THE JOURNAL OF FINANCE • VOL. LXII, NO. 4 • AUGUST 2007

Rational Inattention and Portfolio Selection

LIXIN HUANG and HONG LIU*

ABSTRACT

Costly information acquisition makes it rational for investors to obtain important economic news with only limited frequency or limited accuracy. We show that this rational inattention to important news may make investors over- or underinvest. In addition, the optimal trading strategy is "myopic" with respect to future news frequency and accuracy. We find that the optimal news frequency is nonmonotonic in news accuracy and investment horizon. Furthermore, when both news frequency and news accuracy are endogenized, an investor with a higher risk aversion or a longer investment horizon chooses less frequent but more accurate periodic news updates.

STANDARD PORTFOLIO SELECTION THEORIES assume that investors can costlessly obtain and process all the relevant information that affects investment performance. However, this assumption does not hold for most investors, including the most sophisticated ones. Both information production (data collection and noise reduction) and information processing can be very costly in terms of required time, effort, and expenses. These costs induce limited attention.[1] Indeed, Hong, Torous, and Valkanov (2002), among others, provide evidence of investor inattention to important economic news. Motivated by this observation, Bacchetta and van Wincoop (2005) build an overlapping generations model to explain the forward discount and predictability puzzle with rational inattention.[2] Peng and Xiong (2006) show that inattention can lead to "category-learning" behavior and cross-sectional return predictability.

Some unanswered questions in this literature include how to determine an investor's optimal inattention to valuable information and how rational inattention affects his trading strategy. To address these questions, we develop a novel portfolio selection model whereby an investor with a hyperbolic absolute

*Huang is from Georgia State University and City University of Hong Kong and Liu is from Washington University in Saint Louis. We are very grateful to an anonymous referee, Kerry Back, Suleyman Basak, Alex David, Jérôme Detemple, Gary Gorton, Qiang Kang, Richard Kihlstrom, Yuming Li, Jianjun Miao, Rob Stambaugh (the Editor), Yihong Xia, Guofu Zhou, and seminar participants at the 2004 American Finance Association conference, the 2003 Midwest Finance Association conference, Boston University, the Hong Kong University of Science and Technology, the University of Missouri at Columbia, the University of North Carolina at Chapel Hill, and Washington University in Saint Louis for helpful comments. Huang acknowledges research funding provided by a grant from City University of Hong Kong (project no. 7001870).

[1] See, for example, Kahneman (1973) and Fiske and Taylor (1991) for evidence of limited attention documented in social psychology.

[2] In their model, they assume for tractability that inattentive investors live longer than attentive investors and invest a constant fraction of wealth in stock throughout their entire life.

risk aversion (HARA) preference over his terminal wealth can purchase both periodically released news and continuously released news (with different accuracies and costs) about a predictive state variable that predicts the future expected return.[3]

We derive explicit forms for the value function and the optimal trading strategy. The information cost gives rise to rational inattention to news in the sense that the investor chooses to receive the news with only limited frequency and/or limited accuracy. We show that with rational inattention, the trading strategy is "myopic" with respect to news frequency and accuracy, even for an investor with non-log preferences. That is, as long as the current conditional distribution is the same, the optimal trading strategies will be the same irrespective of whether or not one will have more information in the future and thus, all that matters for the investor's *current* optimal trading strategy is the *current* conditional distribution of the predictive variable. Intuitively, periodic news updates may have two opposing effects on the optimal trading strategy. On the one hand, the investor's estimation risk is smaller with periodic news updates than without them. Therefore, it seems reasonable to conjecture that given the same conditional distribution, the investor would invest more in the case with periodic news updates than in the case without them. On the other hand, since the investor can observe the predictive variable at the beginning of the next period, it also seems reasonable to conjecture that the investor would reduce the current stock investment, decreasing the damage from possible estimation error and waiting to choose a better trading strategy until the next period when he can make a better estimation. We show that both of these conjectures are incorrect. This somewhat surprising independence stems from a fairly simple fact: Access to more accurate or more frequent news in the future improves the investor's utility by a constant multiple across all states and thus does not affect the optimal trading strategy. However, because periodic news coming in the future will almost surely change the conditional distribution, the trajectories of the optimal trading strategies will be almost surely different across the two models with or without periodic news.

Furthermore, we find that trading only at periodic news update times to reduce estimation risk at trading times (e.g., a discrete-time model) results in significant underinvestment and welfare loss. In addition, when the investor underestimates (overestimates) the predictive variable, he tends to underinvest (overinvest). However, if news about the predictive variable is very noisy, the investor may still underinvest even given good news. Finally, our model predicts that trading volume increases at news update times due to the decrease in uncertainty about the predictive variable at these times.

We show that the optimal news frequency increases with the unconditional volatility of the predictive variable and decreases with information cost and risk aversion. Intuitively, a more volatile predictive variable increases the value of news updates, a higher information cost increases the marginal cost of news

[3] In contrast to our approach, Sims (2003) and Moscarini (2004) impose an information-processing capacity constraint on economic agents.

updates, and greater risk aversion decreases the benefit of more frequent news since the investor invests less in stock. In contrast, we find that the optimal news frequency displays nonmonotonic patterns in news accuracy, investment horizon, and the correlation between the stock return and the predictive variable. First, when news is very noisy, it is not very useful and it is not worthwhile to pay for frequent news, whereas when news is very precise, there is no need for very frequent news because the estimation risk based on a precise prior is small. With respect to horizon, a very short investment horizon decreases the benefit of more frequent news, whereas a very long horizon increases the cost of more frequent news. Finally, although a high correlation between the stock return and the state variable makes learning effective and reduces the necessity for more frequent news, a very low correlation also lowers the optimal news frequency because learning becomes ineffective and the value of news updates is reduced.

When we enodogenize news accuracy, we find that an investor with a higher risk aversion or a longer investment horizon chooses less frequent but more accurate periodic news updates. Intuitively, a more risk-averse investor invests less in the stock and is also more averse to estimation risk. As the investment horizon increases, total information costs increase for a given news frequency and accuracy. This induces the investor to decrease news frequency. To partly compensate for the increase in the estimation risk caused by less frequent news updates, the investor increases news accuracy.

We also generalize our model to allow for intertemporal consumption, nonlinear dynamics, and continuous information choice. Numerical results suggest that the results above are robust to these generalizations.

The model closest to ours in the literature is that of Detemple and Kihlstrom (1987). They consider a similar portfolio selection problem with continuous news available for purchase throughout the investment horizon. In contrast, our model allows for both continuous and periodic news and focuses more on the analysis of periodic news since an important manifestation of inattention is the choice of limited news frequency and, in practice, periodic news (e.g., economic statistics announcements) is typically more accurate than continuous news.

Our study is also closely related to the literature on portfolio selection with return predictability (e.g., Kandel and Stambaugh (1996), Xia (2001), and Detemple, Garcia, and Rindisbacher (2003)). Most of the models in this literature assume that all predictive variables are accurately observable at all trading times. For example, the continuous-time models of Xia (2001) and Detemple et al. (2003) assume that all predictive variables are continuously and accurately observable to the investor, and the discrete-time models of Kandel and Stambaugh (1996) and Stambaugh (1999) assume that at the beginning of every period the investor can always accurately observe predictive variables.[4]

[4] One of the justifications for a discrete-time model is the presence of transaction costs. However, Liu (2004) and Liu and Loewenstein (2002) show that in this case the optimal rebalancing frequency is not deterministic but stochastic, and thus is different from the predictive variable observation frequency.

Although the investor in our paper trades continuously, the cost of obtaining and processing information about the predictive variable makes it optimal for him to receive the periodic news only with limited frequency. In addition, the periodic news frequency in our paper is endogenously determined rather than exogenously given as in the existing literature.

Our model is also different from portfolio selection models with return predictability and parameter uncertainty. Kandel and Stambaugh (1996) consider the asset allocation implications of return predictability in the presence of predictive parameter estimation risk. They find that although the predictive regression seems weak when described by usual statistical measures, return predictability can have a substantial influence on stock trading strategy. Xia (2001) also studies the impact of parameter uncertainty on the optimal trading strategy in the presence of return predictability. She shows that uncertainty about the predictive relation leads to a state-dependent relationship between the optimal portfolio choice and the investment horizon. In contrast, we study how information production and processing costs induce inattention to the predictive variable and how this inattention affects the stock trading strategy.

It is worth noting that because the investor in our model needs to combine continuous filtering (from continuously observing the stock price and news) and periodic filtering (from periodic news updates), the problem faced by the investor is different from those problems considered by most of the portfolio selection models with dynamic learning (e.g., Gennotte (1986), Detemple and Kihlstrom (1987), Detemple (1991), Brennan (1998), and Xia (2001)). In most of these models, the unknown variables that an investor tries to learn are never directly observed. For example, Brennan (1998) assumes that the expected return of a stock is constant and unobservable throughout the investment horizon. In our model, in contrast, the investor can obtain periodic news updates about the predictive variable. These periodic updates periodically correct the investor's estimation errors and significantly change the investor's trading strategy. The filtering problem in this paper is also different from that in Duffie and Lando (2001). In their model, investors observe only periodic signals on the unobservables and therefore do not face the continuous filtering problem on top of the discrete filtering, as we do in this paper.

The rest of the paper is organized as follows. In Section I we describe our portfolio selection model with rational inattention. Section II provides the solution in an explicit form. In Section III we calibrate an empirical model using the Center for Research in Security Prices (CRSP) data and consumption-to-wealth ratio data. In Sections IV, V, and VI, we use the calibrated model to study the optimal news frequency, the optimal news accuracies, and the effect of rational inattention on the optimal trading strategy. In Section VII, we generalize the model to allow for intertemporal consumption, nonlinear dynamics, and continuous information choice to verify the robustness of the main results. Section VIII concludes. We provide technical details in the appendices.

I. The Model

There are two assets that an investor can trade continuously in a financial market. The first asset is a risk-free money market account with the constant interest rate of r. The second asset (stock) is risky and its cum-dividend price S_t satisfies

$$\frac{dS_t}{S_t} = (\mu_0 + \mu_1 X_t)\,dt + \sigma_S\,dZ_{1t}, \tag{1}$$

where μ_0, μ_1, and σ_S are all constants, Z_{1t} is a one-dimensional Brownian motion, and X_t is a predictive variable that evolves according to[5]

$$dX_t = (g_0 + g_1 X_t)\,dt + \rho\sigma_X\,dZ_{1t} + \sqrt{1-\rho^2}\sigma_X\,dZ_{2t}, \tag{2}$$

where g_0, g_1, σ_X, and $\rho \in [-1, 1]$ are all constants and Z_{2t} is a one-dimensional Brownian motion that is independent of Z_{1t}.[6]

As opposed to the standard literature, we assume that the predictive variable process X_t is not freely observable to the investor. Instead, there is an information market in which (possibly inaccurate) news about the predictive variable is available only at a cost. The investor can obtain both periodic and continuous news updates. The periodic news is designed to model periodic news announcements such as economic forecasts and newsletters; the continuous news is designed to model in-house information production or inferences from other continuous-information streams such as media coverage of related economic events. For the periodic news updates, we assume that the investor can receive N news updates at $t_i = (i-1)\frac{T}{N}(i = 1, 2, \ldots, N)$, where $T > 0$ is the investor's investment horizon. The ith update y_{t_i} received at $t = t_i$ is a noisy signal of X_t:

$$y_t = X_t + \varepsilon_t, \tag{3}$$

where ε_t represents the noise in the news. We assume that ε_t is serially uncorrelated, independent of any other random variables at any time, and identically normally distributed with mean zero and standard deviation α_ε^{-1}. We interpret α_ε to represent the accuracy of periodic news. Note that if $\alpha_\varepsilon = 0$, it is equivalent to no periodic news update, whereas if $\alpha_\varepsilon = \infty$, it is equivalent to perfectly accurate news updates. The cost for the ith news update is $\beta_i(\alpha_\varepsilon)$, so the total cost for all periodic news updates to be paid at time 0 is equal to $\sum_{i=1}^{N} \beta_i(\alpha_\varepsilon)$.[7] For the continuous news, by paying a cost of $\beta_c(\alpha_\nu)$ at time 0, the investor can receive

[5] Papers on return predictability include Keim and Stambaugh (1986), Lee (1992), McQueen and Roley (1993), Kandel and Stambaugh (1996), Patelis (1997), and Stambaugh (1999).

