the stock price run-up in that period.

Learning can also generate higher equity premia when investors are averse to ambiguity (e.g., Cagetti et al. 2002, Leippold et al. 2008, Epstein & Schneider 2008). When investors worry about model misspecification, their learning must take into account the set of possible alternative models. Model uncertainty is penalized and investors maximize utility over worst-case beliefs. This cautious behavior increases the risk premia in equilibrium.

LEARNING ABOUT THE CONDITIONAL MEAN RETURN

The studies discussed in the previous section let investors learn about fundamentals and analyze the equilibrium implications for expected returns. Another way of relating learning to expected returns is to let investors-econometricians, who do not necessarily set prices, learn about expected returns by observing realized returns and other information.

Let r_{t+1} denote a return from time t to time $t + 1$. This return can be decomposed as

$$r_{t+1} = \mu_t + u_{t+1}, \tag{9}$$

where μ_t is the conditional expected return and u_{t+1} is the unexpected return with mean zero, conditional on all information at time t. In reality, investors observe only a subset of all information, so they do not observe the true value of μ_t. How do rational investors learn about μ_t from realized returns?

If the conditional mean is constant, $\mu_t = \mu$, the updating formula in Equation 3 applies: Unexpectedly high returns increase the posterior mean $\tilde{\mu}_t$, and vice versa. Under noninformative prior beliefs about μ ($\sigma_0 = \infty$ in Equation 1), $\tilde{\mu}_t$ is simply the historical sample mean.

If μ_t varies over time, however, the sample mean is no longer the best estimate of μ_t. Given a process for the unobservable μ_t, we obtain the best estimate of μ_t by optimal filtering. For example, if μ_t follows an AR(1) process with normal shocks,

Equity premium puzzle: standard economic models cannot explain the large gap between the average historical returns on U.S. stocks and government bonds

$$\mu_{t+1} = (1 - \beta)\bar{\mu} + \beta\mu_t + w_{t+1}, \tag{10}$$

Kalman filter:
recursive filter that
estimates the state of a
dynamic system from
a series of noisy
observations

then the Kalman filter implies that the best estimate of μ_t (and hence also the best forecast of r_{t+1}) is a weighted average of all past returns,

$$E(r_{t+1}|F_t) = \sum_{s=0}^{t-1} \kappa_s r_{t-s},$$

where F_t contains the full history of returns up to time t (see Pastor & Stambaugh 2009). The weights in this average, κ_s, crucially depend on ρ_{uw}, the correlation between unexpected returns, u_{t+1} in Equation 9, and innovations in expected returns, w_{t+1} in Equation 10. This correlation is likely to be negative because unexpected increases in discount rates tend to push prices down. If this correlation is sufficiently negative, then recent returns receive negative weights and more distant returns receive positive weights in computing the average.

To understand this result, suppose recent returns have been unusually high. On the one hand, one might think the expected return has risen because a high mean is more likely to generate high realized returns, and μ_t is persistent. On the other hand, one might think the expected return has declined because declines in expected returns tend to be accompanied by high realized returns. When ρ_{uw} is sufficiently negative, the latter effect outweighs the former and recent returns enter negatively when estimating the conditional expected return. At the same time, more distant past returns enter positively because they are more informative about the unconditional mean $\bar{\mu}$ than about recent changes in the conditional mean μ_t.

The above analysis assumes that the information set F_t consists only of past returns. However, investors might use more information to forecast returns. For example, investors might believe that μ_t is given by a linear combination of observable predictors x_t:

$$\mu_t = \alpha + \beta(x_t - \bar{x}), \tag{11}$$

where $\bar{x}$ is the unconditional mean of x_t. Viewing β as unobservable, Xia (2001) uses continuous-time filtering to derive an updating rule for $\tilde{\beta}_t = E[\beta|F_t]$. This rule features time-varying covariance between updates to $\tilde{\beta}_t$ and realized returns. The sign of the covariance depends on whether x_t is above or below $\bar{x}$. When x_t exceeds $\bar{x}$, an unusually high return implies that $\tilde{\beta}_t$ is revised upward, and vice versa.

The assumption in Equation 11 is unlikely to hold exactly. If the true mean μ_t is not a linear function of x_t, the updating rule for μ_t involves not only x_t but also past returns (Pastor & Stambaugh 2009). Past dividends can also be useful in estimating μ_t, as shown by van Binsbergen & Koijen (2008) and Rytchkov (2008). These studies exploit present value relations to estimate not only μ_t but also expected dividend growth rates.

Learning about μ_t also affects long-horizon return volatility. Let $r_{t,t+k} = r_{t+1} + r_{t+2} + \ldots + r_{t+k}$ denote the return in periods $t + 1$ through $t + k$. The variance of $r_{t,t+k}$ conditional on data available at time t, $Var(r_{t,t+k} | F_t)$, depends on uncertainty about μ_{t+j}. Consider the following example from Pastor & Stambaugh (2008): Suppose r_t's are independently and identically distributed with known variance σ^2 and unknown constant mean μ. Conditional on μ, the mean and variance of $r_{t,t+k}$ are $k\mu$ and $k\sigma^2$, respectively. An investor who knows μ faces the same per-period variance, σ^2, regardless of k. However, an investor who does not know μ faces variance that increases with k. Applying the variance decomposition,

$$\text{Var}(r_{t,t+k}|F_t) = \text{E}\{k\sigma^2|F_t\} + \text{Var}\{k\mu|F_t\} = k\sigma^2 + k^2\text{Var}\{\mu|F_t\}.$$

Because μ remains uncertain after seeing the data, $(1/k)\,\text{Var}(r_{t,t+k}|F_t)$ increases with k. Thus, an investor who believes that stock prices follow a random walk but who is uncertain about μ views stocks as riskier in the long run. When μ_t is time-varying, predictability induces both mean reversion, which reduces long-run variance, and additional uncertainty, which increases long-run variance. The overall effect, according to Pastor & Stambaugh (2008), is a higher per-period variance at longer horizons, contrary to conventional wisdom.

