JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS VOL. 40, NO. 2, JUNE 2005
COPYRIGHT 2005, SCHOOL OF BUSINESS ADMINISTRATION, UNIVERSITY OF WASHINGTON, SEATTLE, WA 98195

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

Learning with Information Capacity Constraints

Lin Peng*

Abstract

Motivated by the fact that investors have limited time and attention to process information, this paper provides a continuous-time equilibrium model to analyze the effects of a capacity constraint in the learning process of a representative investor, who optimally allocates her information capacity across multiple sources of uncertainty. Consequently, the cross-sectional structure of information and the resulting asset price dynamics are determined endogenously. The model provides implications on both consumption behavior and the cross-sectional differences in price informativeness in terms of supply of information, speed of price adjustments to fundamental shocks, and price reactions to firm disclosures.

I. Introduction

Learning is crucial to many issues in finance. Investors process information to make their consumption and portfolio decisions, while corporate managers analyze economic conditions to make real investment decisions. Although many studies address the effects of learning, they usually treat information as exogenous, assuming that economic agents can process information instantaneously.

In reality, it takes time and attention for agents to process information and incorporate it into their decisions. Take the example of firm valuation: investors not only need to exert time and effort to analyze a firm's financial statements, but also to study its competitors and industry. In some cases, it takes market participants days or even weeks to react to sudden shocks. For example, on January 17, 1995, the day of the worst earthquake in Japan since 1923, the market reacted only slightly. The strongest reaction was not evident until a week later (Shiller

*lin_peng@baruch.cuny.edu, Department of Economics and Finance, Zicklin School of Business, Baruch College, City University of New York, One Bernard Baruch Way, Box B10-225, New York, NY 10010. This paper is developed from a chapter of my dissertation at Duke University. I appreciate helpful comments from Tim Bollerslev, Joshua Coval (the referee), David Feldman, Campbell Harvey, David Hirshleifer, David Hsieh, Jonathan Karpoff (the editor), Pete Kyle, George Oldfield, Hui Ou-Yang, Emma Rasiel, Ailsa Roell, Chris Sims, S. Viswanathan, Wei Xiong, Paul Zipkin, and seminar participants at Baruch College, Columbia University, Cornell University, Duke University, the Federal Reserve Board of Governors, Indiana University, Penn State University, New York University, Northwestern University, University of North Carolina at Chapel Hill, University of Notre Dame, Yale University, and conference participants at the Frank Batten Young Scholars Conference at the College of William and Mary and the 2003 Summer Meeting of the Econometric Society. I appreciate financial support from a PSC-CUNY research grant. All remaining errors are my responsibility.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

(2000), pp. 79–81). In other cases, even if news is publicly available, it is not incorporated into prices until investors pay attention. According to Huberman and Regev (2001), the publication of an article in the *New York Times* about a new cancer-curing drug from EntreMed attracted great public attention and generated a daily return of more than 300% in its stocks, even though the same story had already been published more than five months earlier in *Nature* and other newspapers. Experimental evidence also suggests that professional analysts often fail to recall and respond appropriately to information in complex financial disclosures, suggesting that the time and attention needed to process financial information is non-trivial (Hirst and Hopkins (1998)).

In this paper, I study the learning process of a representative investor with a capacity (or attention) constraint and analyze the effects of the investor's capacity constraint on consumption behavior and asset price dynamics. Given unobservable fundamental factors in her portfolio, the investor processes information to infer their evolution, and the capacity constraint limits the amount of information that she can process. When there are multiple sources of uncertainty, the investor optimally allocates her capacity across them. As a result, the information and learning dynamics become endogenous: assets may have heterogeneous information environments, therefore their prices may incorporate fundamental shocks at different speeds and respond to firm announcements with different magnitudes, and the investor updates her consumption decisions with a delay.

The model is in continuous time with an infinite horizon. The representative investor is endowed with a portfolio of financial assets whose cash flows are driven by multiple fundamental factors that are unobservable to her. The investor can borrow and lend at a constant interest rate. To maximize her expected lifetime utility, the investor optimally makes her consumption and portfolio decisions based on her inference about the fundamental variables. An exponential utility function and Gaussian distributions for all variables allow for a linear stationary equilibrium in which asset prices are determined through the market-clearing condition.

I model information processing and learning in two steps. In step one, the investor processes information about the fundamental factors and produces a vector of signals as the outcome. This step is subject to a capacity constraint. Following Sims (2003), I measure uncertainty using the concept of entropy from information theory; specifically, I use the rate of entropy reduction to quantify the rate of information processing.[1] I impose the capacity constraint on the investor's learning process as the maximum rate of entropy reduction that can be achieved. In step two, the signals are incorporated into the investor's belief through rational Bayesian updating. While this approach is a simplification of the actual complex learning process, it nonetheless captures the essential feature that learning takes time and attention. This model provides a basis to study the effects of learning or attention constraints on economic decision making. Other models in the existing learning literature usually start from the second step and take the signals as given.

In equilibrium, the investor allocates her limited attention across different sources of uncertainty to minimize the total uncertainty of her portfolio, which

[1] See Cover and Thomas (1991) for an introduction to information theory.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

in turn enables her to make more efficient intertemporal consumption decisions. The capacity constraint on information processing directly leads to a delay in the reaction of the investor's consumption to fundamental shocks, consistent with the documented predictability in aggregate consumption data. The investor's capacity allocation also provides implications for the information environment and price dynamics across assets. I illustrate these implications by analytically solving a case in which the assets have independent fundamentals. To the extent that larger stocks contribute more to the uncertainty in the investor's portfolio, they attract more capacity allocation. As a result, larger stocks tend to have a greater supply of information, their prices incorporate fundamental shocks at a faster speed, and they exhibit less volatility to exogenous announcements from firms. These predictions are consistent with the empirical findings.

The fixed capacity constraint imposed in this paper is not crucial. Alternatively, I can specify an expandable capacity subject to a convex cost function, in which case the investor allocates her capacity after the total capacity is determined. The implications will remain qualitatively similar. The capacity constraint in the model can be interpreted as a limit on the information processing ability of the aggregate financial market. The representative investor serves as a social planner in allocating information capacity. The optimal capacity allocation solved in the model describes the Pareto optimal structure of information processing. Issues related to externalities in a decentralized market are further discussed in Section VIII.

This paper is structured as follows. Section II provides a review of the literature. In Section III, I describe the economy. Section IV solves the asset market equilibrium for a given capacity allocation. Section V sets up a general capacity allocation problem. In Section VI, I solve the capacity allocation problem analytically with independent assets. In Section VII, I discuss the implications. Section VIII provides some further discussion, and Section IX concludes. All proofs to lemmas, theorems, and propositions are provided in the Appendix.

II. Related Studies

There is a growing literature in finance that studies the learning process of investors. Detemple (1986), Dothan and Feldman (1986), Gennotte (1986), and Feldman (1992) apply linear filtering techniques to study asset prices with incomplete information. Veronesi (2000) discusses the relation between asset prices and information quality using non-Gaussian beliefs. Brennan and Xia (2001) show that the learning process of investors can generate excess volatility and thus result in an additional risk premium. Yan (2001) also studies the effects of estimation uncertainty on the equity premium. In these papers, investor information is specified exogenously. My model emphasizes the time and attention involved in information processing.

Recently the economics and finance literature has started to address issues related to the limited ability of economic agents. Kalai and Stanford (1988) study the complexity of strategies and finite rationality. Abreu and Rubinstein (1988) model the computational ability of players in dynamic games using finite automata. Dow (1991) analyzes the search process of a consumer with limited

memory. Sims (2003) applies information theory to study the limited attention of an economic agent and its implications for dynamic programming problems in macroeconomics. Hirshleifer and Teoh (2003) and Hirshleifer, Lim, and Teoh (2002) study firms' disclosure policies when investors have limited attention. My paper is distinct from these papers in that I analyze the capacity allocation problem of investors and the implications for the information environment and price dynamics of assets. In a closely related paper, Peng and Xiong (2003) examine how investors' limited information processing capacities affect the dynamics of asset price volatility.

Issues related to asset price informativeness have also been analyzed in the market microstructure literature (e.g., Grossman and Stiglitz (1980), Glosten and Milgrom (1985), Kyle (1984), (1985), Easley and O'Hara (1987), Admati and Pfleiderer (1988), Holden and Subrahmanyam (1992), and Foster and Viswanathan (1993)). More complete literature reviews are provided in Admati (1991) and O'Hara (1995). While these papers analyze the effects of information on asset prices through trades between informed and uninformed traders, my paper is concerned with a different issue—the ability of investors to process information.