[6] See Appendix A for a more general model that allows for multiple stocks, multiple predictive variables, and time-varying coefficients.

[7] Assuming that paying all the costs up front allows us to keep homogeneity of the value function in the (effective) wealth and thus makes the model tractable. This assumption is also consistent with common practice in subscribing to newsletters or other news services.

continuous news ν_t of accuracy α_ν, where, similar to Detemple and Kihlstrom (1987), ν_t evolves according to

$$d\nu_t = (h_0 + h_1 X_t)\,dt + \sigma_\nu\, dZ_{3t}, \tag{4}$$

where h_0, h_1, and σ_ν are all constants and Z_{3t} is a Brownian motion that is independent of Z_{1t} and Z_{2t}.[8] The accuracy $\alpha_\nu \equiv h_1/\sigma_\nu$ of the continuous news is measured by the signal-to-noise ratio. Following Detemple and Kihlstrom (1987), we assume that both $\beta_c(\cdot)$ and $\beta_i(\cdot)$ are strictly increasing and strictly convex. To simplify exposition, in the rest of the paper, unless we use the words "continuous news," "news" refers to "periodic news."

Before the time 0 news update, the prior of the investor is that X_0 is normally distributed with mean M_{0^-} and variance $V(0^-)$. Let $\mathcal{F}_t$ denote the filtration at time t generated by $\{S_s, \nu_s\}$, $\{y_{t_i}\}$, and the prior $(M_{0^-}, V(0^-))$ for all $s \le t$ and $t_i \le t$. We assume that an investor has HARA preferences over the terminal wealth at time T. Specifically, the investor's utility function is

$$u(W) = \frac{\gamma}{1-\gamma}\left(\frac{\lambda W}{\gamma} + \eta\right)^{1-\gamma},$$

where γ, λ, and η are all constants subject to the restrictions[9]

$$\gamma \neq 1, \lambda > 0, \text{ and } \eta = 1 \text{ if } \gamma = \infty.$$

Given the initial wealth $W_{0^-} > 0$ and the prior $(M_{0^-}, V(0^-))$, the investor's problem is to choose the number $N \in \mathcal{F}_0$ of news updates, the news accuracies $\alpha_\varepsilon, \alpha_\nu \in \mathcal{F}_0$, and a trading strategy $\theta_t \in \mathcal{F}_t$ to maximize expected utility from the terminal wealth, that is,

$$\max_{N,\alpha_\varepsilon,\alpha_\nu,\theta} E[u(W_T)],$$

subject to

$$dW_t = rW_t\,dt + \theta_t(\mu_0 + \mu_1 X_t - r)\,dt + \theta_t \sigma_S\, dZ_{1t}, \tag{5}$$

$$W_0 = W_{0^-} - \sum_{i=1}^{N} \beta_i(\alpha_\varepsilon) - \beta_c(\alpha_\nu), \tag{6}$$

the dynamics (2) of X_t, the news equations (3) and (4), and the constraint that the wealth process W_t is bounded from below, where equation (5) is the budget constraint and equation (6) is the initial wealth after deducting the information cost.

[8] This independency assumption is only for expositional simplicity. Allowing for correlations is a straightforward extension.

[9] The HARA family is rich in the sense that with a suitable adjustment of the parameters one can have a utility function with increasing, decreasing, or constant absolute or relative risk aversion. See Merton (1992, p. 138) for these cases.

II. The Solution

We begin by considering the optimal investment problem for given frequency N and accuracies α_ε and α_ν. We then solve for the optimal N, α_ε, and α_ν that maximize the resulting value function. As in Detemple (1986, 1991) and Gennotte (1986), the investor's problem for given N, α_ε, and α_ν is separable in inference and optimization.[10] In particular, the investor's portfolio selection problem given N, α_ε, and α_ν is equivalent to

$$\max_\theta E\,[u(W_T)],$$

subject to

$$dW_t = rW_t\,dt + \theta_t(\mu_0 + \mu_1 M_t - r)\,dt + \theta_t\sigma_S\,d\hat{Z}_{1t}, \tag{7}$$

where $M_t \equiv E\,[X_t \mid \mathcal{F}_t]$ is the conditional expectation of X_t that, $\forall i = 1, 2, \dots, N$, satisfies

$$dM_t = (g_0 + g_1 M_t)\,dt + \sigma_{M1}(t)\,d\hat{Z}_{1t} + \sigma_{M2}(t)\,d\hat{Z}_{3t}, \quad \forall t \in (t_i, t_{i+1}), \tag{8}$$

where $\sigma_{M1}(t) = \frac{\mu_1}{\sigma_S}V(t) + \rho\sigma_X$, $\sigma_{M2}(t) = \alpha_\nu V(t)$, $V(t) \equiv E\,[(X_t - M_t)^2 \mid \mathcal{F}_t]$ is the conditional variance of X_t satisfying

$$\frac{dV(t)}{dt} = 2g_1 V(t) + \sigma_X^2 - \left(\frac{\mu_1}{\sigma_S}V(t) + \rho\sigma_X\right)^2 - \alpha_\nu^2 V(t)^2, \quad \forall t \in (t_i, t_{i+1}), \tag{9}$$

and $\hat{Z}_{1t}$ and $\hat{Z}_{3t}$ are the (observable) innovation processes satisfying

$$d\hat{Z}_{1t} = \frac{\mu_1}{\sigma_S}(X_t - M_t)\,dt + dZ_{1t},$$

and

$$d\hat{Z}_{3t} = \alpha_\nu(X_t - M_t)\,dt + dZ_{3t}.$$

In addition, at the news update times, conditional mean and conditional variance are updated using Bayes's rule:

$$M_{t_i} = M_{t_i^-} + \frac{V\left(t_i^-\right)}{V\left(t_i^-\right) + \alpha_\varepsilon^{-2}}\left(y_{t_i} - M_{t_i^-}\right), \tag{10}$$

$$V(t_i) = \frac{V\left(t_i^-\right)\alpha_\varepsilon^{-2}}{V\left(t_i^-\right) + \alpha_\varepsilon^{-2}}, \tag{11}$$

where

$$y_{t_i} = X_{t_i} + \varepsilon_{t_i},\, \varepsilon_{t_i} \sim N\left(0, \alpha_\varepsilon^{-2}\right). \tag{12}$$

[10] The separation principle applies trivially because the objective function is independent of the unobservable state variable (see, e.g., Fleming and Rishel (1975, Chap. 4, Sec. 11)).

A brief discussion of the above equations is now in order. Between news updates, the investor infers the conditional distribution of the predictive variable X_t from the observation of stock prices and the continuous news ν_t. By theorem 10.5.1 of Kallianpur (1980), for any $i = 1, 2, \ldots, N$ and time $t \in (t_i, t_{i+1})$, the conditional mean M_t satisfies equation (8) and the conditional variance $V(t)$ satisfies equation (9). Immediately before news y_{t_i} is received at time t_i, the conditional distribution of X_{t_i} is normal with mean $M_{t_i^-}$ and variance $V(t_i^-)$. Upon observing y_{t_i}, this conditional distribution of X_{t_i} is updated. Given equation (12), $M_{t_i^-}$, and $V(t_i^-)$, X_{t_i}, and y_{t_i} are jointly normal and the conditional distribution of X_{t_i} after incorporating news y_{t_i} is normal with mean M_{t_i} and variance $V(t_i)$ as in equations (10) and (11), respectively. It is worth noting that as α_ν decreases to zero, the continuous signal ν_t becomes useless and thus is equivalent to *not* observing ν_t.

A. No Periodic News for the Entire Horizon

In this subsection we assume that no periodic news about the predictive variable X is obtained throughout the entire horizon T (i.e., $\alpha_\varepsilon = 0$). This case provides a basis for developing the solution to the more complicated case with periodic news updates. Similar to Detemple and Kihlstrom (1987), after taking into consideration the information contained in the continuous process ν through equations (8) and (9), the investor's value function only depends on (W, M, t) and is independent of the realized value ν since neither the wealth process W nor the conditional mean process M depends on ν.[11] For a given accuracy of the continuous news α_ν, let J be the value function at t, that is,

$$J(W, M, t; \alpha_\nu) = \max_\theta E[u(W_T) \mid W_t = W, M_t = M].$$

The value function must satisfy the Hamilton–Jacobi–Bellman (HJB) equation

$$J_t + \max_\theta \left\{ \frac{1}{2}\theta^2\sigma_S^2 J_{WW} + \theta(\mu_0 + \mu_1 M - r)J_W + \theta\sigma_S\sigma_{M1}(t)J_{WM} \right\} + rWJ_W + \frac{1}{2}\sigma_M(t)^2 J_{MM} + (g_0 + g_1 M)J_M = 0, \tag{13}$$

and the terminal condition

$$J(W, M, T; \alpha_\nu) = u(W),$$

where $\sigma_M(t)^2 \equiv \sigma_{M1}(t)^2 + \sigma_{M2}(t)^2$. We provide the explicit solution in the following proposition.

Proposition 1: *When there is no periodic news update throughout the entire horizon, the value function is*

[11] Although one may use the conditional volatility $V(t)$ as a state variable in the value function, we combine it with time t for simplicity because $V(t)$ is only a deterministic function of time.

$$J(W, M, t; \alpha_\nu) = U(W, t)e^{(1-\gamma)(c(t)+d(t)M+\frac{1}{2}Q(t)M^2)},$$

where

$$U(W, t) \equiv \frac{\gamma}{1-\gamma}\left(\frac{\lambda W}{\gamma} + \eta e^{-r(T-t)}\right)^{1-\gamma},$$

and the optimal trading strategy satisfies

$$\pi_t^* \equiv \frac{\theta_t^*}{W_t + \frac{\gamma\eta}{\lambda}e^{-r(T-t)}} = \frac{\mu_0 + \mu_1 M_t - r + (1-\gamma)\sigma_S\sigma_{M1}(t)(d(t) + Q(t)M_t)}{\gamma\sigma_S^2}, \tag{14}$$

where Q(t), d(t), and c(t) are as in equations (B4) to (B6) in Appendix B.

Proof: See Appendix C.

The functional form of the value function and the optimal trading strategy follows from the fact that for a HARA investor, $\frac{\gamma\eta}{\lambda}$ is like extra income at time T and the investor's time t effective wealth is equal to his current wealth W_t plus the present value of this extra income ($\frac{\gamma\eta}{\lambda}e^{-r(T-t)}$). To simplify exposition, in the rest of this paper unless specified otherwise we refer to this "effective wealth" simply as "wealth" and π^* simply as the fraction of wealth in stock.

Once we obtain the value function for a fixed α_ν, we then solve the maximization problem

$$\max_{\alpha_\nu} J(W, M, 0; \alpha_\nu)$$

to compute the optimal accuracy α_ν^* for the continuous news. Even though this optimization problem does not appear to have a closed-form solution, it can be easily solved numerically, as we show later in Sections IV, V, and VI.

To better understand the optimal trading strategy, we can rewrite the trading strategy as

$$\begin{aligned}\pi_t^* = &\frac{\mu_0 + \mu_1 M_t - r}{\gamma\sigma_S^2} \\ &+ \frac{(1-\gamma)\rho\sigma_X}{\gamma\sigma_S}(d(t) + Q(t)X_t) \\ &+ \frac{(1-\gamma)\rho\sigma_X Q(t)}{\gamma\sigma_S}(M_t - X_t) \\ &+ \frac{(1-\gamma)\mu_1(d(t) + Q(t)M_t)}{\gamma\sigma_S^2}V(t).\end{aligned} \tag{15}$$

Equation (15) shows that the optimal trading strategy consists of four parts: (1) Myopic trading for wealth growth, (2) hedging against the time-varying investment opportunity set determined by the predictive variable X_t, (3) hedging against the estimation error of the expected stock return, and (4) hedging against the estimation risk of the predictive variable. In effect, the lack of perfect information generates a hedging demand that depends on the estimation error $M_t - X_t$ and the conditional variance $V(t)$ and can therefore cause an investor to over- or underinvest from misestimation.