PORTFOLIO CHOICE

Investors appear to invest too little in stocks. Consider an investor with risk aversion γ who can invest in risky stocks and riskless T-bills. If the mean μ and variance σ of excess stock returns are both constant and known, the investor's optimal stock allocation is given by

$$\text{Myopic Demand} = \frac{\mu}{\gamma\sigma^2}. \tag{12}$$

Based on the historical estimates, $\mu = 7\%$ and $\sigma = 16\%$ per year, the optimal stock allocation is 273% for $\gamma = 1$, 91% for $\gamma = 3$, and 55% for $\gamma = 5$, but households typically invest much less in stocks. This fact could in part be due to learning, as explained below.

If μ is unobservable, investors learn about it by observing realized returns. Even though μ is constant, its posterior mean $\tilde{\mu}_t$ is not, and investors wish to hedge against learning that μ is low (Williams 1977, Detemple 1986, Dothan & Feldman 1986, Gennotte 1986). Gennotte (1986) shows that uncertainty about μ reduces the stock allocation, as the variation in $\tilde{\mu}_t$ generates a negative hedging demand. Following Merton (1971), investors tilt their portfolios to hedge against fluctuations in marginal utility induced by changes in the state variable $\tilde{\mu}_t$. The size of the hedging demand is

$$\text{Hedging Demand} = -\rho_{\tilde{\mu},r}\left(\frac{\sigma_{\tilde{\mu}}}{\sigma}\right)\frac{\partial U_W/\partial\tilde{\mu}}{W\partial U_W/\partial W}, \tag{13}$$

Hedging demand: deviation of the optimal allocation from the myopic allocation due to changes in the investment opportunity set

where $\rho_{\tilde{\mu},r}$ is the correlation between $d\tilde{\mu}_t$ (revisions in $\tilde{\mu}_t$) and instantaneous returns, $\sigma_{\tilde{\mu}}$ is the volatility of $d\tilde{\mu}_t$, σ is return volatility, and U_W is marginal utility with respect to wealth W. The sign of the hedging demand depends on γ. Under power utility with $\gamma > 1$, the hedging demand is negative because $\partial U_W/\partial W < 0$, $\partial U_W/\partial\tilde{\mu} < 0$, and learning about a constant μ induces $\rho_{\tilde{\mu},r} > 0$ (see Equation 4). Intuitively, a negative stock position (relative to the myopic demand) is a good hedge because it profits from unexpectedly low stock returns, which are accompanied by decreases in $\tilde{\mu}_t$ that increase marginal utility. The higher the uncertainty about μ, the higher the value of $\sigma_{\tilde{\mu}}$, and the more negative the hedging demand.

Brennan (1998) shows that the learning-induced hedging demand can be large. For example, with a μ estimate of 8.5%, prior uncertainty about μ of 4.5%, and volatility of $\sigma = 14\%$, an investor with $\gamma = 4$ and a 20-year investment horizon invests only 56% in the stock market, down from 102% when only the myopic demand is considered. Whereas Brennan assumes that μ is constant, Xia (2001) considers time-varying μ_t, with investors learning about the slope β in the predictive relation (Equation 11). The hedging demand now has two components. The first one, which is well understood outside the learning literature, stems from time variation in the predictor x_t. The second component

stems from learning about β, and it involves the covariance between returns and $\tilde{\beta}_t$, as discussed earlier. Xia shows that both hedging demands are economically important.

The portfolio literature under learning has been extended to multiple assets. Brennan & Xia (2001b) assess the importance of the value and size anomalies from the perspective of an investor who is uncertain whether the anomalies are genuine. They find the value anomaly attractive even after incorporating parameter uncertainty. Pastor (2000) provides similar evidence in a single-period context, and also finds the home bias anomaly significant from the investment perspective. Cvitanic et al. (2006) analyze how optimal allocations depend on the correlation between the assets' expected returns. This correlation reduces uncertainty by allowing learning across assets, but it also makes estimation risk more difficult to diversify. Another extension incorporates nonlinear dynamics of μ_t. David (1997) and Honda (2003) solve for optimal allocations when μ_t undergoes unobservable regime shifts. Guidolin & Timmermann (2007) study asset allocation when regime shifts affect not only the mean but the whole return distribution. They empirically identify four regimes and solve for the optimal allocation among four asset classes. They find that unobservable regimes have a large impact on asset allocation.