III. The Economy

I consider a continuous-time model with a representative investor and an infinite horizon. There are m risky assets, and their supply is denoted by the vector $e_{m\times 1}$ measured in number of shares. A risk-free storage technology is available to the investor at a constant rate r.[2] The representative investor is not only the agent who invests and consumes, but she is also an information provider who analyzes and processes raw data. Thus, the investor can be interpreted as the aggregation of all the participants in financial markets including investors, financial analysts, and the news media.

The dividend processes from these assets are observable to the investor, and the dividends per share for each asset are denoted as the vector $\{dD_t\}_{m\times 1}$. The dividend flows are determined by k unobservable independent stochastic factors, $\{F_t\}_{k\times 1}$, according to the following diffusion process, $dD_t = \Phi F_t dt + \Omega_D dz_D$, where $\Phi_{m\times k}$ represents the assets' loadings on the fundamental factors, $\{\Omega_D\}_{m\times m}$ is the matrix of volatility parameters, and $\{z_D\}_{m\times 1}$ is a vector of independent Wiener processes that represent innovations to dividends. Both the drift and innovations of the dividend process are stochastic and unobservable. Without loss of generality, I assume that the dividend innovations across assets are mutually independent, thus Ω_D is diagonal. I also assume that each diagonal element of Ω_D is strictly positive, so that dividend flow does not perfectly reveal the fundamental factors, F_t.

The fundamental factors are mutually independent and evolve according to the following mean-reverting process,

$$dF_t = -\Lambda (F_t - \bar{F})\, dt + \Omega_F dz_F, \tag{1}$$

[2]The constant risk-free rate assumption is used in equilibrium models such as Stapleton and Subrahmanyam (1978), Campbell and Kyle (1993), and Wang (1993). Sundaresan (1983) shows that a constant equilibrium interest rate can be consistent with a closed production economy and CARA utility functions.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

where $\Lambda_{k\times k}$ is a diagonal matrix of the mean-reverting parameters, $\bar{F}_{k\times 1}$ is a vector of the long-run means, $\{\Omega_F\}_{k\times k}$ is a diagonal matrix of volatility parameters, and $\{z_F\}_{k\times 1}$ is a vector of independent standard Wiener processes representing innovations to the fundamental factors. I assume that dz_D and dz_F are independent and that Λ, $\bar{F}$, and Ω_F are constants known to the investor.

The investor can infer the evolution of F_t through two sources of information. One is the observed dividend flow, which is exogenous and provides information that can be readily used by the investor. Examples of dividend information include earnings announcements and other forecasts provided by firms. I assume that this type of information requires no processing and can be readily incorporated into the investor's belief. In addition, the investor can also analyze the fundamental factors on her own. Such activities, including analyzing a firm's financial statements, talking to its managers, and studying its industry, competitors, suppliers, and customers require time and mental capacity before signals can be generated and incorporated into the investor's belief. Examples of these signals are analysts' research reports and an investor's private research outcomes about a firm. To capture the limited time and mental capacity that the investor possesses, I impose a constraint on the amount of information she can process.

I assume that the investor processes information and generates a flow of unbiased but noisy signals about the fundamental factors, $dS_t = F_t dt + \Omega_S dz_S$. The signal flow, dS_t, is a $k \times 1$ vector with its drift term equal to the fundamental vector F_t; $\{\Omega_S\}_{k\times k}$ is the matrix of volatility parameters; and $\{z_S\}_{k\times 1}$, a vector of standard Wiener processes independent of z_F and z_D, represents signal noises. The precision matrix of the signals is denoted by $(\Omega_S\Omega_S^T)^{-1}$.

I model the information processing of independent fundamental factors as separate tasks, similar to the way in which English and mathematics need to be learned separately at school. Therefore, processing information about each factor requires its own capacity allocation,[3] and the resulting noisy signals on each factor are mutually independent. Thus, both $(\Omega_S\Omega_S^T)$ and its inverse, the signal precision matrix, are diagonal matrices, which are endogenously determined by the capacity allocation decision of the investor as described in Section V. Intuitively, when the investor chooses to allocate more capacity to a factor, more information is processed about the factor and the resulting signal has less noise and greater precision.

To solve the equilibrium with information capacity constraints, I follow two steps. In Section IV, I solve the asset market equilibrium for a given capacity allocation. The investor's value function can be expressed as a function of her capacity allocation decision. In Section V, I derive the investor's optimal capacity allocation. Proceeding in these two steps, I solve a linear stationary equilibrium in which the investor's information capacity allocation, her consumption and portfolio decisions, the optimal information structure, and the corresponding asset price dynamics are jointly determined.

[3]It is possible that some capacity can be shared across the learning tasks of different factors, such as acquiring certain common skills in processing information. As long as the sharable capacity has decreasing returns to scale, the investor's capacity allocation will remain qualitatively similar, although the exact magnitude of the allocation might be different.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

IV. The Asset Market Equilibrium Given Capacity Allocation

In this section, I first solve the investor's inference problem given information, then the investor's consumption-portfolio choice problem based on her inference.[4] In equilibrium, the asset prices are determined by the market-clearing condition.

Based on the exogenous dividend flow, $D(t)$, and the endogenous signal flow produced by the investor, $S(t)$, the investor forms an inference about the fundamental vector, F. I denote the information set of the investor at time t as $I_t = \{D(t'), S(t')\}_{t' \leq t}$. I assume that the prior belief of the investor at time zero has a Gaussian distribution. Following Liptser and Shiryayev (1978), I solve the investor's inference problem in the following lemma.

Lemma 1. Conditional on her information set I_t, the investor's inference of the fundamental vector F_t has a Gaussian distribution: $F_t|I_t \sim N(\hat{F}_t, \Sigma)$. The conditional mean, $\{\hat{F}_t\}_{k\times 1}$, evolves according to

$$d\hat{F}_t = -\Lambda\left(\hat{F}_t - \bar{F}\right)dt + \Sigma\Phi^T\Omega_D^{-1}dw_D + \Sigma\Omega_s^{-1}dw_S, \tag{2}$$

where $dw_D = \Omega_D^{-1}(dD_t - \Phi\hat{F}_t dt)$ and $dw_S = \Omega_S^{-1}(dS_t - \hat{F}_t dt)$ are independent and standard Wienner processes, representing surprises from the dividend flow and signal flow, respectively. In a stationary equilibrium, the conditional variance of the belief, $\Sigma_{k\times k}$, is constant and is determined by the following equation,

$$\Omega_F\Omega_F^T - \Lambda\Sigma - \Sigma\Lambda^T - \Sigma\Phi^T\left(\Omega_D\Omega_D^T\right)^{-1}\Phi\Sigma^T - \Sigma\left(\Omega_S\Omega_S^T\right)^{-1}\Sigma^T = 0. \tag{3}$$

The conditional mean of the belief, $\hat{F}_t$, is an unbiased estimator of the true fundamental vector, F_t, based on the investor's information set I_t. The conditional variance Σ is determined by the fundamental volatility, the variance of dividend innovations, and the signal noises. Since the variance of signal noises is determined *endogenously* by the investor's capacity allocation decision (as Section V shows), the investor's belief variance is also affected by the capacity allocation decision.

I assume that the investor has an exponential utility function,

$$u(C) = -\frac{1}{\gamma}e^{-\gamma C},$$

where γ is the risk aversion parameter. Based on her inference of the asset fundamentals, the investor optimally chooses her consumption, C, and portfolio holdings, X, to maximize her expected lifetime utility,

$$J(W, \hat{F}_t, t) = \max_{C,X} E_t \int_{s=t}^{\infty} e^{-\beta(s-t)}u(C_s)ds,$$

subject to the dynamic budget constraint,

$$dW = (rW - C)dt + XdQ,$$

[4]Similar techniques are employed in Detemple (1986), Dothan and Feldman (1986), Gennotte (1986), Feldman (1992), Wang (1993), Brennan and Xia (2001), and Yan (2001).

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

where β is the investor's time preference parameter, r is the constant risk-free rate, and $dQ_{m\times 1}$ is the excess return per share for the risky assets, $dQ = dD + dP - rPdt$. In equilibrium, the prices of the risky assets, P, are determined so that the investor's optimal demand for the risky assets equals their total supply (market-clearing conditions): $X = e$.