B. Periodic News Updates

In this section we extend the above analysis to the case in which periodic news about X_t is received at a cost. This captures the empirical fact that, for a cost (e.g, subscribing to analyst services or investment newsletters), an investor can typically obtain periodic news updates about important predictive variables such as the dividend yield, the gross national product (GNP) growth rate, the consumption–wealth ratio, and the inflation rate.

Suppose the number of updates is N, the news accuracy is α_ε, and the cost is $\sum_{i=1}^{N} \beta_i(\alpha_\varepsilon)$. We solve the investor's optimization problem using an iterative method. At the beginning of the Nth period (i.e., $t = t_N$) the results in Proposition 1 apply, and therefore the value function at any time $t \in [t_N, T]$ is

$$J^N(W, M, t; N, \alpha_\varepsilon, \alpha_\nu) = U(W, t)e^{(1-\gamma)(c^N(t)+d^N(t)M+\frac{1}{2}Q^N(t)M^2)},$$

where $c^N(t), d^N(t)$, and $Q^N(t)$ are the respective counterparts of $c(t), d(t)$, and $Q(t)$ in Proposition 1.

Now suppose the investor is in the $(i-1)$th period, for $i = 2, 3, \ldots, N$. Since the investor does not know M_{t_i} before y_{t_i} is revealed, he first takes the expectation with respect to M_{t_i} conditional on $M_{t_i^-}$ and $V(t_i^-)$. Given expressions (10) and (11), we have

$$\begin{aligned} &J^{i-1}\big(W, M, t_i^-; N, \alpha_\varepsilon, \alpha_\nu\big) \\ &\quad = E\big[J^i\big(W, M_{t_i}, t_i\big); N, \alpha_\varepsilon, \alpha_\nu \,\big|\, M_{t_i^-} = M, V\big(t_i^-\big)\big] \\ &\quad = \int_{-\infty}^{\infty} U\big(W, t_i^-\big)e^{(1-\gamma)(c^i+d^i x+\frac{1}{2}Q^i x^2)} n(x)\,dx \\ &\quad = U\big(W, t_i^-\big)e^{(1-\gamma)(\hat{c}^{i-1}+\hat{d}^{i-1}M+\frac{1}{2}\hat{Q}^{i-1}M^2)}, \end{aligned} \tag{16}$$

where $n(x)$ is the normal density function with mean $M_{t_i^-}$ and variance $V_M(t_i)$, with $V_M(t_i)$ being the conditional variance of M_{t_i} given $M_{t_i^-}$ and $V(t_i^-)$, satisfying

$$V_M(t_i) = \frac{V\left(t_i^-\right)^2}{V\left(t_i^-\right) + \alpha_\varepsilon^{-2}},$$

$$\hat{c}^{i-1} = c^i + \frac{(1-\gamma)(d^i)^2 V_M(t_i)}{2\big(1-(1-\gamma)V_M(t_i)Q^i\big)} - \frac{1}{2(1-\gamma)}\log\big(1-(1-\gamma)Q^i V_M(t_i)\big),$$

$$\hat{d}^{i-1} = \frac{d^i}{1-(1-\gamma)V_M(t_i)Q^i}, \quad \hat{Q}^{i-1} = \frac{Q^i}{1-(1-\gamma)V_M(t_i)Q^i}. \tag{17}$$

Equation (16) provides the terminal conditions for solving the investor's problem in period $(i-1)$. Note that the terminal value function J^{i-1} for period $(i-1)$ has the same exponential form as the value function J^i for period i. The stability of the functional form makes it possible to solve the investor's problem across all periods, as we show below.

PROPOSITION 2: *For a given $N > 0, i = 1, 2, \ldots, N$, the value function at time $t \in [t_i, t_{i+1})$ is*

$$J^i(W, M, t; N, \alpha_\varepsilon, \alpha_\nu) = U(W, t)e^{(1-\gamma)(c^i(t)+d^i(t)M+\frac{1}{2}Q^i(t)M^2)}$$

and the optimal trading strategy is

$$\pi_t^{i*} \equiv \frac{\theta_t^{i*}}{W_t + \frac{\gamma\eta}{\lambda}e^{-r(T-t)}} = \frac{\mu_0 + \mu_1 M_t - r + (1-\gamma)\sigma_S\sigma_{M1}(t)\big(d^i(t) + Q^i(t)M_t\big)}{\gamma\sigma_S^2}, \tag{18}$$

where $Q^i(t)$, $d^i(t)$, and $c^i(t)$ are as in equations (B7) to (B9) in Appendix B.

Proof: See Appendix C.

Once we obtain the value function for fixed N, α_ε, and α_ν, we then solve the maximization problem

$$\max_{N,\alpha_\varepsilon,\alpha_\nu} J^1(W, M, 0; N, \alpha_\varepsilon, \alpha_\nu)$$

to compute the optimal number of news updates N^* and the optimal news accuracies α_ε^* and α_ν^*.

In contrast to the case with no news update, equation (18) implies that the fraction of wealth invested in the stock jumps at observation times due to discrete changes in the conditional mean M_t and the conditional variance $V(t)$ at these times. This is consistent with the empirical evidence that trading volume increases immediately after news arrival (e.g., see Woodruff and Senchack (1988), Cready and Mynatt (1991), Balduzzi, Elton, and Green (1997)). To see the effect of a news update more clearly, we next compare the trading strategy immediately before and immediately after news y_{t_i} is received at $t = t_i$. Immediately before y_{t_i} is received, the trading strategy is

$$\frac{\mu_0+\mu_1 M_{t_i^-}-r}{\gamma\sigma_S^2}+\frac{(1-\gamma)\big(\rho\sigma_X\sigma_S+\mu_1 V\big(t_i^-\big)\big)}{\gamma\sigma_S^2}\big(d^{i-1}\big(t_i^-\big)+Q^{i-1}\big(t_i^-\big)M_{t_i^-}\big). \quad (19)$$

Immediately after y_{t_i} is observed, the trading strategy becomes

$$\frac{\mu_0+\mu_1 M_{t_i}-r}{\gamma\sigma_S^2}+\frac{(1-\gamma)(\rho\sigma_X\sigma_S+\mu_1 V(t_i))}{\gamma\sigma_S^2}\big(d^i(t_i)+Q^i(t_i)M_{t_i}\big).$$

These expressions suggest that the discrete change of π at news update times comes from two sources. The first source is the update of the conditional mean of X_{t_i} from $M_{t_i^-}$ to M_{t_i}, which changes the expected return of the stock. Since the expected value of M_{t_i} is equal to $M_{t_i^-}$, on average this update does not affect π. The second source is the update of the conditional variance of X_{t_i} from $V(t_i^-)$ to $V(t_i)$. Since $V(t_i)$ is always smaller than $V(t_i^-)$, the news update decreases the uncertainty about the predictive variable. Note that the difference between (d^{i-1}, Q^{i-1}) and (d^i, Q^i) is also caused by the change in the conditional volatility, as can be seen from equations (B7) and (B8). Extensive numerical results show that this reduction of the conditional variance tends to increase stock investment at the news update times (see Figure 6 in Section V, for example).

Because of the periodic news updates, the investor can periodically correct his estimation error. This periodic correction may have two opposing effects on the optimal trading strategy. On the one hand, the estimation risk that the investor faces with periodic news is smaller than the risk without periodic news. Accordingly, it seems reasonable to conjecture that given the same current conditional distribution, the investor would invest more with periodic news than without it. On the other hand, at the beginning of the next period the investor will receive a news update that will reduce estimation risk, so it also seems reasonable to conjecture that given the same current conditional distribution, with periodic news the investor would reduce his stock investment to decrease the damage from possible estimation error, and would wait to choose a better trading strategy until the next period when he can make a better estimation. Somewhat surprisingly, the following proposition implies that both of these conjectures are incorrect and that the periodic news update affects the trading strategy *only* through its effect on the current conditional distribution of the predictive variable.

PROPOSITION 3: *For any $i = 1, 2, \ldots, N$, at time $t \in [t_i, t_{i+1})$ the solution $(Q^i(t), d^i(t))$ to the case with periodic news updates and the solution $(Q(t), d(t))$ to the case with no periodic news are related as follows:*

$$Q^i(t)=\frac{Q(t)}{1+(1-\gamma)(V(t)-V^i(t))Q(t)}, \quad (20)$$

$$d^i(t)=\frac{d(t)}{1+(1-\gamma)(V(t)-V^i(t))Q(t)}, \quad (21)$$

where $V(t)$ and $V^i(t)$ are the conditional variances of X_t given no periodic news and given periodic news updates, respectively. In particular, if $V(t) = V^i(t)$, then we have $\boldsymbol{Q}^i(t) = \boldsymbol{Q}(t)$ and $d^i(t) = d(t)$.

Proof: See Appendix C.[12]

Substituting (20) and (21) into (18) in Proposition 2 shows that if $M_t = M_t^i$ and $V(t) = V^i(t)$, then the time t optimal trading strategy with periodic news will be the same as the time t optimal trading strategy without periodic news, that is, $\pi_t^{i*} = \pi_t^*$. This result implies that the time t optimal stock investment π_t^{i*} is independent of how frequently or how accurately the investor will observe the predictive variable in the future. Thus, all that matters for the investor's current optimal trading strategy is the current conditional distribution of the predictive variable. This somewhat surprising "myopic" behavior (with respect to the future information structure) for even non-log investors stems from a fairly simple fact: While access to more accurate or more frequent news in the future does improve the investor's utility, it improves the utility by a constant multiple across all states through $c^i(t)$ and thus does not affect the current optimal trading strategy.

We emphasize that the optimal trading strategies with or without periodic news are the same only when the conditional means and variances are the same across the two models. At times $s \geq t$ the optimal policy π_s^{i*} in the model with periodic information will almost surely differ from the optimal policy π_s^* in the model without periodic information, because periodic information will almost surely change the conditional distribution. Therefore, the trajectories of the optimal trading strategies will be almost surely different across the two models.

From the perspective of computation, Proposition 3 suggests an efficient way of computing the trading strategy with periodic news updates. Specifically, to calculate the optimal trading strategy π_t^{i*} at any time t, we only need to set $M_t = M_t^i$ and $V(t) = V^i(t)$ in equation (14) and then we have $\pi_t^{i*} = \pi_t^*$. This approach reduces N iterations to a one-period computation.

III. An Empirical Calibration

To understand how the information cost affects the optimal news frequency, news accuracies, and trading strategy, we apply the theoretical results derived in Section II to study an empirical model. In this study, we focus primarily on periodic news because one of our main objectives is to examine the determination of the optimal news frequency. In addition, periodic news is usually more accurate than continuous news and is therefore typically a more important

[12] This proposition also holds for a more general model that allows for multiple stocks, multiple predictive variables, and time-varying coefficients (see Appendix A). The proof for this more general case is available from the authors.

information source for an investor. We provide an analysis of optimal continuous news accuracy in Sections VI and VII.

One of the popular predictive variables used in the predictability literature is dividend yield. For example, Fama and French (1988), Hodrick (1992), and Xia (2001) use the dividend yield as the predictive variable in their model calibration. However, as Lettau and Ludvigson (2001) demonstrate, fluctuations in the aggregate consumption–wealth ratio have stronger predictive power and generate better future return forecasts than the dividend yield.[13] The economic intuition is that investors who want to smooth consumption adjust their current consumption if they expect transitory movements in their financial wealth caused by the variation in expected returns. When the expected return rises, a forward-looking investor increases his current consumption, and when the expected return declines, he decreases it.

Consistent with Lettau and Ludvigson (2001), we assume in the following analysis that the consumption–wealth ratio predicts asset returns. We first estimate the joint stochastic process for the stock return and the consumption–wealth ratio. Then we show how the optimal news frequency, news accuracies, and trading strategy are affected by the fundamental parameters such as risk aversion, investment horizon, and information cost.