The learning models discussed above are set in continuous time. There is also a growing discrete-time portfolio literature that relies on Bayesian econometric techniques. This literature typically does not estimate learning-induced hedging demands, but it integrates portfolio choice with empirical estimation of the parameters of the return-generating process. Parameter uncertainty is incorporated by focusing on the predictive distribution of asset returns. Letting θ denote the unknown parameters and F_t denote the data available at time t, the predictive distribution of returns at time $t + k$ is given by

$$p(R_{t+k}|F_t) = \int p(R_{t+k}|\theta, F_t)p(\theta|F_t)d\theta, \qquad (14)$$

where $p(\theta \mid F_t)$ is the posterior distribution of θ. Investors maximize expected utility computed with respect to the predictive distribution. Early contributions to this literature include Zellner & Chetty (1965), Brown (1976), Klein & Bawa (1976), and Bawa et al. (1979). Recent contributions include Kandel & Stambaugh (1996), Barberis (2000), Pastor (2000), Pastor & Stambaugh (2000, 2002b), Tu & Zhou (2004, 2008), Avramov (2004), Brandt et al. (2005), Avramov & Chordia (2006), Avramov & Wermers (2006), Kan & Zhou (2007), and Wachter & Warusawitharana (2009). Whereas the early contributions used noninformative prior distributions, recent contributions increasingly emphasize informative priors motivated by economic theory. Other recent studies analyze portfolio choice of ambiguity-averse investors (e.g., Uppal & Wang 2003, Wang 2005, Garlappi et al. 2007).

INVESTOR BEHAVIOR

Mutual Fund Flows

The way investors allocate their capital to mutual funds might seem puzzling. For example, net capital flows into mutual funds respond positively to past fund performance, even though there is little persistence in performance. Also, the performance-flow relation is convex and stronger for younger funds.

Berk & Green (2004) show that these facts are consistent with rational learning. Their model makes three key assumptions. First, the fund managers' ability is unobservable, and investors learn about it by observing fund returns. Second, this ability exhibits decreasing

returns to scale. Third, rational investors compete for superior returns. To illustrate the model's implications, suppose that a given fund achieved higher-than-expected returns recently. From these returns, investors infer that the fund manager's ability is higher than they previously thought, and they allocate more capital to this fund. This additional capital reduces the fund's ability to generate abnormal returns due to decreasing returns to scale. Given perfect competition in the provision of capital, investors pour capital into the fund until its abnormal performance disappears. As a result, a fund that outperformed in the past will attract new money, but it will not outperform in the future.

The positive performance-flow relation is stronger for younger funds because recent returns of a younger fund represent a bigger portion of the fund's track record, and so they are more informative about the fund manager's ability. Put differently, investors are more uncertain about the ability of funds with shorter track records, so any signal about ability has a bigger impact on the investors' beliefs. The performance-flow relation is convex at least in part because investors expect underperforming funds to change their strategies (Lynch & Musto 2003). Therefore, poor past performance contains less information about future performance than does good past performance. As a result, fund flows are less sensitive to past performance when that performance is poor.

Dangl et al. (2008) extend the Berk-Green model to allow the management company to replace portfolio managers. They derive the optimal replacement strategy and examine fund flows and portfolio risk around manager replacements. Their model rationalizes several empirical facts: (*a*) Managers are more likely to be fired after poor performance, (*b*) manager turnover is more performance-sensitive for younger managers, (*c*) managers with longer tenure tend to manage larger funds and are more likely to retain their jobs, and (*d*) manager replacement is generally preceded by capital outflows and increases in portfolio risk then followed by inflows and decreases in risk. Taylor (2008) develops a related model in which a board of directors learns about CEO skill and repeatedly decides whether to keep or fire the CEO. Taylor estimates his model and finds that very high turnover costs are needed to rationalize the observed rate of forced CEO turnover. Huang, Wei, and Yan (2007) present a rational learning model that relates the asymmetric performance-flow relation for mutual funds to investors' participation costs.

The above studies assume learning by agents in theoretical models, but a learning perspective also seems useful in empirical work. Examples of studies that use Bayesian empirical techniques to analyze the performance of money managers include Baks et al. (2001), Pastor & Stambaugh (2002a), Jones & Shanken (2005), Busse & Irvine (2006), and Kosowski et al. (2007).

Individual Investor Trading

The trading behavior of individual investors exhibits interesting regularities. Individuals lose money by trading, on average, but they trade frequently nonetheless. Individuals' trading intensity depends on their past performance. Poor performance is often followed by exit. More active traders outperform less active ones. Performance exhibits persistence. Explanations offered for these facts range from overconfidence to utility from gambling.

Mahani & Bernhardt (2007) and Linnainmaa (2008) show that these facts are also consistent with rational learning. When individuals are uncertain about their own trading ability, they can learn by trading and observing their profits. Individuals can find it

optimal to trade even if they expect to lose money, as long as the expected short-term loss from trading is offset by the expected gain from learning. Individuals increase their trade sizes after successful trades and decrease them after unsuccessful trades because successful (unsuccessful) trades lead to upward (downward) revisions of perceived ability. More active traders perform better because good news about one's ability leads one to trade more. Linnainmaa finds empirically that the above-mentioned empirical regularities can be explained with moderate uncertainty about trading ability. In contrast, alternative explanations such as overconfidence and risk-seeking seem unable to explain all of the regularities.

How do investors learn from their trading experience? Is their ability a constant subject to learning, as in the models described above, or does it improve as a result of more trading, as in the learning-by-doing literature? Seru et al. (2008) find evidence of both types of learning. They find that poorly performing households are more likely to cease trading, which is consistent with the former type of learning, and they estimate this type of learning to be quantitatively more important than learning by doing.