With a constant absolute risk aversion preference, the investor's demand for risky assets is independent of her wealth, and so are the equilibrium asset prices. This utility specification, together with the assumption that all variables are jointly normal, permits a linear solution for the asset market equilibrium as shown in the following theorem.[5]

Theorem 1. Given the investor's inferences of the fundamental factors, $F_t|I_t \sim N(\hat{F}_t, \Sigma)$, the equilibrium asset price increases with $\hat{F}_t$ according to

$$P(\hat{F}) = r^{-1}\Phi\bar{F} + A(\hat{F} - \bar{F}) - \gamma\Omega\Omega^T e, \tag{4}$$

where $A = \Phi(\Lambda + rI_k)^{-1}$. The variance matrix of the assets' excess returns, dQ, is

$$\Omega\Omega^T = A\Omega_F\Omega_F^T A^T + \Omega_D\Omega_D^T + 2rA\Sigma A^T. \tag{5}$$

The investor's consumption increases linearly with her wealth,

$$C = rW - \frac{1}{r\gamma}\left(r - \beta - \frac{1}{2}\mu^T(\Omega\Omega^T)^{-1}\mu\right), \tag{6}$$

where her wealth is given by

$$W(\hat{F}) = K + r^{-1}e^T\Phi\bar{F} + e^T\Phi(\Lambda + rI_k)^{-1}(\hat{F} - \bar{F}) - \gamma e^T\Omega\Omega^T e, \tag{7}$$

with K donating her cash holdings. The investor's value function in equilibrium is

$$J = -\frac{1}{r\gamma}\exp\left[-r\gamma\left(K + e^T A\hat{F}\right) - \gamma e^T A\Lambda\bar{F} + \frac{r-\beta}{r} + \frac{1}{2}r\gamma^2 e^T\Omega\Omega^T e\right]. \tag{8}$$

Since the investor is holding risky assets with unobservable fundamentals, the learning process helps the investor form a more precise estimate of her wealth as in equation (7). Therefore, learning helps the investor make better intertemporal consumption decisions, as it is optimal for the investor to consume linearly with her current wealth (equation (6)).

Equation (8) shows that the investor's value function decreases with $e^T\Omega\Omega^T e$, the portfolio return variance. To maximize her value function, the investor must minimize this variance. There are three terms in $\Omega\Omega^T$, as equation (5) shows. The first two terms are exogenous and are due to the variance of the fundamental factors and the variance of the dividend processes. The third term is related to the variance of the investor's belief about the fundamental factors, which is endogenously determined by the investor's capacity constraint on learning.[6] Thus, maximizing the investor's value function is equivalent to

$$\min\ e^T A\Sigma A^T e. \tag{9}$$

[5] Similar assumptions are used by Campbell and Kyle (1993) and Wang (1993).

[6] In an extreme case, when the investor has infinite information capacity, she can process an infinite amount of information, her inference of the fundamental factors becomes perfect, and Σ is zero. In this case, the third term becomes zero.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

This term represents the variance in the investor's belief about the value of her portfolio. Note that Σ is determined by $(\Omega_S \Omega_S^T)^{-1}$ through equation (3). In the next section, I solve for the investor's capacity allocation decision to maximize this objective.

V. Capacity Allocation Problem

The capacity constraint determines a maximum rate of information in the signal flow processed by the investor. Following Sims (2003), I use the concept of entropy from information theory to quantify the rate of information. [7] In information theory, the uncertainty of a random variable X with a continuous probability density function, $f(x)$, can be measured by its entropy, which is defined as

$$\mathrm{H}(X) \quad = \quad -\mathrm{E}\log[f(X)] \quad = \quad -\int (\log f(x)) f(x) dx.$$

If X has a Gaussian distribution, $N(\bar{X}, \sigma^2)$, its entropy depends only on the logarithm of the volatility parameter, $\mathrm{H}(X) = \log \sigma + 0.5 \log(2\pi e)$, where e is the base of the exponential function. It is intuitive that for a Gaussian distribution higher uncertainty is associated with a larger variance.

The conditional entropy of X given another random variable Y is the expected entropy of the conditional distribution, averaged over the conditioning random variable Y. Let $f(x,y)$ be the joint probability density function of X and Y, $f(x|y)$ be the conditional distribution of X conditional on Y, and $g(y)$ be the marginal density function of Y. The conditional entropy of X given Y is

$$\mathrm{H}(X|Y) \quad = \quad \int \mathrm{H}(X|y) g(y) dy \quad = \quad -\int \int (\log f(x|y)) f(x,y) dx dy.$$

The amount of information that Y contains about X is defined as the reduction in the uncertainty of X due to the knowledge of Y,

$$\mathrm{I}(X\,;\,Y) \quad = \quad \mathrm{H}(X) - \mathrm{H}(X|Y) \quad = \quad \int \int f(x,y) \log \frac{f(x,y)}{f(x)g(y)} dx dy.$$

A nice and intuitive property of this information measure is that it is invariant to any nonzero linear transformations of both variables X and Y.

Lemma 2. If random variables X and Y are transformed by any nonzero constants a and b, respectively, the amount of information contained in bY about aX is the same as the amount of information contained in Y about X: $\mathrm{I}(aX\,;\,bY) = \mathrm{I}(X\,;\,Y)$.

This property establishes entropy reduction as a useful information measure that is independent of the scale of the underlying variables. It is particularly important for the multi-asset equilibrium, since different assets may have different

[7] For a reference on information theory, see Cover and Thomas (1991). Entropy is a widely used concept in many scientific fields. In statistics, entropy is used for estimation of spectral densities. In data compression, entropy is the lower bound on the average length of the shortest description of the random variable. In computer science, entropy is related to the complexity of a string. Entropy is also used in option pricing (see Stutzer (2000) for a review).

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

scales and a parsimonious measure of information across assets is needed. Other measures of information such as the precision of a signal may not possess this property.

Using entropy reduction as the measure of information, I impose the information capacity constraint as the maximum rate of entropy reduction that can be achieved by the investor. Essentially, the capacity constraint is the maximum bits per second (bps) in the signals processed by the investor. The following proposition states the investor's capacity allocation problem.

Proposition 1. Given a fixed information capacity κ ($\kappa \geq 0$), the investor chooses an optimal variance structure for the signal flows to solve

$$\min_{(\Omega_S \Omega_S^T)^{-1}} \; e^T \Phi (\Lambda + rI_k)^{-1} \Sigma (\Lambda + rI_k)^{-1} \Phi^T e \tag{10}$$

$$\text{subject to} \quad \text{(a)} \;\; \tfrac{1}{2} \operatorname{tr}\left(\Sigma \left(\Omega_S \Omega_S^T \right)^{-1} \right) \leq \kappa, \quad \text{and}$$
$$\text{(b)} \;\; \left(\Omega_S \Omega_S^T \right)^{-1} \text{ is positive semi-definite.} \tag{11}$$

The variance of the investor's belief, Σ, is determined by $(\Omega_S \Omega_S^T)^{-1}$ through equation (3).

Constraint (a) is the capacity constraint on the rate of entropy reduction caused by the signals. In constraint (b), I allow $(\Omega_S \Omega_S^T)^{-1}$ to be positive semi-definite since the investor may choose not to process any information about a factor. Once the investor determines the optimal information structure of the signals, $(\Omega_S \Omega_S^T)^{-1}$, she can solve both her inference problem and the consumption-portfolio problem as illustrated in Section IV. Consequently, I obtain a multi-asset equilibrium where the amount of information and asset prices are jointly determined.

VI. The Equilibrium with Optimal Capacity Allocation

In the remainder of the paper, I focus on the equilibrium and its implications for the case where assets are independent, and I briefly discuss the more general case with correlated assets in Section VIII. The framework developed in the previous section for solving the investor's capacity allocation problem does not yield analytical solutions in general. For an arbitrary fundamental correlation structure of assets, the factor loading matrix Φ is not necessarily diagonal. Since the investor uses the correlated dividends to update her belief about fundamental factors, the resulting belief variance matrix Σ, as a solution to the quadratic equation (3), is non-diagonal. Therefore, in the capacity allocation problem specified in equations (10) and (11), both the objective function and the constraint are non-linear functions of the choice variable $(\Omega_S \Omega_S^T)^{-1}$. This greatly complicates the problem and solving it would require numerical methods.

When assets are independent, both $(\Omega_S \Omega_S^T)^{-1}$ and $(\Omega_D \Omega_D^T)^{-1}$ are diagonal and consequently Σ is also diagonal. This allows one to derive analytic solutions to the capacity allocation problem and offers interesting implications on the

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

consumption behavior and the cross-sectional properties of information processing and price dynamics. Learning about asset-specific factors and their cross-sectional implications are economically important. In reality, financial analysts devote a significant amount of effort to follow individual companies in order to provide investors with firm-specific information. This information has also been one of the main drivers of the volatility of individual stocks. Campbell, Lettau, Malkiel, and Xu (2001) document an increasing trend of idiosyncratic volatility over the last 30 years, and show that the number of stocks needed to achieve a given level of diversification increases over time. Given the importance of asset-specific information, it becomes necessary to study both the learning process about these asset-specific factors and how prices incorporate asset-specific information.