Specifically, we estimate the model described by equations (1) and (2), where the stock return is the quarterly CRSP value-weighted return of stocks traded on the New York Stock Exchange (NYSE) and X_t represents the estimated trend deviation variable, cay_t, for the consumption-wealth ratio as used by Lettau and Ludvigson (2001) to predict the stock return.[14] The parameters $\mu_0, \mu_1, g_0, g_1, \sigma_S, \sigma_X$, and ρ are all constants to be estimated.

The estimation period is from 1952 to 2001. We select this period because of the availability of the predictive variable cay_t. The risk-free interest rate is the mean of the continuously compounded real quarterly T-bill return, which is equal to 0.0034.[15] Since only quarterly data of cay_t are available, following Duffie, Pan, and Singleton (2000) and Singleton (2001) we derive the characteristic function and use the maximum likelihood method to estimate the parameter values (see Appendix D for details).[16] Using this procedure, we obtain the following estimates:

[13] There have been some arguments against the estimation approach used by Lettau and Ludvigson (2001) (see Brennan and Xia (2002) and (Hahn and Lee (2001)). But Inoue and Kilian (2004) suggest that this approach may be the correct one to improve test power.

[14] Implicitly, we assume that *historical quarterly* consumption–wealth ratio time-series data are precisely observable to econometricians and the resulting estimates are available to investors. We abstract away from calibrating the noise in the historical quarterly data. Quarterly data of cay_t can be downloaded from Martin Lettau's web page at New York University.

[15] Our estimate of the real risk-free interest rate is similar in magnitude to those used in the literature (e.g., Campbell, Lo, and MacKinlay (1997)). We also use higher real interest rates, like those in Xia (2001), to check the sensitivity of our results to our interest rate estimate and find that our main results are robust to these changes.

[16] We also use *monthly* stock return data and quarterly cay_t data to estimate the model. Specifically, we maximize the likelihood function derived after taking into account the monthly Bayesian updates on the predictive variable from the monthly observation of the stock return. The qualitative results remain the same.

$$g_0 = 0.117, \quad g_1 = -0.180, \quad \mu_0 = -1.301, \quad \mu_1 = 2.040,$$
$$\sigma_S = 0.0801, \quad \sigma_X = 0.00747, \text{ and } \rho = -0.620.$$

For the numerical analyses that we conduct in subsequent sections, we set the default risk-aversion coefficient, γ, to 5 (as estimated, e.g., by Xia (2001)), $\lambda = 1$, $\eta = 0.01$, and the initial value of the predictive variable X_0 equal to its sample mean 0.649.

We assume that the prior variance $V(0^-)$ is equal to the steady-state conditional variance $\bar{V} = 0.000147$ and the prior mean M_{0^-} is equal to the initial value X_0. In addition, as a default setting we assume that each news update is unbiased and thus the conditional mean M_0 after the initial news release is also equal to X_0.[17] The investor's initial wealth W_{0^-} is normalized to be equal to one.

In Sections IV and V we focus on the analysis of the optimal news frequency. To reflect that continuous news is typically less accurate and more costly, we assume that the news accuracies are exogenously fixed at $\alpha_\varepsilon = 100$ and $\alpha_\nu = 10$ and that the information costs are fixed at $\beta_i(\alpha_\varepsilon) = \beta = 0.05\%$ (five basis points) for every periodic news update and $\beta_c(\alpha_\nu) = \beta_c = 0.1\%$ for the continuous news.[18] In Sections VI and VII we study the case in which the optimal frequency and the optimal accuracies are all endogenously determined.

IV. Analysis of the Optimal News Frequency

In this section, we use the model calibrated above to examine how various fundamental parameters affect the optimal news frequency.

A. Risk Aversion

Figure 1 plots the optimal news update frequency (measured in quarters by the time τ between news updates) as a function of the risk-aversion coefficient γ. This figure shows that as an investor becomes more risk averse, he chooses less frequent news updates. This is consistent with the standard finding that a more risk-averse investor invests less in stock and therefore benefits less from news updates. The figure also illustrates the intuitive result that as the information cost β increases, the optimal news frequency decreases. For example, an investor with a risk-aversion coefficient of five chooses $\tau^* = 0.32$ (about one update per month) when the cost is five basis points, and he chooses $\tau^* = 0.53$ (about two updates per quarter) when the cost goes up to 10 basis points. The wedge between the two lines illustrates that an increase in the information cost induces a larger decrease in the optimal frequency for a more risk-averse investor.

[17] Implicitly, we assume that a long history of data on the stock price and the predictive variable is available and therefore the steady-state conditional variance $\bar{V}$ can be estimated and used as the initial conditional variance.

[18] Varying these parameter values does not change qualitative results.

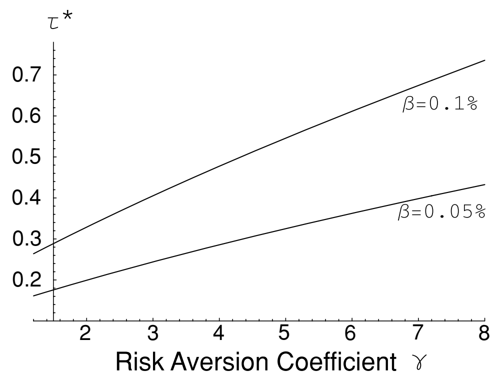

Figure 1. The optimal news frequency against risk aversion. Parameter values: $T = 20, \mu_0 = -1.301, \mu_1 = 2.040, g_0 = 0.117, g_1 = -0.180, \sigma_S = 0.0801, \sigma_X = 0.00747, \rho = -0.620, r = 0.00340, X_0 = 0.649, M_0 = 0.649, V(0^-) = 0.000147, \beta_c = 0.1\%, \alpha_\varepsilon = 100, \alpha_\nu = 10, \eta = 0.01$, and $\lambda = 1$.

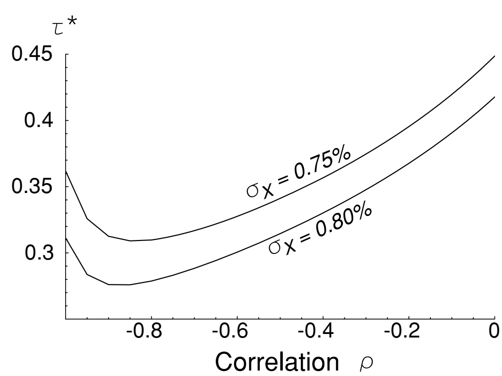

Figure 2. The optimal news frequency against correlation. Parameter values: $\gamma = 5, T = 20, \mu_0 = -1.301, \mu_1 = 2.040, g_0 = 0.117, g_1 = -0.180, \sigma_S = 0.0801, r = 0.00340, X_0 = 0.649, M_0 = 0.649, V(0^-) = 0.000147, \beta = 0.05\%, \beta_c = 0.1\%, \alpha_\varepsilon = 100, \alpha_\nu = 10, \eta = 0.01$, and $\lambda = 1$.

B. Correlation

Figure 2 plots the optimal news update frequency as a function of the correlation ρ between the stock return and the predictive variable. This figure shows that the optimal news frequency is a nonmonotonic function of the correlation. Intuitively, when the stock return and the predictive variable are highly (negatively) correlated, the investor can infer information about the predictive variable from observing the stock return and therefore does not need very frequent news updates. As the correlation decreases, the stock return becomes less informative and the investor chooses more frequent news updates to reduce the estimation risk. However, when the correlation becomes too low, learning is not effective and the benefit from extra information obtained from more frequent news updates becomes smaller. Accordingly, the investor starts to decrease the news frequency to reduce information costs.

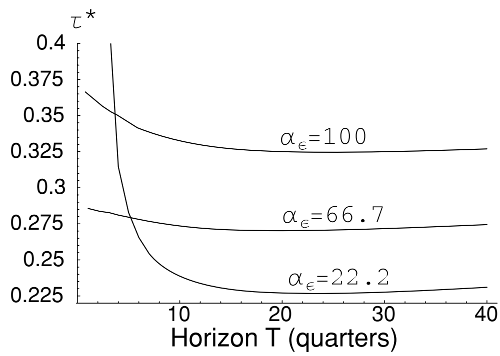

Figure 3. The optimal news frequency against horizon. Parameter values: $\gamma = 5, \mu_0 = -1.301, \mu_1 = 2.040, g_0 = 0.117, g_1 = -0.180, \sigma_S = 0.0801, \sigma_X = 0.00747, \rho = -0.620, r = 0.00340, X_0 = 0.649, M_0 = 0.649, V(0^-) = 0.000147, \beta = 0.05\%, \beta_c = 0.1\%, \alpha_\nu = 10, \eta = 0.01$, and $\lambda = 1$.

This figure also shows that as the unconditional volatility of the predictive variable increases, the investor increases the news update frequency to reduce the estimation risk.

C. Investment Horizon

Figure 3 plots the optimal news update frequency as a function of the investment horizon T. This figure shows that the optimal news frequency is also nonmonotonic in the investment horizon. When the investment horizon is short, the investor cannot gain much from news updates, so he only needs low-frequency news updates. An increase in the investment horizon has two opposing effects. On the one hand, as the investment horizon increases, the gain from more frequent news updates also increases. This effect tends to increase the optimal news frequency. On the other hand, as the investment horizon increases, for a given frequency the total information cost increases because the number of news updates becomes larger. This effect tends to decrease the optimal news frequency. The figure shows that although the first effect dominates when the horizon is short, the second effect gradually takes over as the investment horizon becomes longer. The interaction of the two opposing effects produces the nonlinear patterns in the figure.

Figure 3 also shows that for a short investment horizon the optimal news frequency is nonmonotonic in news accuracy.[19] Specifically, as the news becomes less accurate, the optimal news frequency first increases and then decreases. The underlying intuition is similar to that for the nonmonotonic pattern in the correlation ρ displayed in Figure 2. When news is accurate, the investor does not need very frequent news updates. As news becomes less reliable, more frequent news is important to reduce estimation risk. However, when news becomes too

[19] The nonmonotonicity of the optimal news frequency in the news accuracy actually holds for all investment horizons, with the critical news accuracy level (at which the slope is zero) decreasing in the horizon.

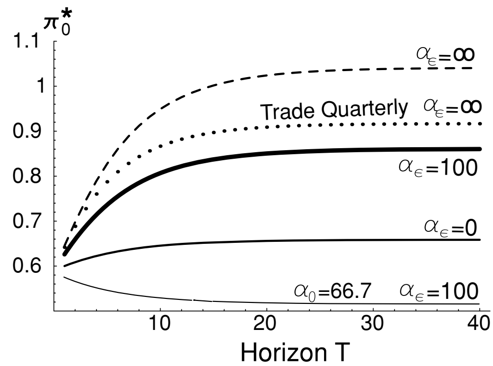

Figure 4. The fraction of wealth in stock against horizon. Parameter values: $\gamma = 5, \mu_0 = -1.301, \mu_1 = 2.040, g_0 = 0.117, g_1 = -0.180, \sigma_S = 0.0801, \sigma_X = 0.00747, \rho = -0.620, r = 0.00340, X_0 = 0.649, M_{0^-} = 0.649, V(0^-) = 0.000147, \beta = 0.05\%, \beta_c = 0.1\%, \alpha_\nu = 10, \eta = 0.01$, and $\lambda = 1$.

noisy, additional news updates increase costs without much benefit, and thus the optimal news frequency decreases.

V. Analysis of the Optimal Trading Strategy

In this section, we use the calibrated model to examine how rational inattention affects the optimal trading strategy.

A. Horizon

As we discuss above, Propositions 2 and 3 imply that the optimal investment in the stock at time t only depends on the conditional distribution $(M_t, V(t))$, but not on the *future* news frequency or accuracy. Figure 4 plots the initial fraction of wealth (π_0^*) in the stock against horizon T.[20] Starting with the same prior $(M_{0^-}, V(0^-))$ (except for the lowest curve to be discussed later), the figure illustrates the optimal stock investments immediately after a time 0 news update with accuracy σ_ε. The figure shows that uncertainty about the predictive variable significantly reduces investment in the stock across all horizons and the reduction increases with the horizon. When the horizon is 1 year, the difference in investments between the case with $\sigma_\varepsilon = 0$ and the case with $\sigma_\varepsilon = 100$ is about 10%. As the horizon increases to 5 years, this difference increases to about 20%.