Trading Volume

Why do investors trade so much? Why is trading volume correlated with volatility? Learning combined with information asymmetry can shed light on these questions. However, heterogeneous information alone cannot induce trading; given the no-trade theorem, trading requires additional motives, such as liquidity (e.g., Kyle 1985, Admati & Pfleiderer 1988, Wang 1993), hedging (e.g., Wang 1994), different prior beliefs (e.g., Detemple & Murthy 1994, Zapatero 1998, Basak 2000, Buraschi & Jiltsov 2006), or different interpretation of common signals (e.g., Harrison & Kreps 1978, Harris & Raviv 1993, Kandel & Pearson 1995, Scheinkman & Xiong 2003, David 2008b).

Wang (1994) helps us understand the correlation between trading volume and return volatility. His model features informed agents, who trade for both informational (speculative) and noninformational (hedging) reasons, and uninformed agents, who trade for noninformational reasons only. When the informed agents sell stocks, the stock price must drop to induce the uninformed agents to buy. As information asymmetry increases, the uninformed agents demand a larger discount to cover the risk of trading against private information. Therefore, trading volume is positively correlated with return volatility, and the correlation increases with information asymmetry. Wang's model also implies that hedging-motivated trading induces return reversals, whereas speculative trading induces return continuations. Llorente et al. (2002) find empirical support for these predictions.

Another way of modeling trading relies on differences in beliefs. In Scheinkman & Xiong (2003), heterogeneous beliefs arise from the presence of overconfident agents who believe their information is more accurate than it really is. These agents observe the same signals but, due to their behavioral bias, they interpret the signals differently. The resulting fluctuations in the differences of beliefs induce trading. The amount of trading in this model can be large, even infinite.

ENTREPRENEURIAL FINANCE

IPO: initial public offering

Firm profitability tends to rise before the firm's initial public offering (IPO) and fall after the IPO. Common explanations for these facts include irrationality and asymmetric

information. Pastor et al. (2009) show that these facts are also consistent with a rational symmetric-information model of learning. This model features two types of agents: investors, who are well diversified, and an entrepreneur, whose wealth is tied up in a private firm. All agents learn about the average profitability of the private firm by observing realized profits. The entrepreneur solves for the optimal time to go public, trading off diversification benefits of going public against benefits of private control. The model produces a cutoff rule whereby going public is optimal when the firm's expected future profitability is sufficiently high. Therefore, expected profitability must go up before the IPO. According to Bayes' rule, agents revise their expectations upward only if they observe realizations higher than expected. As a result, realized profitability exceeds expected profitability at the time of the IPO, and thus profitability is expected to drop after the IPO.

The model also predicts that the post-IPO drop in profitability is larger for firms with more volatile profitability and firms with less uncertain average profitability. These predictions also follow from Bayesian updating. Agents revise their expectations by less if their prior uncertainty is lower (because prior beliefs are stronger) and if signal volatility is higher (because signals are less precise). In both cases, realized profitability must rise more sharply to pull expected profitability above the IPO cutoff. As a result, the expected post-IPO drop in profitability is larger when volatility is higher and when uncertainty is lower. These predictions are supported empirically. Volatility and uncertainty can be separated by estimating the stock price reaction to earnings announcements, which is strong when uncertainty is high and volatility is low. Firms with weaker stock price reactions experience larger post-IPO drops in profitability, as predicted by the model. Because the volatility and uncertainty predictions seem unique to learning, this evidence suggests that learning is at least partly responsible for the observed profitability patterns around IPOs.

Sorensen (2008) develops a model of learning by investing, extending the multiarmed bandit model literature (e.g., Gittins 1989). In his model, each investment brings not only a monetary payoff but also more information, which helps improve future investment decisions. Investors learn from their own investment returns. Their optimal strategy trades off exploiting investments with known high payoffs and exploring investments with uncertain payoffs but a higher option value of learning. Sorensen estimates his model on U.S. data from venture capital (VC) investments. He finds that VCs' investment decisions are affected not only by immediate returns but also by the option value of learning. He also finds that VCs who engage in more learning are more successful. A related model is presented in Bergemann et al (2008).

Empirically, the performance of VC funds managed by the same general partner (GP) exhibits high persistence (unlike the performance of mutual funds). This fact raises the question why successful GPs do not raise their fees or fund size to capture all the surplus, as in Berk & Green (2004). Hochberg et al. (2008) rationalize VC performance persistence in a learning model in which investors learn about a GP's skill over time. The idea is that limited partners (LPs) who invest in a GP's fund learn more about the GP's skill than do other investors. This asymmetric learning enables incumbent LPs to hold up the highly-skilled GP when he raises his next fund because other potential investors would interpret incumbent LPs' failure to reinvest as a negative signal about the GP's skill. Due to their hold-up power, incumbent LPs continue to earn high net-of-fee returns in their investments in the follow-on funds of the same GP. In contrast, perfor-

VC: venture capital

VCs: venture capitalists

GP: general partner

LP: limited partner

mance persistence is weaker for mutual funds where asymmetric information between the incumbent investors and outsiders is smaller. Hochberg et al. also predict that LPs should earn higher returns in follow-on funds than in first-time funds, and there should be persistence in the LP composition across the funds run by the same GP. These predictions are supported empirically in a large sample of U.S. VC funds. Glode & Green (2009) also develop a model that relies on learning to rationalize performance persistence in private equity partnerships.