Suppose that each of the m risky assets is driven by a fundamental factor independent of those of the other assets. Then the asset factor loading matrix, Φ, becomes diagonal, and assets are equivalent to fundamental factors. The investor treats information processing of individual assets as separate tasks, one for each asset. Both Σ and $(\Omega_S\Omega_S)^{-1}$ become diagonal matrices. Denote $e = \text{diagonal}(e_1, \ldots, e_m)$, $\Phi = \text{diagonal}(\phi_1, \ldots, \phi_m)$, $\Lambda = \text{diagonal}(\lambda_1, \ldots, \lambda_m)$, $\Sigma = \text{diagonal}(\Sigma_1, \ldots, \Sigma_m)$, $\Omega_S\Omega_S = \text{diagonal}(\sigma_{S1}^2, \ldots, \sigma_{Sm}^2)$, $\Omega_F\Omega_F = \text{diagonal}(\sigma_{F1}^2, \ldots, \sigma_{Fm}^2)$, and $\Omega_D\Omega_D = \text{diagonal}(\sigma_{D1}^2, \ldots, \sigma_{Dm}^2)$. Let κ_i be the ith diagonal element in matrix $\frac{1}{2}\Sigma(\Omega_S\Omega_S^T)^{-1}$, $\kappa_i = \Sigma_i\sigma_{Si}^{-2}/2$. Thus, κ_i represents the amount of information capacity that the investor allocates to asset (factor) i.

The capacity allocation problem in Proposition 1 can be rewritten as

$$\min_{\{\kappa_i\}} \quad \sum_{i=1}^{m} \frac{e_i^2\phi_i^2}{(\lambda_i + r)^2}\Sigma_i(\kappa_i) \tag{12}$$

$$\text{subject to} \quad \text{(a)} \ \sum_{i=1}^{m} \kappa_i \leq \kappa,$$

$$\text{(b)} \ \kappa_i \geq 0 \ \text{ for } \ i = 1, \ldots, m,$$

where Σ_i is given by

$$\Sigma_i(\kappa_i) = \sqrt{(\lambda_i + \kappa_i)^2\sigma_{Di}^4 + \sigma_{Fi}^2\sigma_{Di}^2} - (\lambda_i + \kappa_i)\sigma_{Di}^2. \tag{13}$$

Constraint (a) requires that the sum of the capacity allocated to each asset cannot exceed the total capacity, while constraint (b) imposes that the investor cannot allocate negative capacity to an asset, that is, information cannot be "unlearned." The positive root of equation (3) is Σ_i in equation (13).

It can be shown that Σ_i decreases monotonically with κ_i. As more capacity is allocated to asset i, there is less uncertainty remaining in the investor's belief. If the investor allocates zero capacity to asset i, Σ_i is at its maximum as she relies solely on information in the dividend flow to form inferences about the asset's fundamentals. In the other extreme, if the investor allocates infinite capacity to asset i, she will obtain a perfect signal on the asset's fundamental factor. Furthermore, Σ_i is a convex function of κ_i, that is, the marginal reduction in the belief variance decreases as capacity increases.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

Theorem 2. There exists a unique solution to the investor's capacity allocation problem with independent assets. The allocation $\{\kappa_i^*\}$ satisfies

$$-\frac{e_i^2\phi_i^2}{(\lambda_i+r)^2}\Sigma_i'(\kappa_i^*) \leq \eta, \text{ with equality if } \kappa_i^* > 0, \quad \forall\, i=1,\ldots,m, \quad \text{and}$$

$$\sum_{i=1}^{m}\kappa_i^* = \kappa,$$

where $\eta \geq 0$ is the Lagrange multiplier of the total capacity constraint.

The Lagrange multiplier η can be interpreted as the marginal value of capacity in reducing uncertainty. Asset i receives no capacity if the marginal reduction of uncertainty from increasing κ_i is less than η. Given the total capacity constraint, η also represents an implicit cost of allocating capacity to an asset. The implicit information cost complements the explicit cost analyzed by Verrecchia (1982).

The optimal capacity allocation allows me to analyze how innovations to the fundamental factor, F_i, are incorporated into the investor's belief, $\hat{F}_i$, over time, as in the following proposition.

Proposition 2. The joint dynamics of F_i and $\hat{F}_i$ are described by

$$F_{it} - \bar{F}_i = (F_{i0}-\bar{F}_i)\mathrm{e}^{-\lambda_i t} + \int_{s=0}^{t}\mathrm{e}^{-\lambda_i(t-s)}\sigma_{Fi}dz_{Fi}(s), \tag{14}$$

$$\begin{aligned}\hat{F}_{it} - \bar{F}_i = {} & (F_{i0}-\bar{F}_i)\mathrm{e}^{-\lambda_i t} + (\hat{F}_{i0}-F_{i0})\mathrm{e}^{-(\lambda_i+h_i)t} \\ & + \int_{s=0}^{t}\mathrm{e}^{-(\lambda_i+h_i)(t-s)}\frac{\Sigma_i}{\sigma_{Di}}dz_{Di}(s) \\ & + \int_{s=0}^{t}\left[\mathrm{e}^{-\lambda_i(t-s)}\left(1-\mathrm{e}^{-h_i(t-s)}\right)\right]\sigma_{Fi}dz_{Fi}(s) \\ & + \int_{s=0}^{t}\mathrm{e}^{-(\lambda_i+h_i)(t-s)}\frac{\Sigma_i}{\sigma_{Si}}dz_{Si}(s),\end{aligned} \tag{15}$$

where h_i characterizes the adjustment speed in the investor's inference and it increases with the capacity allocation, κ_i^*,

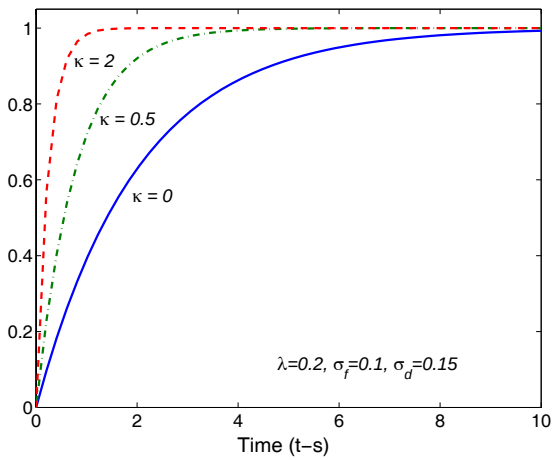
$$h_i = \sqrt{(\lambda_i+\kappa_i^*)^2 + \sigma_{Fi}^2/\sigma_{Di}^2} + \kappa_i^* - \lambda_i. \tag{16}$$

The term associated with the fundamental innovation $dz_{F_i}(s)$ in equation (14) represents its remaining effect on the fundamental factor F_i at time t. Due to mean reversion, the effect decays exponentially at the rate λ_i. The term associated with $dz_{F_i}(s)$ in equation (15) represents its remaining effect on $\hat{F}_i$, the investor's

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

inference about F_i. It has an additional factor, $1 - e^{-h_i(t-s)}$, compared with the corresponding term in equation (14). This factor is between zero and one, and it reflects the delayed response, or underreaction, in the investor's inference to fundamental shocks relative to that in the perfect information case. Figure 1 plots the delayed response factor, $1 - e^{-h_i(t-s)}$, for three different values of κ_i. When the shock arrives at time s, the reaction is zero since the investor does not observe the shock. Over time, the shock is gradually incorporated through the learning process. The learning speed depends positively on the parameter h_i, which in turn increases with κ_i^*. An increased capacity allocation allows the investor to process more information and infer the unobservable fundamental innovations at a faster speed.

FIGURE 1
Delayed Reaction to Fundamental Shocks



VII. Implications

In this section, I build on the equilibrium analysis developed in Section VI and discuss the effect of an information capacity constraint on consumption behavior and the cross-sectional properties of information and asset price dynamics.

A. The Effect on Consumption Behavior

Studies in the last 25 years examining the implications of the permanent income hypothesis on aggregate consumption behavior yield considerable deviations. Hall (1978) showed that if agents have full information and form rational expectations, their consumption decision would reflect shocks to their future income and changes in consumption should be unpredictable. However, Campbell and Deaton (1989) and West (1989) point out that the aggregate consumption is too smooth to be explained by permanent income theory. Campbell and Deaton (1989) argue that consumption responds to innovations in income with a delay.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

Campbell and Mankiw (1990) show that changes in consumption can be predicted by lagged variables.

My model provides a plausible economic rationale for consumption's slow adjustment to shocks. As equation (6) shows, the investor's optimal consumption rule follows a linear function of current wealth, similar to that of the permanent income models. In contrast to the permanent income models, the investor in my model does not perfectly observe the exact fundamental values of her asset holdings; instead, she must make her consumption decision based on her *perceived* wealth level. Given limited information processing capacity, her inference is imperfect. As a result, her consumption only adjusts to fundamental shocks to her wealth with a delay.