This figure also shows that the optimal fraction of wealth in the stock increases with the horizon, consistent with the typical life-cycle investment advice from a financial advisor. However, the life-cycle pattern can be reversed if some of the default parameter values are different. As an example, after we

[20] As we mention before, to simplify exposition we simply use "wealth" to refer to the "effective wealth" defined immediately after Proposition 1.

decrease the initial news accuracy $\alpha_0 \equiv 1/\sqrt{V(0^-)}$ from 82.5 to 66.7, the optimal fraction of wealth invested in the stock *decreases* with the investment horizon.[21] Intuitively, there are two offsetting effects at work. On the one hand, because of the negative correlation between the stock return and the predictive variable, the hedging benefit increases as the horizon increases; the stock investment also increases with the horizon. We refer to this effect as the *hedging effect*. On the other hand, as the noise in observation increases, the investor becomes less certain about the expected return of the stock and on that ground reduces his stock investment. As the horizon increases, the impact of this uncertainty grows, therefore the reduction in the stock investment is greater for an investor with a longer horizon. We refer to this effect as the *uncertainty effect*. Depending on the relative magnitudes of these two opposing effects, the stock investment can increase or decrease as the horizon increases. In the presence of parameter uncertainty, Xia (2001) finds that the horizon pattern of the optimal stock investment depends on the current value of the continuously observable predictive variable. In contrast, we show that the magnitude of the uncertainty about a predictive variable can reverse the horizon effect of return predictability. The reverse pattern illustrated in this figure suggests that the horizon effect is sensitive to the existence of information costs and news inaccuracy.

Given that news is only received at discrete times in the case with periodic news updates, an alternative strategy is to trade only at the news update times, as in a discrete-time model. Clearly, this is a suboptimal strategy because it imposes an exogenous constraint on the trading frequency. What would be the effect of this discrete trading constraint on the trading strategy? Would it cause an investor to invest more or less? To address these questions, in Figure 4 we also plot the optimal trading strategy of an investor who can only trade at quarterly news update times, assuming quarterly news is accurate. This figure shows that trading only at news update times would result in significant underinvestment for an investor with a long horizon. For example, the investor with a 5-year horizon who can only trade at news update times would invest about 12% less in the stock than an investor who can trade continuously. When the investor is restricted to trading only at news update times, and thus cannot update his position as new information from the stock price comes in, he is more cautious and invests significantly less in the stock.[22]

B. Estimation Error

Figure 5 plots the initial trading strategy difference ($\Delta\pi_0^*$) between an investor who receives a news update at time 0 and an investor who perfectly

[21] This result suggests that if the conditional volatility is stochastic, as in a non-Gaussian setting, the life-cycle pattern may be reversed when uncertainty is high. If the correlation ρ between the stock return and the predictive variable is positive, this life-cycle pattern may also be reversed.

[22] We also find that the welfare loss (in terms of certainty-equivalent wealth) of the investor from adopting this suboptimal strategy can be very significant (as high as 8% of the initial wealth for an investor with a 5-year horizon and a risk-aversion coefficient of 5). We omit the figure on this result to save space.

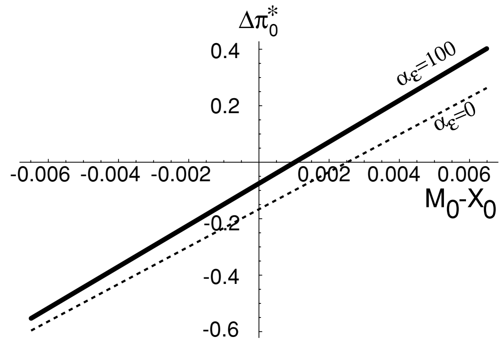

Figure 5. The fraction of wealth in stock against initial estimation error. Parameter values: $\gamma = 5, T = 20, \mu_0 = -1.301, \mu_1 = 2.040, g_0 = 0.117, g_1 = -0.180, \sigma_S = 0.0801, \sigma_X = 0.00747, \rho = -0.620, r = 0.00340, X_0 = 0.649, V(0^-) = 0.000147, \beta = 0.05\%, \beta_c = 0.1\%, \alpha_\nu = 10, \eta = 0.01$, and $\lambda = 1$.

observes the predictive variable against the initial estimation error $(M_0 - X_0)$ for two news accuracy levels: $\alpha_\varepsilon = 100$ and $\alpha_\varepsilon = 0$. This figure demonstrates the impact of estimation error on the optimal trading strategy and shows how it changes with uncertainty about the estimation. When a line is above (below) zero, the investor overinvests (underinvests) compared with the optimal investment under perfect observation. In general, the figure shows that when the investor underestimates the predictive variable, he underinvests, and when he overestimates, he overinvests. However, there is a region in which $M_0 - X_0 > 0$, but the $\alpha_\varepsilon = 100$ and $\alpha_\varepsilon = 0$ curves are below the horizontal axis. This suggests that if news is not very accurate (e.g., $\alpha_\varepsilon = 100$), then the investor might still underinvest even when he overestimates. Intuitively, news inaccuracy discounts the reliability of the predictive variable estimate and lowers the investor's investment in the stock.

C. Portfolio Revision at News Update Times

With periodic news updates, the investor updates his estimate of the predictive variable and discretely adjusts his portfolio according to the discrete changes in the conditional mean and conditional variance. Since the conditional expectation immediately before a news update is an unbiased estimate of the conditional expectation after the news update ($E(M_{t_i}) = M_{t_i^-}$), on average the change in the portfolio weights caused by the difference in the conditional means before and after an observation is equal to zero. However, the conditional variance also has an impact on the portfolio weights. Whenever a news update is received, the updated conditional variance decreases immediately. The reduction in uncertainty about the stock return makes the stock more attractive, and thus the investor increases his stock investment. Figure 6, which plots the expected percentage increase in stockholdings ($E[\frac{\pi^*_{\tau^*} - \pi^*_{\tau^*-}}{\pi^*_{\tau^*-}}] \times 100$) at

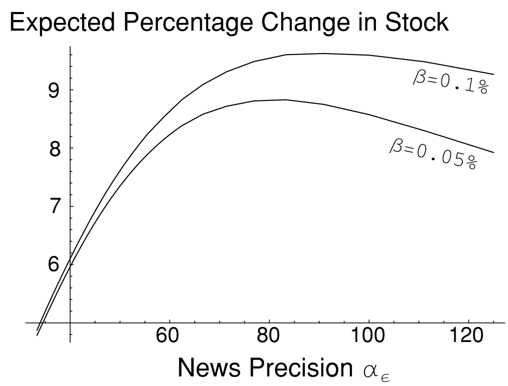

Figure 6. The expected percentage change in stock fraction against news accuracy. Parameter values: $\gamma = 5, T = 20, \mu_0 = -1.301, \mu_1 = 2.040, g_0 = 0.117, g_1 = -0.180, \sigma_S = 0.0801, \sigma_X = 0.00747, \rho = -0.620, r = 0.00340, X_0 = 0.649, M_{0^-} = 0.649, V(0^-) = 0.000147, \beta_c = 0.1\%, \alpha_\nu = 10, \eta = 0.01$, and $\lambda = 1$.

the first news update time for an investor with a risk-aversion coefficient of 5 and an investment horizon of 5 years, confirms this intuition. For example, when $\alpha_\varepsilon = 100$, after the first news update on average the investor invests an additional 8.5% of wealth in the stock if $\beta = 0.05\%$ and an additional 9.5% of wealth in the stock if $\beta = 0.1\%$. The figure also shows that the expected percentage change in the stockholding first increases and then decreases with news accuracy α_ε. A decrease in the news accuracy α_ε has two opposing effects. On the one hand, a noisier news update results in larger conditional volatility just before the next news update, which implies a greater reduction in the conditional volatility when the next news arrives with the same accuracy. On the other hand, for a fixed level of conditional volatility noisier news is less helpful in reducing volatility. Figure 6 shows that when α_ε is large, the first effect dominates and the expected adjustment increases as α_ε decreases. When α_ε is small, however, the second effect dominates and the expected adjustment decreases as α_ε decreases.

In addition, an increase in the information cost β increases the percentage change in the stock investment at the optimal news update time and the resulting increase is greater for more accurate news. Intuitively, as the information cost increases, the optimal news frequency decreases, and thus a news update causes a greater reduction in the conditional volatility.

VI. Endogenous News Accuracies

In Sections IV and V we assume that an investor can only choose the news frequency but not the news accuracies. However, it is usually the case that one can improve the accuracy of news up to a certain limit by paying a higher information production cost. An interesting question is then: What are the optimal news frequency and accuracies from an investor's point of view given the tradeoffs among frequency, accuracies, and information costs? To shed some light on this question, we now consider the joint choice of optimal news frequency and optimal news accuracies. For the analysis in this section, we assume that the cost function for periodic news, which is the same across all news updates, is

$$\beta(\alpha_\varepsilon) = \left(\frac{0.0006\alpha_\varepsilon}{1 - 0.0025\alpha_\varepsilon}\right)^{16/9}, \tag{22}$$

and the cost function for the continuous news is[23]

$$\beta_c(\alpha_\nu) = \left(\frac{0.0004\alpha_\nu}{1 - 0.1\alpha_\nu}\right)^{16/9}. \tag{23}$$

We numerically solve for the optimal news frequency τ^*, the optimal periodic news accuracy α_ε^*, and the optimal continuous news accuracy α_ν^* that maximize the investor's value function.

Figure 7 plots the optimal frequency, the optimal accuracies, and the optimal trading strategy as functions of the risk-aversion coefficient. The figure shows that as an investor becomes more risk averse, he tends to choose less frequent but more accurate periodic news updates. Intuitively, because a more risk-averse investor invests less in stock, more frequent news is less helpful to him and thus, he chooses less accurate high-frequency news (e.g., continuous news) and more accurate low frequency periodic news.

Figure 8 shows that as investment horizon increases, the investor also chooses less frequent but more accurate periodic news updates plus more accurate continuous news. Recall that when the news accuracy is exogenously given, the optimal news frequency is nonmonotonic in the horizon (see Figure 3). With endogenous news accuracy, the investor has a tradeoff between news frequency and news accuracy. In the case depicted in this figure, it is relatively cheaper to increase accuracy rather than frequency. In addition, as in the case with exogenous news accuracy, the optimal fraction of wealth invested in stock also increases with horizon. Since the investor chooses more accurate news as the horizon increases, this life-cycle pattern is strengthened when news accuracies are endogenized.

VII. Extensions

To verify the robustness of the main results we obtain above, in this section we generalize our model to allow for intertemporal consumption, nonlinear dynamics, and continuous choice of information accuracy. Specifically, we assume

[23] We present results based on these cost functions because they are increasing and convex, imply an upper bound for possible news accuracies (i.e., $\alpha_\varepsilon < 400$ and $\alpha_\nu < 10$), and yield an interior solution. The difference between β and β_c is chosen to capture the typical feature that the upper bound for the continuous news accuracy is lower than that for periodic news accuracy and the cost for continuous news is higher than that for periodic news. We employ other functional forms, and for those functional forms that yield interior solutions, we obtain the same qualitative results as those reported in this section.

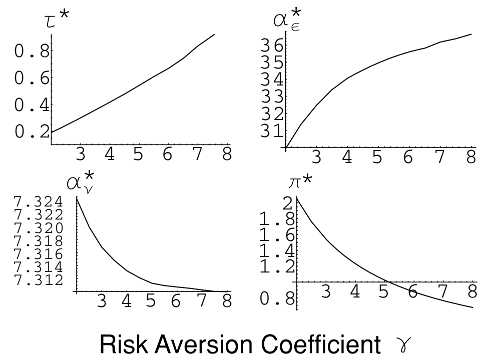

Figure 7. Optimal news frequency, accuracies, and trading strategy against the risk-aversion coefficient. Parameter values: $T = 20, \mu_0 = -1.301, \mu_1 = 2.040, g_0 = 0.117, g_1 = -0.180, \sigma_S = 0.0801, \sigma_X = 0.00747, \rho = -0.620, r = 0.00340, X_0 = 0.649, M_{0^-} = 0.649, V(0^-) = 0.000147, \alpha_\nu = 10, \eta = 0.01$, and $\lambda = 1$.