FUTURE ISSUES

1. Much work on the role of learning in finance still lies ahead. Some promising directions are evident in recent work that is not examined in this review. For example, in most existing learning models, agents learn by observing cash flows or asset returns, but they could also learn from the prices of derivative securities (e.g., Dubinsky & Johannes 2006, Beber & Brandt 2009, Johannes et al. 2009). Other interesting topics not covered here include endogenous information acquisition (e.g., Veldkamp 2006; Peng & Xiong 2006; van Nieuwerburgh & Veldkamp 2008, 2009), non-Bayesian learning (e.g., Gervais & Odean 2001, Brav & Heaton 2002, Piazzesi & Schneider 2007), learning by doing (e.g., Arrow 1962, Berk et al. 2004), informational cascades (e.g., Welch 1992), incomplete information equilibria (e.g., Feldman 2007), and higher-order beliefs (e.g., Allen et al. 2006, Banerjee et al. 2009).

2. Another promising direction is to separate systematic and idiosyncratic uncertainty, which have different implications for asset prices. Although idiosyncratic uncertainty increases both return volatility and asset valuations, systematic uncertainty increases volatility but decreases valuations. Time variation in the two types of uncertainty produces dynamic relations between prices, expected returns, and volatility. Separating the two types of uncertainty, perhaps with the help of option prices, could shed new light on the asset price dynamics.

3. Future work can also analyze strategic information generation. We have discussed learning from exogenously specified signals, but what agents observe may depend on the actions of other agents whose objectives are different. For example, corporate insiders may manipulate earnings, which are used by outside investors as signals about average profitability. It seems interesting to analyze dynamic agency models with asymmetric information. More generally, we need more dynamic learning models in corporate finance.

4. New learning models should be held to high standards. For each model, one should identify testable predictions that are unique to learning so the model can be empirically distinguished from alternatives. It is also important to assess the magnitude of the learning-induced effects, either by calibration or by structural estimation. Examples of the latter approach include Linnainmaa (2008), Sorensen (2008), and Taylor (2008). We expect to see more structural estimation of learning models down the road.

APPENDIX

Let $f_t(g)$ denote the probability density function of g at time t, with $\Pr(r > g) = 1$. The stock price is given by

$$P_t = E_t\left[\int_t^\infty e^{-r(\tau-t)}D_\tau d\tau\right] = \int_{-\infty}^r E\left[\int_t^\infty e^{-r(\tau-t)}D_\tau d\tau|g\right]f_t(g)dg.$$

Conditional on g, $D_\tau = D_t e^{(g-\sigma^2/2)(\tau-t)+\sigma(W_\tau-W_t)}$, with $W_\tau - W_t \sim N(0, \tau - t)$. Therefore,

$$
\begin{aligned}
P_t &= \int_{-\infty}^r \int_t^\infty E\left[e^{-r(\tau-t)}D_t e^{(g-\sigma^2/2)(\tau-t)+\sigma(W_\tau-W_t)}|g\right]d\tau f_t(g)dg \\
&= D_t\int_{-\infty}^r\int_t^\infty e^{-(r-g)(\tau-t)}d\tau f_t(g)dg = D_t\int_{-\infty}^r \frac{1}{r-g}f_t(g)dg,
\end{aligned}
\tag{15}
$$

which is Equation 7. When g is observable, $f_t(g)$ is degenerate and we obtain Equation 5.

The volatility in Equation 8 obtains from Equation 15 as follows: Let $f_t(g)$ represent the normal distribution with mean $\tilde{g}_t$ and variance $\tilde{\sigma}_t^2$, except for the truncation $g < r$. Approximate the dynamics of $\tilde{g}_t$ and $\tilde{\sigma}_t^2$ by Equation 4 with $ds_t = dD_t/D_t$, so that $d\tilde{g}_t \approx m_t(dD_t/D_t - \tilde{g}_t dt)$. This is an approximation because Equation 4 holds exactly only when $f_t(g)$ is non-truncated normal. Let $F(\tilde{g}_t, \tilde{\sigma}_t^2) \equiv \log(P_t/D_t)$. From Ito's Lemma,

$$\frac{dP_t}{P_t} = \frac{dD_t}{D_t} + \left(\frac{\partial F(\tilde{g}_t, \tilde{\sigma}_t^2)}{\partial \tilde{g}_t}\right)d\tilde{g}_t + o(dt),$$

where $o(dt)$ denotes deterministic terms of order dt. Substituting for $d\tilde{g}_t$ and rearranging,

$$\frac{dP_t}{P_t} \approx \frac{dD_t}{D_t} \times \left[1 + \left(\frac{\partial F(\tilde{g}_t, \tilde{\sigma}_t^2)}{\partial \tilde{g}_t}\right)m_t\right] + o(dt).$$

Taking standard deviations of both sides, we obtain return volatility in Equation 8.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

Admati AR, Pfleiderer P. 1988. A theory of intraday patterns: volume and price variability. *Rev. Financ. Stud.* 1:3–40

Ai H. 2007. *Information quality and long-run risk: asset pricing and welfare implications*. Work. Pap., Duke Univ.