The investor's reaction speed increases monotonically with information capacity as discussed in Proposition 2. In the limit, when capacity is infinite and the investor is able to instantaneously infer any fundamental innovations, her consumption immediately adjusts to changes in her wealth. With a finite capacity, the investor chooses to consume based on her best knowledge. Although consumption is still a random walk in the probability measure based on the information processed by the investor, namely, $E(C_{t+\Delta t}|\hat{F}_t) = C_t$, consumption can be predicted by an econometrician who possesses a greater information processing ability. If the econometrician employs variables to form a better inference of the fundamental shocks to consumers' wealth, these variables would naturally have predictive power for future consumption. This seeming violation of the permanent income hypothesis may be driven by the differences in information sets or, equivalently, information processing capacities between consumers and the econometrician.

The optimal capacity allocation I derive in my model provides the social optimal allocation with the most efficient use of information processing. In reality, externalities that result in inefficient allocations may lead to even slower reactions of aggregate consumption to fundamental shocks. Thus, the consumption rule in my model provides an upper bound on how fast aggregate consumption can adjust to innovations in the economy.

B. Cross-Sectional Implications

The investor's capacity allocation determines the cross-sectional differences in the information environments of stocks as well as their price dynamics. In this subsection, I discuss these issues by focusing on the effects generated by the cross-sectional differences across assets' fundamental volatilities.[8]

For simplicity, suppose that all risky assets share the same parameters such as $\{\bar{F}, \lambda, \sigma_D\}$, except for the fundamental volatility. I normalize the fundamental volatility per share of an asset (σ_F) to be equal, as well as that of each asset's fundamental factor loading (ϕ). The difference in the total fundamental volatility

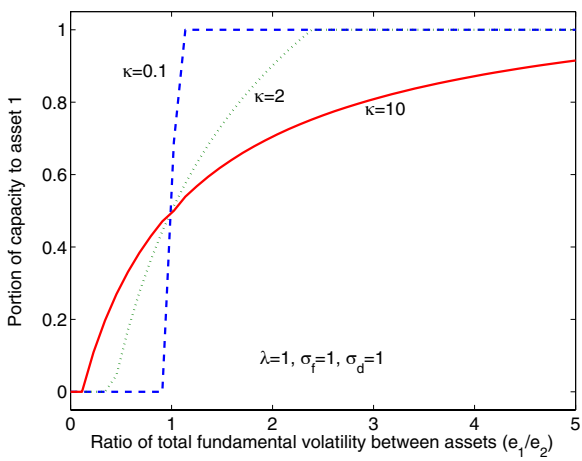
[8]Another parameter that might be of interest is the mean reversion speed of assets' fundamentals, λ_i. A smaller value of λ_i implies that shocks to firm i are more persistent and thus more relevant for the long-run fundamentals. It can be shown that, controlling for the other parameters, the investor allocates more capacity to a stock with a smaller value of the mean reversion parameter. For the sake of space, I omit the related implications.

across assets is reflected in the difference in the number of shares of assets, e_i.[9] The following proposition establishes the difference in the capacity allocation generated by the differences in fundamental volatility.

Proposition 3. For any two assets i and j in the portfolio, if asset i has a greater total fundamental volatility, then $\kappa_i \geq \kappa_j$. As a result, there is less uncertainty in the investor's belief about asset i: $\Sigma_i \leq \Sigma_j$.

Figure 2 further illustrates the capacity allocation in a two-asset example with three different levels of total capacity. When the ratio between the two assets' total fundamental volatility, e_1/e_2, has an intermediate value, both assets receive positive allocation, and the percentage of capacity allocated to asset 1 increases monotonically with e_1/e_2. When e_1 is either too small or too large relative to e_2, corner solutions emerge: the investor allocates all or none of the capacity to asset 1. In particular, when the investor is extremely capacity constrained (small κ), slight differences in fundamental volatility can result in dramatically different capacity allocations. When the total capacity increases, the capacity allocation becomes less sensitive to the difference in fundamental volatility.

FIGURE 2
Optimal Capacity Allocation in a Two-Asset Example



1. Supply of Information

Proposition B suggests that larger stocks, which usually have more total fundamental volatility, are likely to receive more attention from investors. This is consistent with casual observations on the supply of public information. In response to investor demand for information, large stocks and industry leaders typi-

[9]This rescaling does not change the capacity allocation problem since capacity measured using entropy is scale independent (see Proposition 2). Also note that the number of shares, e_i, after the normalization could be different from the actual number of shares outstanding depending on the normalization factor. Note that e_i measures the units of fundamental volatility and captures the relative difference in the total fundamental volatility of assets. Subsequently, all per share variables should be interpreted as per unit of fundamental volatility.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

cally receive much more attention from the financial media and the popular press. They also receive more coverage from sell-side analysts, such as more frequent recommendations and earnings forecasts. In fact, Grant (1985) finds that NYSE firms have a significantly greater number of interim news items in *The Wall Street Journal* than OTC firms. Shores (1990) investigates Nasdaq companies and finds that coverage by *The Wall Street Journal* increases monotonically with firm size, with a correlation coefficient of 0.46. Barth, Kasznik, and McNichols (2001) show that the correlation between the logarithm of the market value of stocks and the number of analysts covering a firm is 0.79.[10]

It is useful to note that news from sources such as *The Wall Street Journal* and the *New York Times* and reports from many sell-side analysts are public information and are widely accessible by many market participants. The demand for such information is largely driven by the need of investors to reduce uncertainty faced in their portfolio and consumption decisions, as analyzed in this model. The supply of this public information is different from the acquisition of nonpublic information emphasized by microstructure models in which speculative trading profits are the key determinant.

2. Speed of Price Adjustments

A larger capacity allocation to a stock is directly reflected in the speed of its price adjustments to fundamental shocks. According to the discussion in Section VI, the investor gradually processes information to incorporate the effects of fundamental shocks into her inferences about the assets' fundamentals. As a result, the equilibrium prices, which are determined by the investor's belief as in Theorem 1, also exhibit a delayed reaction to fundamental shocks. In particular, this delay decreases with the capacity allocated to an asset. Since larger stocks tend to attract more information capacity, their prices should incorporate fundamental shocks at faster speeds and therefore become more informative. This is consistent with the empirical findings. Collins, Kothari, and Rayburn (1987) find that while larger firms' stock prices contain valuable information about future earnings, there is no significant relation between prices and future earnings for small firms. Freeman (1987) finds that security prices of firms in the largest size quartile start to reflect future earnings three months earlier than those in the smallest quartile, and this difference in price informativeness is statistically significant.

3. Effect of the Firm's Disclosures

The capacity allocation to a stock also affects the magnitude of its price reaction to surprises in the firm's disclosures. From equations (2) and (4), changes in a stock's price can be expressed as

$$\begin{aligned} dP_i(\hat{F}_i) &= \frac{1}{\lambda_i + r} d\hat{F}_{i,t} \\ &= -\frac{\lambda_i}{\lambda_i + r}\left(\hat{F}_{i,t} - \bar{F}_i\right) dt + \frac{1}{\lambda_i + r}\frac{\Sigma_i}{\sigma_{D,i}^2} dw_{D,i} + \frac{2\kappa_i}{\lambda_i + r} dw_{S,i}. \end{aligned}$$

[10]There are other factors that might also determine the degree of analyst coverage, such as institutional ownership, the investment banking or market making opportunities for the brokerage house that the analyst is associated with, and cost of coverage (see, for example, Brennan and Hughes (1991)).

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

The dividend shocks, $dw_{D,i}$, can be interpreted as information revealed by firms' voluntary disclosures such as earnings announcements. The magnitude of price reaction to dividend shocks is proportional to $\Sigma_i/\sigma^2_{D,i}$. It is easy to show that this magnitude decreases with the capacity allocation, κ_i. Intuitively, if κ_i is higher the investor has less uncertainty about the fundamental factor $F_{i,t}$, and she will put a smaller weight on dividend news when updating her inferences. Hence, the price responds less to dividend surprises. In this sense, information processed by the investor preempts information in the firm's future public disclosures and smoothes out the responses of stock prices.

To the extent that larger firms tend to receive more capacity allocation, my model implies that their prices react less to earnings surprises of a given magnitude. Atiase (1985), Shores (1990), and Lobo and Mahmoud (1989) show that stock return variability at earnings announcements is inversely related to firm size and analyst coverage. Freeman (1987) further shows that for a given level of earnings surprise, small stocks have greater price reactions than large stocks.