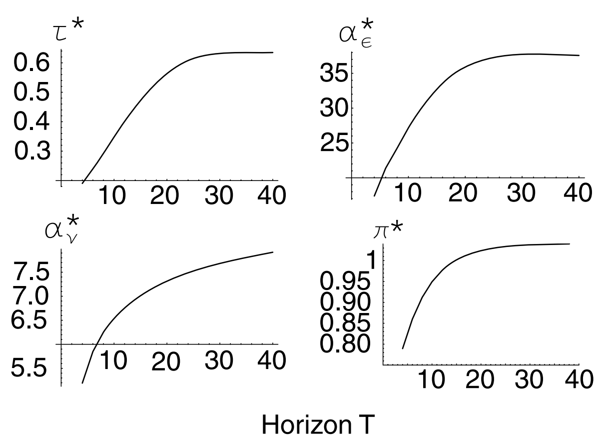

Figure 8. News frequency, accuracies, and trading strategy against horizon. Parameter values: $\gamma = 5, \mu_0 = -1.301, \mu_1 = 2.040, g_0 = 0.117, g_1 = -0.180, \sigma_S = 0.0801, \sigma_X = 0.00747, \rho = -0.620, r = 0.00340, X_0 = 0.649, M_{0^-} = 0.649, V(0^-) = 0.000147, \alpha_\nu = 10, \eta = 0.01$, and $\lambda = 1$.

that an investor derives utility from both intertemporal consumption c_t and terminal wealth W_T. For the continuous news ν, the investor can continuously choose any accuracy $\alpha_{\nu t}$ by paying a cost at the rate of $\beta_c(\alpha_{\nu_t})$; for the periodic news, the investor can choose any accuracy α_ε by paying at time 0 a cost of $\beta_i(\alpha_\varepsilon)$ for the ith observation. The investor's problem is to choose the number of news updates N, the periodic news accuracy α_ε, the continuous news accuracy $\alpha_{\nu t} \in \mathcal{F}_t$, the consumption $c_t \in \mathcal{F}_t$, and the trading strategy $\theta_t \in \mathcal{F}_t$ to maximize the expected utility from consumption and terminal wealth, that is,

$$\max_{N,\alpha_\varepsilon,\sigma_\nu,c,\theta} \boldsymbol{E}\left[(1-\omega)\int_0^T e^{-\delta t}u_1(c_t)\,dt + \omega e^{-\delta T}u_2(W_T)\right],$$

subject to the budget constraint,

$$dW_t = rW_t\,dt + \theta_t(\mu_0 + \mu_1 X_t - r)\,dt + \theta_t\sigma_S\,dZ_{1t} - c_t\,dt - \beta_c(\alpha_{\nu t})\,dt, \quad (24)$$

the nonlinear dynamics of X_t,

$$dX_t = \left(g_0 + g_1 X_t + g_2 X_t^2\right)dt + \rho\sigma_X\,dZ_{1t} + \sqrt{1-\rho^2}\sigma_X\,dZ_{2t}, \quad (25)$$

the periodic news equation (3), the continuous news equation (4), and the initial wealth after deducting the information cost for the periodic news,

$$W_0 = W_{0^-} - \sum_{i=1}^{N}\beta_i(\alpha_\varepsilon), \quad (26)$$

where $\delta > 0$ is the time discount factor, $\omega \in [0, 1]$ is the weight of the utility from the terminal wealth, g_2 is the nonlinearity coefficient, and

$$u_i(x) = \frac{\gamma_i}{1-\gamma_i}\left(\frac{\lambda_i x}{\gamma_i} + \eta_i\right)^{1-\gamma_i}, \quad i \in \{1, 2\}.$$

Unfortunately, given the complexity of this generalized problem, it seems very difficult, if not impossible, to obtain an explicit solution. In addition, with the HJB partial differential equation being four-dimensional (with state variables (W, M, V, t)) and highly nonlinear, even solving the HJB numerically becomes quite challenging. Accordingly, we numerically solve a discretized version of the investor's problem using the projection method as described in Judd (1998). Specifically, we divide T into n time intervals with length h. The investor's objective function then becomes

$$\max_{N,\alpha_\varepsilon,\alpha_{\nu j},c_j,\theta_j} E\left[(1-\omega)\sum_{j=0}^{n} e^{-\delta j h}u_1(c_j)h + \omega e^{-\delta n h}u_2(W_n)\right],$$

and the wealth dynamics become

$$W_{j+1} = W_j + (rW_j + \theta_j(\mu_0 + \mu_1 M_j - r) - c_j - \beta_c(\alpha_{\nu j}))h + \theta_j\sigma_S\sqrt{h}z_1,$$

where z_1 is a standard normal random variable.

Discretizing the continuous-time filtering equations derived by Basin (2003), we have the conditional mean dynamics[24]

$$M_{j+1} = M_j + \left(g_0 + g_1 M_j + g_2 M_j^2\right)h + \left(\frac{\mu_1}{\sigma_S}V_j + \rho\sigma_X\right)\sqrt{h}z_1 + \alpha_{\nu j}V_j\sqrt{h}z_3,$$

[24] Given the nonlinear dynamics of the predictive variable X_t, it does not seem feasible to derive the exact filter for the discretized model. The commonly used extended Kalman filter for nonlinear filtering requires another layer of approximation (in addition to time discretization) by linearizing the nonlinear dynamics (see Chen (1993)). In comparison, direct discretization of the continuous-time filter equations has only one layer of approximation (i.e., time discretization) and allows for the incorporation of continuous information. This approach has also been widely used in the existing literature (e.g., Clark (1978), James, Krishnamurthy, and Le Gland (1996)).

and the conditional variance dynamics

$$V_{j+1} = V_j + \left(2g_1 V_j + 4g_2 M_j V_j + \sigma_X^2 - \left(\frac{\mu_1}{\sigma_S} V_j + \rho\sigma_X\right)^2 - \alpha_{\nu j}^2 V_j^2\right) h,$$

where z_3 is a standard normal random variable independent of z_1.

Note that with nonlinear dynamics for the predictive variable, the conditional variance is no longer deterministic. Instead, it depends on the conditional mean M and consequently is stochastic over time. For a given number of news updates N and a given accuracy α_ε, we then solve this discrete-time problem iteratively from period $n-1$ backward, taking into account the periodic news y as before. Finally, we maximize over N and α_ε to obtain the solution.

To examine the robustness of the main qualitative results shown in the previous sections, we first reestimate the nonlinear system of equations (1) and (25) for the parameter values of $\mu_0, \mu_1, \sigma_S, g_0, g_1, g_2, \sigma_X$, and ρ using the methodology derived in Shoji and Ozaki (1998). We then numerically compute the optimal frequency, optimal trading strategy, optimal consumption, and/or optimal accuracies for cases corresponding to Figures 1 to 8 and report a subset of these results in Table I.[25]

Table I shows that the main qualitative results remain the same in the presence of intertemporal consumption, nonlinear dynamics, and continuous news accuracy choice. Specifically, in Panels A and B, we solve for the optimal periodic news frequency τ^*, the optimal trading strategy π^*, and the optimal consumption c^* for a given constant periodic news cost β, a continuous news accuracy α_ν, and a continuous news cost β_c, similar to Figures 1 to 6. Comparison of row 1 with rows 2 and 3 of Panel A shows that, consistent with Figure 1, the optimal periodic news frequency decreases as risk aversion or news costs increase. Consistent with Figure 2 (comparing row 1 to rows 4 to 6), the optimal frequency is increasing in σ_X and nonmonotonic in the correlation coefficient ρ because of the change in the informativeness of the stock return. The results shown in rows 1, 7, and 8 suggest that, similar to Figure 4, the optimal fraction of wealth (π^*) still increases with horizon and news accuracy. Rows 9 and 10 suggest that overestimation (underestimation) tends to cause overinvestment (underinvestment), the same pattern as that illustrated in Figure 5. Rows 12 and 13 report results for the cases with exogenously given quarterly periodic news frequency and the same initial news accuracy $\alpha_{\varepsilon 0} = 100$ but different news accuracy for future periodic updates. Comparing rows 7, 12, and 13 suggests that, consistent with Proposition 3, the optimal trading strategy is not sensitive to future news frequency or news accuracy as long as the current news accuracy is the same.

In Panel B, we report the expected percentage changes ($E[\frac{\Delta\pi^*_{\tau^*}}{\pi^*_{\tau^*-}}] \times 100$) in the fraction of wealth in stock at the first news update time after time 0 for cases with different news accuracies and costs. Consistent with Figure 6, the expected percentage change is nonmonotonic in news accuracy and increases as the news costs increase.

[25] More extensive numerical results confirm the same findings as reported in Table I.

Table I

Optimal Periodic News Frequency, Accuracies, Trading Strategy, and Consumption

The table shows the optimal frequency (τ^*), the optimal fraction of wealth in stock π^*, and the optimal consumption (c^*). Panel B shows π^* just before and immediately after the second periodic news update time τ^* and the percentage changes. Panel C also shows the optimal news accuracies. For Panels A and B, the base case parameter values are $\gamma_1 = \gamma_2 = 5, \lambda_1 = \lambda_2 = 1, \eta_1 = \eta_2 = 0.01, T = 20, \mu_0 = -1.275, \mu_1 = 1.998, g_0 = -0.724, g_1 = 2.423, g_2 = -2.012, \sigma_S = 0.0802, \sigma_X = 0.00748, \rho = -0.617, r = 0.00340, \beta = 0.03\%, \beta_c = 3 \times 10^{-6}, \alpha_\varepsilon = 100, \alpha_\nu = 10, \delta = 0.01, \omega = 0.5, X_0 = 0.649, M_{0-} = 0.649,$ and $V(0^-) = 0.000139$. For Panel C, $\beta, \alpha_\varepsilon$, and α_ν are all endogenous with $\beta(\alpha_\varepsilon) = (0.0001\alpha_\varepsilon/(1 - 0.001\alpha_\varepsilon))^{16/9}$ and $\beta_c(\alpha_\nu) = (0.000004\alpha_\nu/(1 - 0.1\alpha_\nu))^{16/9}$.[a]

Panel A: Optimal News Frequency, Trading Strategy, and Consumption

Row	Parameter Value	τ^*(Quarters)	π^*	c^*
1	Base case	0.2	0.76	0.030
2	$\gamma_2 = 6$	1.6	0.61	0.020
3	$\beta = 0.035\%$	1.6	0.72	0.030
4	$\sigma_X = 0.70\%$	1.6	0.71	0.030
5	$\rho = -0.95$	1.2	1.06	0.036
6	$\rho = -0.55$	1.6	0.69	0.030
7	$T = 15$	1.9	0.71	0.045
8	$\alpha_\varepsilon = 90$	1.6	0.71	0.030
9	$m_0 = 0.650$	0.2	0.83	0.030
10	$m_0 = 0.648$	0.2	0.69	0.029
11	$\eta_1 = \eta_2 = 0.05$	1.6	0.71	0.011
12	$T = 15, \alpha_{\varepsilon t} = 100, t > 0$ (quarterly news)	NA	0.71	0.044
13	$T = 15, \alpha_{\varepsilon t} = 90, t > 0$ (quarterly news)	NA	0.71	0.044

Panel B: Impact of News on Stock Investment

Row	Parameter Value	$\pi^*_{\tau^*-}$	$\pi^*_{\tau^*}$	$\Delta\pi^*_{\tau^*}/\pi^*_{\tau^*-} \times 100$
14	Base case	0.76	0.79	4.9
15	$\alpha_\varepsilon = 70$	0.72	0.76	5.4
16	$\alpha_\varepsilon = 50$	0.69	0.73	4.9
17	$\beta = 0.035\%, \alpha_\varepsilon = 100$	0.63	0.73	16.0

Panel C: Optimal Frequency, Accuracies, Trading Strategy, and Consumption

Row	Parameter Value	τ^*(Quarters)	π^*	α^*_ν	α^*_ε	c^*
18	Base case	0.6	0.72	9.39	95.90	0.030
19	$\gamma_2 = 6$	1.0	0.60	9.37	96.00	0.020
20	$T = 15$	1.1	0.70	9.38	95.80	0.044

[a]To facilitate comparison we choose the same news cost functional forms as before. Some parameters are different because in this generalized model the investor pays the continuous news cost continuously. We also conduct the same analysis for other functional forms and obtain the same qualitative results.