Allen F, Morris S, Shin HS. 2006. Beauty contests, bubbles, and iterated expectations in asset markets. *Rev. Financ. Stud.* 19:719–52

Arrow K. 1962. The economic implications of learning by doing. *Rev. Econ. Stud.* 29:155–73

Avramov D. 2002. Stock return predictability and model uncertainty. *J. Financ. Econ.* 64:423–58

Avramov D. 2004. Stock return predictability and asset pricing models. *Rev. Financ. Stud.* 17:699–738

Avramov D, Chordia T. 2006. Predicting stock returns. *J. Financ. Econ.* 82:387–415

Avramov D, Wermers R. 2006. Investing in mutual funds when returns are predictable. *J. Financ. Econ.* 81:339–77

Baks KP, Metrick A, Wachter JA. 2001. Should investors avoid all actively managed mutual funds? A study in Bayesian performance evaluation. *J. Finance* 56:45–85

Banerjee S, Kaniel R, Kremer I. 2009. Price drift as an outcome of differences in higher-order beliefs. *Rev. Financ. Stud*. In press

Bansal R, Yaron A. 2004. Risks for the long run: a potential resolution of asset pricing puzzles. *J. Finance* 59:1481–509

Barberis N. 2000. Investing for the long run when returns are predictable. *J. Finance* 55:225–64

Basak S. 2000. A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk. *J. Econ. Dyn. Control* 24:63–95

Battalio RH, Schultz P. 2006. Options and the bubble. *J. Finance* 61:2071–102

Bawa VS, Brown SJ, Klein RW. 1979. *Estimation Risk and Optimal Portfolio Choice*. Amsterdam: North-Holland

Beber A, Brandt MW. 2009. Resolving macroeconomic uncertainty in stock and bond markets. *Rev. Finance* 13:1–45

Bergemann D, Hege U, Peng L. 2008. *Venture capital and sequential investments*. Work. Pap., Yale Univ.

Berk JB, Green RC. 2004. Mutual fund flows and performance in rational markets. *J. Polit. Econ.* 112:1269–95

Berk JB, Green RC, Naik V. 2004. Valuation and return dynamics of new ventures. *Rev. Financ. Stud.* 17:1–35

Brandt MW, Goyal A, Santa-Clara P, Stroud JR. 2005. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *Rev. Financ. Stud.* 18:831–73

Brandt MW, Zeng Q, Zhang L. 2004. Equilibrium stock return dynamics under alternative rules of learning about hidden states. *J. Econ. Dyn. Control* 28:1925–54

Brav A, Heaton JB. 2002. Competing theories of financial anomalies. *Rev. Financ. Stud.* 15:575–606

Brennan MJ. 1998. The role of learning in dynamic portfolio decisions. *Eur. Finance Rev.* 1:295–306

Brennan MJ, Xia Y. 2001a. Stock return volatility and equity premium. *J. Monet. Econ.* 47:249–83

Brennan MJ, Xia Y. 2001b. Assessing asset pricing anomalies. *Rev. Financ. Stud.* 14:905–42

Brown SJ. 1976. *Optimal portfolio choice under uncertainty: A Bayesian approach*. PhD thesis. Univ. Chicago

Buraschi A, Jiltsov A. 2006. Model uncertainty and option markets with heterogeneous beliefs. *J. Finance* 61:2841–97

Busse J, Irvine PJ. 2006. Bayesian alphas and mutual fund persistence. *J. Finance* 61:2251–88

Cagetti M, Hansen LP, Sargent T, Williams N. 2002. Robustness and pricing with uncertain growth. *Rev. Financ. Stud.* 15:363–404

Coles J, Loewenstein U. 1988. Equilibrium pricing and portfolio composition in the presence of uncertain parameters. *J. Financ. Econ.* 22:279–303

Cremers M. 2002. Stock return predictability: A Bayesian model selection perspective. *Rev. Financ. Stud.* 15:1223–49

Croce MM, Lettau M, Ludvigson SC. 2006. *Investor information, long-run risk, and the duration of risky cash flows*. Work. Pap., New York Univ.

Cvitanic J, Lazrak A, Martellini L, Zapatero F. 2006. Dynamic portfolio choice with parameter uncertainty and the economic value of analysts' recommendations. *Rev. Financ. Stud.* 19:1113–56

Dangl T, Wu Y, Zechner J. 2008. Market discipline and internal governance in the mutual fund industry. *Rev. Financ. Stud.* 21:2307–43

David A. 1997. Fluctuating confidence in stock markets: implications for return and volatility. *J. Financ. Quant. Anal.* 32:427–62

David A. 2008a. Inflation uncertainty, asset valuations, and the credit spreads puzzle. *Rev. Financ. Stud.* 21:2487–534

David A. 2008b. Heterogeneous beliefs, speculation, and the equity premium. *J. Finance* 63:41–83

David A, Veronesi P. 2002. *Option prices with uncertain fundamentals: theory and evidence on the dynamics of implied volatilities*. Work. Pap., Univ. Chicago

David A, Veronesi P. 2008. *Inflation and earnings uncertainty and volatility forecasts: a structural form approach*. Work. Pap., Univ. Chicago

Detemple JB. 1986. Asset pricing in a production economy with incomplete information. *J. Finance* 41:383–91

Detemple JB, Murthy S. 1994. Intertemporal asset pricing with heterogeneous beliefs. *J. Econ. Theory* 62:294–320

Diether KB, Malloy CJ, Scherbina A. 2002. Differences of opinion and the cross section of stock returns. *J. Finance* 57:2113–41

Donaldson RG, Kamstra M. 1996. A new dividend forecasting procedure that rejects bubbles in asset prices: the case of 1929's stock crash. *Rev. Financ. Stud.* 9:333–83

Dothan MU, Feldman D. 1986. Equilibrium interest rates and multiperiod bonds in a partially observable economy. *J. Finance* 41:369–82

Dubinsky A, Johannes M. 2006. *Earnings announcements and equity options*. Work. Pap., Columbia Univ.