4. Alternative Mechanisms

The cross-sectional differences in the price informativeness of stocks may also depend on the amount of private information acquired by speculators. Since larger stocks may have a greater number of noise (or liquidity) traders, they might also induce more information acquisition or more aggressive trading by speculators. Whether this leads to more informative prices in equilibrium is determined by the trade-off between the increase in noise trading and the corresponding increase in informed trading. The net effect is to depend crucially on market structure. Kyle (1984) studies an imperfectly competitive equilibrium with risk-neutral informed traders and market makers. He shows that increased noise trading may give rise to more entry of informed traders and their competition improves price informativeness. On the other hand in a competitive equilibrium with risk-averse traders, Grossman and Stiglitz (1980) demonstrate that when noise trading increases, the increase in the proportion of informed traders will exactly offset the effect of noise trading, thus the equilibrium price informativeness remains unchanged (Theorem 4 in their paper). In models with imperfect competition and a fixed number of risk-neutral informed traders, Kyle (1984), (1985) also shows that, although informed traders trade more aggressively as the amount of noise trading increases, the price informativeness does not change in equilibrium.

Nevertheless, it is an important empirical issue to disentangle the cross-sectional effects caused by investors' capacity constraints from those caused by private information acquisition. Such an analysis would require careful controls for market structure, fundamental volatility, and intensity of noise trading.

VIII. Further Discussion

Investors' capacity allocations can also affect the correlation of stock prices, but to explicitly analyze such an effect in the current framework requires solving a series of nonlinear equations that are typically not tractable as discussed in Section VI. Intuitively, when there are common factors among asset fundamentals, a

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

capacity-constrained investor allocates attention across both common and firm-specific factors. When aggregate uncertainty becomes large, the investor may choose to shift more attention to common factors, resulting in an amplification of comovements at such times. When the total capacity is below a certain level, the investor will only analyze the common factors and ignore firm-specific factors that result in categorization behavior such as style investing discussed by Barberis and Shleifer (2002). These implications are analyzed in Peng and Xiong (2004).

While my model studies a stationary equilibrium, the results are also useful in predicting how the capacity allocation of investors may react to changes in the fundamental uncertainty of assets. For example, a distressed company, which experiences a decline in its market capitalization, may face more uncertainty per dollar of the capitalization. Whether it is worthwhile to devote more capacity to the company relative to other firms with a similar size depends on the total uncertainty the company faces. When the capacity allocation changes over time in response to changes in the economic environment, it may also be important to consider the cost in adjusting capacity allocation. If investors have more experience in analyzing one type of company, reallocating capacities to other types of companies may be costly.

This paper highlights the effects of capacity-constrained learning in a model with a representative investor. While it is a challenging task to solve the problem of the capacity allocation of investors in a model with many investors, the current framework provides a useful foundation to analyze additional implications that may be generated by the interaction among investors. An interesting feature is that an investor can choose to free ride on the information revealed through prices, rather than to process information herself. However, learning from prices is not exactly free, since doing so requires knowledge of the structure of the market including the presence and risk preference of other traders, their information quality, and the intensity of liquidity trading (see Grossman and Stiglitz (1980)). It also takes time and attention for investors to analyze market conditions that may change randomly in reality. Thus, in a decentralized model an investor's problem involves allocating attention between processing her own fundamental information and extracting the information of others from prices. Such a model may generate further implications on trading among investors in addition to the information dynamics analyzed in the current model.

Another important feature of the investor's learning process is about "learning what to learn." This model, along with the standard rational expectations equilibrium models, assumes that the investor knows the structure of the economy and, more precisely, the processes and parameters that determine the evolution of asset fundamentals. Investors in the real world often face uncertainty about their model of the economy, thus they may not know which types of information are worth processing. Recent theoretical work applies the techniques of robust control to analyze agents' robust decision rules under potential model misspecification (see Hansen and Sargent (2004) for a review). Facing model uncertainty, a capacity-constrained investor may devote her attention to reducing model errors, as well as to decreasing fundamental uncertainty.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

IX. Conclusion

This paper provides a continuous-time equilibrium model to analyze the learning process of a representative investor with a limited information processing capacity. Facing multiple sources of fundamental uncertainty, the investor optimally allocates her capacity across various sources to minimize her wealth uncertainty and to make intertemporal consumption decisions more efficient. The paper shows that the capacity constraint generates delayed consumption behavior. In addition, the model predicts that assets with greater total fundamental volatility will attract more capacity allocation from the investor. Thus, they generate a greater supply of information, their prices incorporate fundamental shocks at a faster speed, and their prices exhibit less volatility to exogenous announcements by firms. These implications are consistent with empirical findings in the literature.

Appendix for Proofs

A. Proof of Lemma 1

I apply Theorem 12.7 in Liptser and Shiryayev (1978) to derive that

$$\begin{aligned} d\hat{F} &= -\Lambda(\hat{F}-\bar{F})dt + \Sigma\Phi^T\left(\Omega_D\Omega_D^T\right)^{-1}(dD-\Phi\hat{F}dt) \\ &\quad + \Sigma\left(\Omega_S\Omega_S^T\right)^{-1}(dS-\hat{F}dt), \\ d\Sigma_t &= -\Lambda\Sigma - \Sigma\Lambda^T + \Omega_F\Omega_F^T - \Sigma\Phi^T\left(\Omega_D\Omega_D^T\right)^{-1}\Phi\Sigma^T \\ &\quad - \Sigma\left(\Omega_S\Omega_S^T\right)^{-1}\Sigma^T. \end{aligned} \tag{A-1}$$

Note that only $D(t)$ and $S(t)$ are in the investor's information set I_t. The innovations $z_F(t)$, $z_D(t)$, and $z_S(t)$ are not. I denote the normalized surprises from the dividend flow and the signal flow as

$$dw_D = \Omega_D^{-1}\left(dD_t - \Phi\hat{F}_t dt\right), \quad dw_S = \Omega_S^{-1}\left(dS_t - \hat{F}_t dt\right). \tag{A-2}$$

These surprises are in the information set, I_t, and are mutually independent standard Wiener processes. I can rewrite the investor's conditional belief $\hat{F}$ as a diffusion process in I_t, and obtain equation (2). In a stationary equilibrium, Σ_t becomes constant and its value is determined so that the right-hand side of equation (A-1) is zero.

B. Proof of Theorem 1

To solve the equilibrium, I conjecture an equilibrium price function and solve the investor's consumption and portfolio problems. In equilibrium, the conjectured price function is verified using the market-clearing condition. This approach is similar to that of Campbell and Kyle (1993) and Wang (1993).

Given the investor's information set, I_t, the state variables in the asset markets are the investor's conditional beliefs, $\hat{F}_t$. I conjecture that asset prices are affine functions of $\hat{F}$,

$$P(\hat{F}) = A\hat{F} + B, \tag{A-3}$$

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

where A is a $m \times k$ matrix, and B is a $m \times 1$ vector. Based on this price function, the excess return process can be derived as a diffusion process in the investor's information set,

$$dQ = dP + dD - rPdt = \mu dt + \Omega dw, \tag{A-4}$$

where

$$\mu = (-A\Lambda + \Phi - rA)\hat{F} + A\Lambda\bar{F} - rB, \tag{A-5}$$

$$\Omega = \left(A\Sigma\Phi^T\Omega_D^{-1} + \Omega_D \;\; A\Sigma\Omega_S^{-1}\right), \quad \text{and} \tag{A-6}$$

$$w = \left(w_D^T, w_S^T\right)^T,$$

where $\mu_{m\times 1}$ represents the risk premia, $\Omega_{m\times(m+k)}$ is the matrix of return volatility, and $w_{(m+k)\times 1}$ represents the innovations to the return processes.

Following Merton (1971), I solve the investor's problem using the Bellman equation,

$$0 = \max_{C,X} u(C) - \beta J + J_W(rW - C + X^T\mu) + \frac{1}{2}J_{WW}X^T\Omega\Omega^T X.$$

The following first-order condition can be derived as

$$u'(C) = J_W,$$

$$X = \frac{J_w}{-J_{WW}}(\Omega\Omega^T)^{-1}\mu.$$

The value function is solved as

$$J(W) = -\frac{e^{-r\gamma W}}{r\gamma}\exp\left[\frac{1}{r}\left(r - \beta - \frac{1}{2}\mu^T(\Omega\Omega^T)^{-1}\mu\right)\right], \tag{A-7}$$

$$X = \frac{1}{r\gamma}(\Omega\Omega^T)^{-1}\mu, \quad \text{and} \tag{A-8}$$

$$C = rW - \frac{1}{r\gamma}\left(r - \beta - \frac{1}{2}\mu^T(\Omega\Omega^T)^{-1}\mu\right). \tag{A-9}$$

The market-clearing condition $X = e$ implies that the equilibrium risk premium is

$$\mu = r\gamma\Omega\Omega^T e. \tag{A-10}$$

Using the definition of $\Omega\Omega^T$ from equation (A-6), combined with the stationary level of the variance-covariance matrix of the belief, Σ, in equation (3), I decompose the variance-covariance matrix of the asset returns, $\Omega\Omega^T$, and express it as in equation (5).