In Panel C, we also endogenize the news frequency and news accuracies for both the periodic and the continuous news. Rows 18 to 20 suggest that the optimal periodic news frequency decreases with risk aversion and horizon. This is consistent with Figure 7 but different from Figure 8. The difference from Figure 8 confirms the nonlinearity of news frequency in investment horizon as shown in Figure 3. The tradeoff between news accuracy and frequency in this example makes it cheaper to decrease news frequency for a short horizon, as in Figure 3. Also, the optimal periodic news accuracy increases with both risk aversion and horizon. In contrast, the optimal continuous news accuracy decreases with risk aversion but increases with investment horizon. Panel C also implies that the optimal consumption decreases with risk aversion and horizon. Intuitively, as the investor gets more risk averse, he saves more to guard against bad shocks in the market, and as the consumption horizon increases, the investor needs to spread consumption across more years and therefore the consumption level declines.

VIII. Conclusions

In this paper, we show that because of information production and processing costs, inattention to important economic news that affects investment performance may be rational. Rational inattention significantly changes the optimal trading strategy and the investor may over- or underinvest. Optimal news frequency (attention frequency) displays nonmonotonic patterns in news accuracy and investment horizon. We also find that the optimal trading strategy is myopic with respect to news frequency and news accuracy even for an investor with non-log preferences. Finally, an investor with a higher risk aversion or a longer investment horizon chooses less frequent but more accurate periodic news updates.

Appendix A: A More General Model with Multiple Risky Assets, Multiple Predictive Variables, and Time-Varying Parameters

In this appendix, we extend our analysis to a more general model with multiple risky assets, multiple predictive variables, and time-varying parameters.

Let there be $n + 1$ assets that are continuously traded in the market. We assume $n < m$ so that there are no redundant assets. The first asset is a money market account that is locally risk free. The other n assets ("stocks") are risky. There are k predictive variables that predict the expected returns of these assets. The first k_1 predictive variables X_t^u are only observable (possibly with some error) at discrete times $t_1(0), t_2, \ldots, t_N \in [0, T]$. The remaining $k_2 = k - k_1$ predictive variables X_t^o are continuously and accurately observable. Let B_t denote the price of the locally riskless asset and S_t be the cum-dividend stock price vector at time t. We assume that B_t evolves as follows:

$$\frac{dB_t}{B_t} = \left(r_0(t) + r_2(t)X_t^o\right) dt,$$

where $r_0(t)(1 \times 1)$ and $r_2(t)(1 \times k_2)$ are deterministic functions of time. In addition, the prices of the risky assets satisfy

$$\frac{dS_t}{S_t} = \left(\mu_0(t) + \mu_1(t)X_t^u + \mu_2(t)X_t^o\right)dt + \sigma_S(t)\,dZ_t,$$

where $\mu_0(t)(n \times 1)$, $\mu_1(t)(n \times k_1)$, $\mu_2(t)(n \times k_2)$, and $\sigma_S(t)(n \times m)$ are deterministic functions of time t and dS_t/S_t represents element-by-element division. For the dynamics of the predictive variables, we assume that

$$dX_t^u = \left(g_0(t) + g_1(t)X_t^u + g_2(t)X_t^o\right)dt + \sigma_u(t)\,dZ_t$$

and

$$dX_t^o = \left(h_0(t) + h_1(t)X_t^u + h_2(t)X_t^o\right)dt + \sigma_o(t)\,dZ_t,$$

where $g_0(t)(k_1 \times 1), g_1(t)(k_1 \times k_1), g_2(t)(k_1 \times k_2), \sigma_u(t)(k_1 \times m), h_0(t)(k_2 \times 1)$, $h_1(t)(k_2 \times k_1)$, $h_2(t)(k_2 \times k_2)$, and $\sigma_o(t)(k_2 \times m)$ are all deterministic functions of time t. To save notation, henceforth we suppress argument t wherever confusion is unlikely.

Assume that at any observation time t_i $(i = 1, 2, \ldots, N)$, a noisy signal on X^u reveals that $X^u(t_i)$ is normally distributed with mean M_{t_i} ("initial mean") and deterministic variance v_{t_i} ("initial variance").

Suppose that $\sigma\sigma^\top$ is invertible so that there are no redundant assets. Let M_t be the time $t \in [t_i, t_{i+1})$ expectation of X_t^u conditional on observing S_t and X_t^o (so $M(t_i) = m_{t_i}$). Define

$$\sigma(t) = \begin{bmatrix} \sigma_S \\ \sigma_o \end{bmatrix}$$

and

$$d\hat{Z}_t = (\sigma\sigma^\top)^{-1/2}\left(\begin{bmatrix} \mu_1 \\ h_1 \end{bmatrix}\left(X_t^u - M_t\right)dt + \sigma\,dZ_t\right),$$

where $\hat{Z}_t$ is the (observable) innovation process. By Theorem 10.5.1 of Kallianpur (1980), M_t then evolves according to

$$dM_t = \left(g_0 + g_1M_t + g_2X_t^o\right)dt + \sigma_M(t)\,d\hat{Z}_t,$$

where

$$\sigma_M(t) = \left(V(t)\begin{bmatrix} \mu_1 \\ g_1 \end{bmatrix}^\top + \sigma_u\sigma^\top\right)(\sigma\sigma^\top)^{-1/2}, \tag{A1}$$

and $V(t)$ is the conditional variance of X_t^u satisfying

$$\frac{dV(t)}{dt} = g_1 V(t) + V(t)^\top g_1^\top + \sigma_u \sigma_u^\top$$
$$- \left(V(t) \begin{bmatrix} \mu_1 \\ h_1 \end{bmatrix}^\top + \sigma_u \sigma^\top \right) (\sigma\sigma^\top)^{-1} \left(V(t) \begin{bmatrix} \mu_1 \\ h_1 \end{bmatrix}^\top + \sigma_u \sigma^\top \right)^\top,$$

subject to $V(t_i) = v_{t_i}$.

We then have that the investor's portfolio selection problem in this more general case is equivalent to

$$\max_{\theta(t):t\in[0,T]} E[u(W(T))],$$

subject to the budget constraint

$$dW_t = r(X_t, t)W_t\, dt + \theta(t)^\top W_t(\mu(X_t, t) - r(X_t, t)\bar{\mathbf{1}})\, dt + \theta(t)^\top W_t \sigma_S(t)\, d\hat{Z}_t, \tag{A2}$$

and

$$dX_t = \mu_X(X_t, t)\, dt + \sigma_X(t)\, d\hat{Z}_t,$$

where

$$r(X_t, t) = r_0(t) + r_2(t)X_t^o,$$
$$\mu(X_t, t) = \mu_0(t) + \mu_1(t)M_t + \mu_2(t)X_t^o,$$
$$\mu_X(X_t, t) = \begin{bmatrix} g_0(t) + g_1(t)M_t + g_2 X_t^o \\ h_0(t) + h_1(t)M_t + h_2 X_t^o \end{bmatrix},$$

$$X_t = \begin{bmatrix} M_t \\ X_t^o \end{bmatrix}, \tag{A3}$$

and

$$\sigma_X(t) = \begin{bmatrix} \sigma_M(t) \\ \sigma_o(t) \end{bmatrix}.$$

A.1 The Solution to the Case in Which the Predictive Variables Are Unobservable for the Entire Horizon

We conjecture that

$$J(W, X, t) = \frac{\gamma}{1-\gamma} \left(\frac{\lambda W}{\gamma} + \eta e^{-r(T-t)} \right)^{1-\gamma} f(X, t)^{1-\gamma}. \tag{A4}$$

Assume that there are no redundant stocks and therefore $\sigma_S\sigma_S^\top$ is invertible. Then the optimal fraction of wealth invested in the stocks is given by

$$\pi = \frac{1}{\gamma}\left(\sigma_S\sigma_S^\top\right)^{-1}\left(\mu - r\bar{\mathbf{1}} + (1-\gamma)\sigma_S\sigma_X^\top\frac{f_X}{f}\right). \tag{A5}$$

Substituting the conjectured form of the solution into the HJB equation, we have

$$\begin{aligned} f_t &+ \frac{1}{2}Tr\left(\sigma_X\sigma_X^\top f_{XX^\top}\right) + \left(\frac{1-\gamma}{\gamma}(\mu - r\bar{\mathbf{1}})^\top\left(\sigma_S\sigma_S^\top\right)^{-1}\sigma_S\sigma_X^\top + \mu_X^\top\right) f_X \\ &+ \left(\frac{1}{2\gamma}(\mu - r\bar{\mathbf{1}})^\top\left(\sigma_S\sigma_S^\top\right)^{-1}(\mu - r\bar{\mathbf{1}}) + r\right) f + \frac{f_X^\top\Phi_1 f_X}{2f} = 0, \end{aligned} \tag{A6}$$

where

$$\Phi_1 = \sigma_X\left(\frac{(1-\gamma)^2}{\gamma}\sigma_S^\top\left(\sigma_S\sigma_S^\top\right)^{-1}\sigma_S - \gamma I\right)\sigma_X^\top.$$

We now further conjecture that

$$f(X,t) = e^{c(t) + X^\top d(t) + \frac{1}{2}X^\top Q(t)X}, \tag{A7}$$

where $Q(t)$ is set to be symmetric. Substituting equation (A7) into (A6), we obtain the following ordinary differential equations (ODEs) for c, d, and Q:

$$c'(t) + \frac{1}{2}d(t)^\top\Phi_1 d(t) + \frac{1}{2}Tr\left(\Phi_2\left(Q(t) + d(t)d(t)^\top\right)\right) + \Phi_3 d(t) + \varphi_c = 0, \tag{A8}$$

$$d'(t) + Q(t)(\Phi_1 + \Phi_2)d(t) + \Phi_4 d(t) + Q(t)\Phi_3^\top + \varphi_d = 0, \tag{A9}$$

and

$$Q'(t) + Q(t)(\Phi_1 + \Phi_2)Q(t) + \Phi_4 Q(t) + Q(t)\Phi_4^\top + \varphi_Q = 0, \tag{A10}$$

subject to the terminal conditions

$$c(T) = 0, \quad d(T) = 0, \quad \text{and} \quad Q(T) = 0,$$

where

$$\Phi_2 = \sigma_X \sigma_X^\top, \quad \Phi_3 = \frac{1-\gamma}{\gamma}\delta_0^\top \left(\sigma_S \sigma_S^\top\right)^{-1} \sigma_S \sigma_X^\top + \eta_0^\top,$$

$$\Phi_4 = \frac{1-\gamma}{\gamma}\delta_1^\top \left(\sigma_S \sigma_S^\top\right)^{-1} \sigma_S \sigma_X^\top + \eta_1^\top, \quad \varphi_c = \frac{1}{2\gamma}\delta_0^\top \left(\sigma_S \sigma_S^\top\right)^{-1}\delta_0 + r_0,$$

$$\varphi_d = \frac{1}{\gamma}\delta_1^\top \left(\sigma_S \sigma_S^\top\right)^{-1}\delta_0 + [r_1 \quad 0]^\top, \quad \varphi_Q = \frac{1}{\gamma}\delta_1^\top \left(\sigma_S \sigma_S^\top\right)^{-1}\delta_1,$$

$$\delta_0 = \mu_0 - r_0 \bar{\mathbf{1}}, \quad \delta_1 = (\mu_1, \mu_2 - \bar{\mathbf{1}} r_2),$$

$$\eta_0 = \begin{bmatrix} g_0 \\ h_0 \end{bmatrix}, \quad \text{and} \quad \eta_1 = \begin{bmatrix} g_1 & g_2 \\ h_1 & h_2 \end{bmatrix}.$$

The ODEs (A9) and (A10) are matrix Riccati equations, which can be solved using the standard method. By plugging the solutions of $d(t)$ and $Q(t)$ into equation (A8), $c(t)$ can be solved.