Duffie D, Lando D. 2001. Term structure of credit spreads with incomplete accounting information. *Econometrica* 69:633–64

Epstein L, Schneider M. 2008. Ambiguity, information quality and asset prices. *J. Finance* 63:197–228

Feldman D. 2007. Incomplete information equilibria: separation theorems and other myths. *Ann. Oper. Res.* 151:119–49

Garlappi L, Uppal R, Wang T. 2007. Portfolio selection with parameter and model uncertainty: a multi-prior approach. *Rev. Financ. Stud.* 20:41–81

Gennotte G. 1986. Optimal portfolio choice under incomplete information. *J. Finance* 41:733–46

Gervais S, Odean T. 2001. Learning to be overconfident. *Rev. Financ. Stud.* 14:1–27

Gittins J. 1989. *Multi-armed Bandit Allocation Indices*. New York: Wiley

Glode V, Green RC. 2009. *Information spillovers and performance persistence in private equity partnerships*. Work. Pap., Univ. of Penn.

Guidolin M, Timmermann AG. 2007. Asset allocation under multivariate regime switching. *J. Econ. Dyn. Control* 31:3503–44

Harris M, Raviv A. 1993. Differences of opinion make a horse race. *Rev. Financ. Stud.* 6:473–506

Harrison J, Kreps D. 1978. Speculative behavior in a stock market with heterogeneous expectations. *Q. J. Econ.* 92:323–36

Hochberg Y, Ljungqvist A, Vissing-Jorgensen A. 2008. *Informational hold-up and performance persistence in venture capital*. Work. Pap., Northwestern Univ.

Honda T. 2003. Optimal portfolio choice for unobservable and regime-switching mean returns. *J. Econ. Dyn. Control* 28:45–78

Hong H, Scheinkman J, Xiong W. 2006. Asset float and speculative bubbles. *J. Finance* 61:1073–117

Hong H, Scheinkman J, Xiong W. 2008. Advisors and asset prices: a model of the origins of bubbles. *J. Financ. Econ.* 89:268–89

Huang J, Wei KD, Yan H. 2007. Participation costs and the sensitivity of fund flows to past performance. *J. Finance* 62:1273–311

Johannes M, Polson N, Stroud J. 2009. Optimal filtering of jump-diffusions: extracting latent states from asset prices. *Rev. Financ. Stud.* 22:2759–99

Johnson TC. 2001. Return dynamics when persistence is unobservable. *Math. Finance* 11:415–45

Johnson TC. 2004. Forecast dispersion and the cross section of expected returns. *J. Finance* 59:1957–78

Johnson TC. 2007. Optimal learning and new technology bubbles. *J. Monet. Econ.* 54:2486–511

Jones CS, Shanken J. 2005. Mutual fund performance with learning across funds. *J. Financ. Econ.* 78:507–52

Kan R, Zhou G. 2007. Optimal portfolio choice with parameter uncertainty. *J. Financ. Quant. Anal.* 42:621–56

Kandel E, Pearson N. 1995. Differential interpretation of public signals and trade in speculative markets. *J. Polit. Econ.* 103:831–72

Kandel S, Stambaugh RF. 1996. On the predictability of stock returns: an asset allocation perspective. *J. Finance* 51:385–424

Klein RW, Bawa VS. 1976. The effect of estimation risk on optimal portfolio choice. *J. Financ. Econ.* 3:215–31

Kosowski R, Naik NY, Teo M. 2007. Do hedge funds deliver alpha? A Bayesian and bootstrap analysis. *J. Financ. Econ.* 84:229–64

Kyle AS. 1985. Continuous auctions and insider trading. *Econometrica* 53:1315–36

Leippold M, Trojani F, Vanini P. 2008. Learning and asset prices under ambiguous information. *Rev. Financ. Stud.* 21:2565–97

Lettau M, Ludvigson SC, Wachter JA. 2008. The declining equity premium: What role does macroeconomic risk play? *Rev. Financ. Stud.* 21:1653–87

Lewellen J, Shanken J. 2002. Learning, asset-pricing tests, and market efficiency. *J. Finance* 57:1113–45

Li CW, Xue HF. 2008. *A Bayesian's "bubble."* Work. Pap., La. State Univ.

Li G. 2005. Information quality, learning and stock market returns. *J. Financ. Quant. Anal.* 40:595–620

Linnainmaa J. 2008. *Why do households trade so much*? Work. Pap., Univ. Chicago

Llorente G, Michaely R, Saar G, Wang J. 2002. Dynamic volume-return relation of individual stocks. *Rev. Financ. Stud.* 15:1005–47

Lynch AW, Musto DK. 2003. How investors interpret past fund returns. *J. Finance* 58:2033–58

Mahani RS, Bernhardt D. 2007. Financial speculators' underperformance: learning, self-selection, and provision of liquidity. *J. Finance* 62:1313–40

Massa M, Simonov A. 2005. Is learning a dimension of risk? *J. Bank. Finance* 29:2605–32

Merton RC. 1971. Optimum consumption and portfolio rules in a continuous-time model. *J. Econ. Theory* 3:373–413

Merton RC. 1974. On the pricing of corporate debt: the risk structure of interest rates. *J. Finance* 29:449–70

Ozoguz A. 2009. Good times or bad times? Investors' uncertainty and stock returns. *Rev. Financ. Stud.* In press

Pastor L. 2000. Portfolio selection and asset pricing models. *J. Finance* 55:179–223

Pastor L, Stambaugh RF. 2000. Comparing asset pricing models: an investment perspective. *J. Financ. Econ.* 56:335–81