Based on equations (A-5) and (A-10), the constant vectors A and B can be derived as

$$A = \Phi(\Lambda + rI_k)^{-1} \quad \text{and} \quad B = r^{-1}A\Lambda\bar{F} - \gamma\Omega\Omega^T e. \tag{A-11}$$

Thus, the equilibrium asset prices are given in equation (4).

To determine the optimal variance-covariance structure of the signal flow under the capacity constraint, I examine the relation between the investor's value function and the variance-covariance structure in the signal flow. The return variance enters the investor's value function from two sources that equation (A-7) shows. The first is the squared Sharpe ratio, $\mu^T(\Omega\Omega^T)^{-1}\mu$. The second is the investor's wealth, W, which is the sum of the investor's cash holding, K, and asset holding, $e^T P$,

$$W = K + e^T P = K + r^{-1}e^T\Phi\bar{F} + e^T\Phi(\Lambda + rI_k)^{-1}(\hat{F} - \bar{F}) - \gamma e^T\Omega\Omega^T e. \tag{A-12}$$

Substituting the investor's wealth and the equilibrium risk premium, the value function can be expressed as in equation (8).

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

C. Proof of Lemma 2

Define $X' = aX$ and $Y' = bY$. The joint density function of X' and Y' is

$$f_{X'Y'}(x', y') = \frac{1}{ab} f(x'/a, y'/b). \tag{A-13}$$

The marginal density functions of X' and Y' are

$$f_{X'}(x') = \frac{1}{a} f(x'/a), \quad f_{Y'}(y') = \frac{1}{b} g(y'/b).$$

Then the amount of information of bY to aX is given by

$$\begin{aligned} \mathrm{I}(aX|bY) = \mathrm{I}(X'|Y') &= \int\int f_{X'Y'}(x', y') \log \frac{f_{X'Y'}(x', y')}{f_{X'}(x') g_{Y'}(y')} dx' dy' \\ &= \int\int \frac{1}{ab} f(x'/a, y'/b) \log \frac{f(x'/a, y'/b)}{f(x'/a) g(y'/b)} dx' dy' \\ &= \int\int f(x, y) \log \frac{f(x, y)}{f(x) g(y)} dx dy = \mathrm{I}(X; Y). \end{aligned}$$

D. Proof of Proposition 1

The entropy of a vector of the random variable $X = (x_1, x_2, \ldots, x_n)$, with a joint normal distribution, $N(\bar{X}, \Sigma_X)$, is

$$\mathrm{H}(X) = \frac{1}{2} \log |\Sigma_X| + \frac{n}{2} \log(2\pi e).$$

Since the investor's belief has a multivariate Gaussian distribution, the rate of change in the entropy of the investor's belief is

$$\begin{aligned} \frac{1}{2} \frac{d \log |\Sigma|}{dt} &= \frac{1}{2} \mathrm{vec}\left(\frac{d}{d\Sigma} \log |\Sigma| \right)^T \mathrm{vec}\left(\frac{d\Sigma}{dt} \right) \\ &= \frac{1}{2} \mathrm{vec}(\Sigma^{-1})^T \mathrm{vec}\left(\frac{d\Sigma}{dt} \right) = \frac{1}{2} \mathrm{tr}\left((\Sigma^{-1})^T \frac{d\Sigma}{dt} \right). \end{aligned}$$

In this derivation, I use the formula, $d(\log |\Sigma|)/d\Sigma = (\Sigma^{-1})^T$.

Substituting $d\Sigma/dt$ with equation (A-1), I have

$$\begin{aligned} \frac{1}{2} \frac{d \log |\Sigma|}{dt} &= \frac{1}{2} \mathrm{tr}\left((\Sigma^{-1})^T \Omega_F \Omega_F^T \right) - \frac{1}{2} \mathrm{tr}((\Sigma^{-1})^T (\Lambda\Sigma + \Sigma\Lambda^T)) \\ &\quad - \frac{1}{2} \mathrm{tr}\left(\Phi^T \left(\Omega_D \Omega_D^T \right)^{-1} \Phi \Sigma^T \right) - \frac{1}{2} \mathrm{tr}\left(\left(\Omega_S \Omega_S^T \right)^{-1} \Sigma^T \right). \end{aligned}$$

The last term represents the rate of entropy reduction caused by the signals processed by the investor. The amount of information processed by the investor per unit of time is bounded by her information capacity, $\frac{1}{2} \mathrm{tr}(\Sigma(\Omega_S \Omega_S^T)^{-1}) \leq \kappa$.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

E. Proof of Theorem 2

Since the choice set for $(\kappa_1, \ldots, \kappa_k)$ specified by the capacity constraint is bounded and closed, the solution to the optimization problem always exists according to the Weierstrass theorem. Furthermore, the set is also convex. The objective to be minimized is a linear combination of Σ_i, and Σ_i is a convex function of κ_i. Thus, the objective is jointly convex with respect to $(\kappa_1, \ldots, \kappa_k)$. As a result, the minima must be unique.

I prove by contradiction that the capacity constraint is binding for any optimal solution to the capacity allocation problem. Define the optimal solution as $X^* = (\kappa_1^*, \ldots, \kappa_k^*)$. If $\sum_{i=1}^{k} \kappa_i^* < \kappa$, then I can always find a positive number δ, so that $\kappa_m' = \kappa_m^* + \delta$ and $\sum_{i\neq m}^{k} \kappa_i^* + \kappa_m' \leq \kappa$. I define the new solution vector as $X' = (\kappa_i^*, \text{ for } i \neq m;\ \kappa_m', \text{ for } i = m)$ and the objective function in equation (12) as $Q(X)$. Since Σ_m decreases with κ_m, I have $\Sigma_m(\kappa_m^*) > \Sigma_m(\kappa_m')$. Then $Q(X^*) > Q(X')$. Thus, any solution where the capacity constraint is not binding cannot be an optimal solution to the capacity allocation problem.

The first-order conditions are a direct application of Kuhn-Tucker conditions.

F. Proof of Proposition 2

Combining equations (1) and (2), I have the following linear system of diffusion equations,

$$dF_i = -\lambda_i (F_i - \bar{F}_i)\, dt + \sigma_{Fi} dz_{Fi}, \quad \text{and} \tag{A-14}$$

$$d\hat{F}_i = \left[h_i (F_i - \bar{F}_i) - (\lambda_i + h_i)\left(\hat{F}_i - \bar{F}_i\right)\right] dt + \frac{\Sigma_i}{\sigma_{Di}} dz_{Di} + \frac{\Sigma_i}{\sigma_{Si}} dz_{Si}, \tag{A-15}$$

where h_i is a constant as defined in equation (16). It is easy to verify F_i in equation (14) as the solution to equation (A-14). To solve for $\hat{F}_i$, I define $g_i = \hat{F}_i - F_i$ as the error of the investor's belief. From equations (A-14) and (A-15), I have

$$dg_i = -(\lambda_i + h_i) g_i dt - \sigma_{Fi} dz_{Fi} + \frac{\Sigma_i}{\sigma_{Di}} dz_{Di} + \frac{\Sigma_i}{\sigma_{Si}} dz_{Si}.$$

The solution for the error term is

$$g_{it} = g_{i0} \mathrm{e}^{-(\lambda_i + h_i)t} - \int_{s=0}^{t} \mathrm{e}^{-(\lambda_i + h_i)(t-s)} \sigma_{Fi} dz_{Fi}(s) + \int_{s=0}^{t} \mathrm{e}^{-(\lambda_i + h_i)(t-s)} \frac{\Sigma_i}{\sigma_{Di}} dz_{Di}(s) + \int_{s=0}^{t} \mathrm{e}^{-(\lambda_i + h_i)(t-s)} \frac{\Sigma_i}{\sigma_{Si}} dz_{Si}(s).$$

The error mean reverts at the rate $\lambda_i + h_i$. Since $\hat{F}_i = F_i + g_i$, I can solve for the investor's belief $\hat{F}_i$ as given in equation (15).

G. Proof of Proposition 3

I prove this by contradiction. Assume that $\kappa_i \leq \kappa_j$. The first-order condition in Theorem 2 implies that $e_i^2 \Sigma_i'(\kappa_i) = e_j^2 \Sigma_j'(\kappa_j)$ for any interior solutions. From differentiating equation (13), I obtain $\Sigma'(\kappa) < 0$ and $\Sigma''(\kappa) > 0$. With all the other parameters being the same, $\kappa_i \leq \kappa_j$ implies that $\Sigma_i'(\kappa_j) \leq \Sigma_j'(\kappa_j) < 0$. Since $e_i > e_j$, $e_i^2 \Sigma_i'(\kappa_i) < e_j^2 \Sigma_j'(\kappa_j)$ and this contradicts the first-order condition. Therefore, the investor always allocates more capacity to asset i: $\kappa_i > \kappa_j$. This directly implies that $\Sigma_i < \Sigma_j$.

References

Abreu, D., and A. Rubinstein. "The Structure of Nash Equilibrium in Repeated Games with Finite Automata." *Econometrica,* 56 (1988), 1259–1281.

Admati, A. "The Informational Role of Prices—A Review Essay." *Journal of Monetary Economics,* 28 (1991), 347–361.

Admati, A., and P. Pfleiderer. "A Theory of Intraday Patterns: Volume and Price Variability." *Review of Financial Studies,* 1 (1988), 3–40.

Atiase, R. K. "Predisclosure Information, Firm Capitalization, and Security Price Behavior around Earnings Announcements." *Journal of Accounting Research,* 23 (1985), 21–36.

Barberis, N., and A. Shleifer. "Style Investing." *Journal of Financial Economics,* 68 (2003), 161–199.

Barth, M.; R. Kasznik; and M. McNichols. "Analyst Coverage and Intangible Assets." *Journal of Accounting Research,* 39 (2001), 1–34.

Brennan, M., and P. Hughes. "Stock Prices and the Supply of Information." *Journal of Finance,* 46 (1991), 1665–1691.

Brennan, M., and Y. Xia. "Stock Price Volatility and Equity Premium." *Journal of Monetary Economics,* 47 (2001), 249–283.

Campbell, J., and A. Deaton. "Why is Consumption So Smooth?" *Review of Economic Studies,* 56 (1989), 357–374.

Campbell, J. Y., and A. S. Kyle. "Smart Money, Noise Trading and Stock Price Behavior." *Review of Economic Studies,* 60 (1993), 1–34.

Campbell, J. Y.; M. Lettau; B. G. Malkiel; and Y. Xu. "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk." *Journal of Finance,* 56 (2001), 1–43.

Campbell, J. Y., and G. Mankiw. "Permanent Income, Current Income and Consumption." *Journal of Business and Economic Statistics,* 8 (1990), 265–279.

Collins, D.; S. P. Kothari; and J. Rayburn. "Firm Size and the Information Content of Prices with Respect to Earnings." *Journal of Accounting and Economics,* 9 (1987), 111–138.

Cover, T., and J. Thomas. *Elements of Information Theory.* New York, NY: John Wiley & Sons (1991).

Detemple, J. "Asset Pricing in a Production Economy with Incomplete Information." *Journal of Finance,* 41 (1986), 383–391.

Dothan, M., and D. Feldman. "Equilibrium Interest Rates and Multiperiod Bonds in a Partially Observable Economy." *Journal of Finance,* 41 (1986), 369–382.

Dow, J. "Search Decision with Limited Memory." *Review of Economic Studies,* 58 (1991), 1–14.

Easley, D., and M. O'Hara. "Price, Trade Size, and Information in Securities Markets." *Journal of Financial Economics,* 19 (1987), 69–90.

Feldman, D. "Logarithmic Preferences, Myopic Decisions, and Incomplete Information." *Journal of Financial and Quantitative Analysis,* 27 (1992), 619–629.

Foster, F. D., and S. Viswanathan. "A Theory of the Intraday Variations in Volume, Variance and Trading Costs in Security Markets." *Review of Financial Studies,* 3 (1990), 593–624.

Freeman, R. "The Association between Accounting Earnings and Security Returns for Large and Small Firms." *Journal of Accounting and Economics,* 9 (1987), 195–228.

Gennotte, G. "Optimal Portfolio Choice under Incomplete Information." *Journal of Finance,* 41 (1986), 733–746.

Glosten, L., and P. Milgrom. "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics,* 13 (1985), 71–100.

Grant, E. "Market Implications of Differential Amounts of Interim Information." *Journal of Accounting Research,* 18 (1985), 255–268.

Grossman, S. J., and J. E. Stiglitz. "On the Impossibility of Informationally Efficient Markets." *American Economic Review,* 70 (1980), 393–408.

Hall, R. E. "Stochastic Implications of the Life Cycle Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy,* 86 (1978), 971–987.

Hansen, L. P., and T. J. Sargent. "Misspecification in Recursive Macroeconomic Theory." Monograph, Univ. of Chicago and New York Univ. (2004).

Hirshleifer, D.; S. Lim; and S. H. Teoh. "Disclosure to a Credulous Audience: The Role of Limited Attention." Unpubl. Working Paper, Ohio State Univ. (2002).

Hirshleifer, D., and S. H. Teoh. "Limited Attention, Financial Reporting and Disclosure." *Journal of Accounting and Economics,* 36 (2003), 337–386.

Hirst, D. E., and P. E. Hopkins. "Comprehensive Income Reporting and Analysts' Valuation Judgments." *Journal of Accounting Research,* 36 (1998), 47–75.

Holden, C. W., and A. Subrahmanyam. "Long-Lived Private Information and Imperfect Competition." *Journal of Finance,* 47 (1992), 247–270.

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press

Huberman, G., and T. Regev. "Contagious Speculation and a Cure for Cancer: A Nonevent that Made Stock Prices Soar." *Journal of Finance,* 56 (2001), 387–396.
Kalai, E., and W. Stanford. "Finite Rationality and Interpersonal Complexity in Repeated Games." *Econometrica,* 56 (1988), 397–410.
Kyle, A. S. "Market Structure, Information, Futures Markets, and Price Formation." *International Agriculture Trade: Advanced Readings in Price Formation, Market Structure, and Price Instability,* G. Story, A. Schmitz, and A. Sarris, eds. Boulder and London: Westview Press (1984).
Kyle, A. S. "Continuous Auctions and Insider Trading." *Econometrica,* 53 (1985), 1315–1336.
Liptser, R. S., and A. N. Shiryayev. *Statistics of Random Processes.* New York, NY: Spring-Verlag (1977).
Lobo, G. J., and A. A. W. Mahmoud. "Relationship between Differential Amounts of Prior Information and Security Return Variability." *Journal of Accounting Research,* 27 (1971), 116–134.
Merton, R. "Optimal Consumption and Portfolio Rules in a Continuous-Time Model." *Journal of Economic Theory,* 3 (1971), 373–413.
O'Hara, M. *Market Microstructure Theory.* Boston, MA: Blackwell Publishers (1995).
Peng, L., and W. Xiong. "Time to Digest and Volatility Dynamics." Unpubl. Working Paper, Baruch College and Princeton Univ. (2003).
Peng, L., and W. Xiong. "Investor Attention, Overconfidence and Category Learning." Unpubl. Working Paper, Baruch College and Princeton Univ. (2004).
Sims, C. A. "Implications of Rational Inattention." *Journal of Monetary Economics,* 50 (2003), 665–690.
Shores, D. "The Association between Interim Information and Security Returns Surrounding Earnings Announcements." *Journal of Accounting Research,* 28 (1990), 164–181.
Shiller, R. *Irrational Exuberance.* Princeton, NJ: Princeton Univ. Press (2000).
Stapleton R., and M. G. Subrahmanyam. "A Multiperiod Equilibrium Asset Pricing Model." *Econometrica,* 46 (1978), 1077–1096.
Stutzer, M. J. "Simple Entropic Derivation of a Generalized Black-Scholes Option Pricing Model." *Entropy,* 2 (2000), 70–77.
Sundaresan, M. "Constant Absolute Risk Aversion Preferences and Constant Equilibrium Interest Rates." *Journal of Finance,* 38 (1983), 205–212.
Veronesi, P. "How Does Information Quality Affect Stock Returns?" *Journal of Finance,* 55 (2000), 807–837.
Verrecchia, R. E. "Information Acquisition in a Noisy Rational Expectations Economy." *Econometrica,* 50 (1982), 1415–1430.
Wang, J. "A Model of Intertemporal Asset Prices under Asymmetric Information." *Review of Economic Studies,* 60 (1993), 249–282.
West, K. D. "The Insensitivity of Consumption to News About Income." *Journal of Monetary Economics,* 21 (1988), 17–33.
Yan, H. "Estimation Uncertainty and the Equity Premium." Unpubl. Working Paper, Univ. of Texas at Austin (2001).

https://doi.org/10.1017/S0022109000002325 Published online by Cambridge University Press