Pastor L, Stambaugh RF. 2002a. Mutual fund performance and seemingly unrelated assets. *J. Financ. Econ.* 63:315–49

Pastor L, Stambaugh RF. 2002b. Investing in equity mutual funds. *J. Financ. Econ.* 63:351–80

Pastor L, Stambaugh RF. 2008. *Are stocks really less volatile in the long run*? Work. Pap., Univ. Chicago

Pastor L, Stambaugh RF. 2009. Predictive systems living with imperfect predictors. *J. Financ.* 64:1583–628

Pastor L, Taylor LA, Veronesi P. 2009. Entrepreneurial learning, the IPO decision, and the post-IPO drop in firm profitability. *Rev. Financ. Stud.* 22:3005–46

Pastor L, Veronesi P. 2003. Stock valuation and learning about profitability. *J. Finance* 58:1749–89

Pastor L, Veronesi P. 2006. Was there a Nasdaq bubble in the late 1990s? *J. Financ. Econ.* 81:61–100

Pastor L, Veronesi P. 2009. Technological revolutions and stock prices. *Am. Econ. Rev.* 99:1451–83

Peng L, Xiong W. 2006. Investor attention, overconfidence, and category learning. *J. Financ. Econ.* 80:563–602

Piazzesi M, Schneider M. 2007. Equilibrium yield curves. In *NBER Macroeconomics Annual 2006*, ed. D Acemoglu, K Rogoff, M Woodford, pp. 389–442. Cambridge, MA: MIT Press

Rytchkov O. 2008. *Filtering out expected dividends and expected returns*. Work. Pap., Univ. Tex.

Scheinkman J, Xiong W. 2003. Overconfidence and speculative bubbles. *J. Polit. Econ.* 111:1183–219

Seru A, Shumway T, Stoffman N. 2008. Learning by trading. *Rev. Financ. Stud.* In press

Sorensen M. 2008. *Learning by investing: Evidence from venture capital*. Work. Pap., Columbia Univ.

Stambaugh RF. 1999. Predictive regressions. *J. Financ. Econ.* 54:375–421

Taylor LA. 2008. *Why are CEOs rarely fired? Evidence from structural estimation*. Work. Pap., Univ. Penn.

Timmermann AG. 1993. How learning in financial markets generates excess volatility and predictability of stock returns. *Q. J. Econ.* 108:1135–45

Timmermann AG. 1996. Excess volatility and predictability of stock prices in autoregressive dividend models with learning. *Rev. Econ. Stud.* 63:523–57

Tu J, Zhou G. 2004. Data-generating process uncertainty: What difference does it make in portfolio decisions? *J. Financ. Econ.* 72:385–421

Tu J, Zhou G. 2008. *Incorporating economic objectives into Bayesian priors: portfolio choice under parameter uncertainty*. Work. Pap., Wash. Univ., St. Louis

Uppal R, Wang T. 2003. Model misspecification and under-diversification. *J. Finance* 58:2465–86

van Binsbergen JH, Koijen RSJ. 2008. *Predictive regressions: a present-value approach*. Work. Pap., Duke Univ.

van Nieuwerburgh S, Veldkamp L. 2008. *Information acquisition and under-diversification*. Work. Pap., New York Univ.

van Nieuwerburgh S, Veldkamp L. 2009. Information immobility and the home bias puzzle. *J. Finance.* 64:1187–215

Veldkamp L. 2006. Media frenzies in markets for financial information. *Am. Econ. Rev.* 96:577–601

Veronesi P. 1999. Stock market overreaction to bad news in good times: a rational expectations equilibrium model. *Rev. Financ. Stud.* 12:975–1007

Veronesi P. 2000. How does information quality affect stock returns? *J. Finance* 55:807–37

Veronesi P. 2004. The peso problem hypothesis and stock market returns. *J. Econ. Dyn. Control* 28:707–25

Wachter JA, Warusawitharana M. 2009. Predictable returns and asset allocation: Should a skeptical investor time the market? *J. Econ.* 148:162–78

Wang J. 1993. A model of intertemporal asset prices under asymmetric information. *Rev. Econ. Stud.* 60:249–82

Wang J. 1994. A model of competitive stock trading volume. *J. Polit. Econ.* 102:127–68

Wang Z. 2005. A shrinkage approach to model uncertainty and asset allocation. *Rev. Financ. Stud.* 18:673–705

Weitzman M. 2007. Subjective expectations and asset-return puzzles. *Am. Econ. Rev.* 97:1102–30

Welch I. 1992. Sequential sales, learning, and cascades. *J. Finance* 47:695–732

Williams JT. 1977. Capital asset prices with heterogeneous beliefs. *J. Financ. Econ.* 5:219–39

Xia Y. 2001. Learning about predictability: the effects of parameter uncertainty on dynamic asset allocation. *J. Finance* 56:205–46

Yu F. 2005. Accounting transparency and the term structure of credit spreads. *J. Financ. Econ.* 75:53–84

Zapatero F. 1998. Effect of financial innovation on market volatility when beliefs are heterogeneous. *J. Econ. Dyn. Control* 22:597–626

Zellner A, Chetty VK. 1965. Prediction and decision problems in regression models from the Bayesian point of view. *J. Am. Stat. Assoc.* 60:608–16

Contents

Errata

An online log of corrections to *Annual Review of Financial Economics* articles
may be found at http://financial.annualreviews.org