A.2 The Solution to the Case in Which the Predictive Variables Are Periodically Observable

Now we assume that a signal Y_t of X_t^u can be observed periodically at $t = t_i\,(i = 1, 2, \ldots, N)$. Moreover, Y_t satisfies

$$Y_t = X_t^u + \varepsilon_t, \tag{A11}$$

where ε_t represents the noise in the signal. We assume that ε_t is normally distributed with mean zero and variance–covariance matrix Σ, is not serially correlated, and is independent of any other random variables at any time.

Immediately before the observation of the predictive variables at t_i, $X_{t_i}^u$ is normally distributed with mean M_{t_i} and variance V_{t_i}. By Bayes's rule, the distribution of $X_{t_i}^u$ conditional on Y_{t_i} is also normal with mean

$$m_{t_i} = M_{t_i} + V_{t_i}(V_{t_i} + \Sigma)^{-1}(Y_{t_i} - M_{t_i}) \tag{A12}$$

and variance–covariance matrix

$$v_{t_i} = V_{t_i}(V_{t_i} + \Sigma)^{-1}\Sigma.$$

We conjecture that the value function at $t \in [t_{i-1}, t_i)$ in period $(i-1)$ is

$$J^{i-1}(W, X, t) = \frac{\gamma}{1-\gamma}\left(\frac{\lambda W}{\gamma} + \eta e^{-r(T-t)}\right)^{1-\gamma} f^{i-1}(X, t)^{1-\gamma}, \tag{A13}$$

with

$$f^{i-1}(X, t) = e^{c^{i-1}(t) + X^\top d^{i-1}(t) + \frac{1}{2}X^\top Q^{i-1}(t) X}, \tag{A14}$$

where X_t is as defined in equation (A3), and $c^{i-1}(t)$, $d^{i-1}(t)$, and $Q^{i-1}(t)$ are the coefficients in period $(i-1)$. To solve the value function, we first need to pin down the terminal conditions by calculating the expected value of J^i for given M_{t_i} and V_{t_i}.[26]

Let E^m denote the expectation operator with respect to M_{t_i} for given M_{t_i} and V_{t_i}. According to equation (A12), M_{t_i} is normally distributed with mean M_{t_i} and variance $\Lambda = V_{t_i}(V_{t_i} + \Sigma)^{-1}V_{t_i}$. Let

$$c = c^i(t_i), \tag{A15}$$

$$d = \begin{bmatrix} d_m \\ d_o \end{bmatrix} = d^i(t_i), \tag{A16}$$

and

$$Q = \begin{bmatrix} q_{mm} & q_{mo} \\ q_{mo}^\top & q_{oo} \end{bmatrix} = Q^i(t_i) \tag{A17}$$

be the parameter values of J^i at the beginning of ith period t_i and

$$\hat{X}_t = \begin{bmatrix} m_t \\ X_t^o \end{bmatrix}. \tag{A18}$$

We then have

$$\begin{aligned} E^m\big[f^i(\hat{X}, t_i)^{1-\gamma}\big] &= E^m\big[e^{(1-\gamma)(c+\hat{X}^\top d+\frac{1}{2}\hat{X}^\top Q\hat{X})}\big] \\ &= E^m\big[e^{(1-\gamma)(c+d_m^\top m+d_o^\top X^o+\frac{1}{2}(m^\top q_{mm}m+2m^\top q_{mo}X^o+(X^o)^\top q_{oo}X^o))}\big] \\ &= e^{(1-\gamma)(c+d_o^\top X^o+(X^o)^\top q_{oo}X^o)}E^m\big[e^{(1-\gamma)(d_m^\top m+m^\top q_{mo}X^o+\frac{1}{2}m^\top q_{mm}m)}\big] \\ &= \big(e^{\hat{c}^{i-1}+X^\top\hat{d}^{i-1}+\frac{1}{2}X^\top\hat{Q}^{i-1}X}\big)^{1-\gamma}, \end{aligned} \tag{A19}$$

where

$$\begin{aligned} \hat{c}^{i-1} &= c + \frac{1}{2}(1-\gamma)d_m^\top(\mathbf{I} - (1-\gamma)\Lambda q_{mm})^{-1}\Lambda d_m \\ &\quad - \frac{1}{2(1-\gamma)}\log[Tr(\mathbf{I} - (1-\gamma)\Lambda q_{mm})], \end{aligned} \tag{A20}$$

$$\hat{d}^{i-1} = \begin{bmatrix} d_m^{i-1} \\ d_o^{i-1} \end{bmatrix} = \begin{bmatrix} (\mathbf{I} - (1-\gamma)q_{mm}\Lambda)^{-1}d_m \\ d_o + (1-\gamma)q_{mo}^\top(\mathbf{I} - (1-\gamma)\Lambda q_{mm})^{-1}\Lambda d_m \end{bmatrix}, \tag{A21}$$

[26] Strictly speaking, M_{t_i} and V_{t_i} are values at t_i^-, immediately before Y_{t_i} is observed, and m_{t_i} and v_{t_i} are values at t_i, immediately after Y_{t_i} is observed.

and

$$\hat{Q}^{i-1} = \begin{bmatrix} q_{mm}^{i-1} & q_{mo}^{i-1} \\ q_{om}^{i-1} & q_{oo}^{i-1} \end{bmatrix}$$

$$= \begin{bmatrix} (\mathbf{I} - (1-\gamma)q_{mm}\Lambda)^{-1}q_{mm} & (\mathbf{I} - (1-\gamma)q_{mm}\Lambda)^{-1}q_{mo} \\ q_{mo}^{\top}(\mathbf{I} - (1-\gamma)\Lambda q_{mm})^{-1} & q_{oo} + (1-\gamma)q_{mo}^{\top}(\mathbf{I} - (1-\gamma)\Lambda q_{mm})^{-1}\Lambda q_{mo} \end{bmatrix}. \tag{A22}$$

The expected value function calculated above is of the objective function for the $(i-1)$th period and therefore provides the terminal conditions.

Therefore, $c^{i-1}(t)$, $d^{i-1}(t)$, and $Q^{i-1}(t)$ must solve the ODEs (A8), (A9), and (A10), with $c(t)$, $d(t)$, and $Q(t)$ replaced by $c^{i-1}(t)$, $d^{i-1}(t)$, and $Q^{i-1}(t)$, respectively, subject to the terminal conditions

$$c^{i-1}(t_i) = \hat{c}^{i-1}, \quad d^{i-1}(t_i) = \hat{d}^{i-1}, \quad \text{and } Q^{i-1}(t_i) = \hat{Q}^{i-1}.$$

In this way, the investor's problem can be solved iteratively, and in each iteration we only need to solve the same system of ODEs with different terminal conditions.

Appendix B. The One Stock, One State Variable Case

In the case of one stock and one predictive variable, as described by equations (1), (2), and (4), the conditional variance V_t of the predictive variable satisfies

$$\frac{dV(t)}{dt} = a_1 V(t)^2 + a_2 V(t) + a_3,$$

where $a_1 = -(\frac{\mu_1}{\sigma_S})^2 - \alpha_\nu^2$, $a_2 = (2g_1 - \frac{2\rho\sigma_X\mu_1}{\sigma_S})$, and $a_3 = (1-\rho^2)\sigma_X^2$. The solution is

$$V(t) = \frac{-a_2 - \eta}{2a_1} + \frac{\eta(2a_1V_0 + a_2 + \eta)}{a_1\big(2a_1V_0 + a_2 + \eta - (2a_1V_0 + a_2 - \eta)e^{\eta t}\big)},$$

where $\eta = \sqrt{a_2^2 - 4a_1a_3}$ ($a_2^2 - 4a_1a_3 > 0$ follows trivially from $a_1 < 0$ and $a_3 > 0$). As t goes to infinity, $V(t)$ converges to $\frac{-a_2 - \eta}{2a_1}$.

Plugging $\sigma_{M1}^2(t) = (\frac{\mu_1}{\sigma_S}V(t) + \rho\sigma_X)^2$, $\sigma_{M2}^2(t) = (\alpha_\nu V(t))^2$, and $\sigma_M^2(t) = \sigma_{M1}^2(t) + \sigma_{M2}^2(t)$ into the ODEs, we have

$$Q'(t) + \varphi_1(t)Q^2(t) + \varphi_2(t)Q(t) + \varphi_3 = 0, \tag{B1}$$

$$d'(t) + \varphi_1(t)Q(t)d(t) + \frac{1}{2}\varphi_2(t)d(t) + \varphi_4(t)Q(t) + \varphi_5 = 0, \tag{B2}$$

and

$$c'(t) + \frac{1}{2}\varphi_1(t)d^2(t) + \frac{1}{2}\sigma_M^2(t)Q(t) + \varphi_4(t)d(t) + \varphi_6 = 0, \tag{B3}$$

where

$$\varphi_1(t) = \frac{(1-\gamma)^2}{\gamma}\sigma_{M1}^2(t) + (1-\gamma)\sigma_M^2(t), \quad \varphi_2(t) = 2\left(\frac{1-\gamma}{\gamma}\frac{\mu_1}{\sigma_S}\sigma_{M1}(t) + g_1\right),$$

$$\varphi_3 = \frac{1}{\gamma}\left(\frac{\mu_1}{\sigma_S}\right)^2, \quad \varphi_4(t) = \frac{1-\gamma}{\gamma}\frac{\mu_0 - r}{\sigma_S}\sigma_{M1}(t) + g_0,$$

$$\varphi_5 = \frac{\mu_1(\mu_0 - r)}{\gamma\sigma_S^2}, \quad \text{and} \quad \varphi_6 = \frac{1}{2\gamma}\left(\frac{\mu_0 - r}{\sigma_S}\right)^2 + r.$$

For the no-periodic-news case, (Q, d, c) is the solution to (B1) to (B3) subject to the terminal conditions $Q(T) = d(T) = c(T) = 0$. We can also obtain a series solution for Q by transforming the nonlinear ODE for Q into a linear homogeneous ODE. The solution for (Q, d, c) in terms of constants b_j is[27]

$$Q(t) = \frac{\gamma\sigma_M'(t)}{(1-\gamma)\sigma_M(t)^2}\frac{\sum_{j=0}^{\infty} j b_j(\sigma_M(t) - \sigma_M(T))^{j-1}}{\sum_{j=0}^{\infty} b_j(\sigma_M(t) - \sigma_M(T))^j}, \tag{B4}$$

$$d(t) = \int_t^T e^{\int_t^s\left(\varphi_1(k)Q(k) + \frac{1}{2}\varphi_2(k)\right)dk}(\varphi_4(s)Q(s) + \varphi_5)\,ds, \tag{B5}$$

and

$$c(t) = \int_t^T \left(\frac{1}{2}\varphi_1(s)d^2(s) + \frac{1}{2}\sigma_M^2(s)Q(s) + \varphi_4(s)d(s) + \varphi_6\right)ds. \tag{B6}$$

For the periodic news case, $(Q^{i-1}(t), d^{i-1}(t), c^{i-1}(t))(t \in [t_{i-1}, t_i))$ is the solution to (B1) to (B3) with $V(t)$ replaced by $V^{i-1}(t)$, subject to the terminal conditions $Q^{i-1}(t_i) = \hat{Q}^{i-1}, d^{i-1}(t_i) = \hat{d}^{i-1}, c^{i-1}(t_i) = \hat{c}^{i-1}$, where $(\hat{Q}^{i-1}, \hat{d}^{i-1}, \hat{c}^{i-1})$ is as defined in (17). Similar to the no-periodic-news case, we can also obtain a series solution as follows:

$$Q^{i-1}(t) = \hat{Q}^{i-1} + \frac{\gamma\sigma_M'(t)}{(1-\gamma)\sigma_M(t)^2}\frac{\sum_{j=0}^{\infty} j b_j^{i-1}(\sigma_M(t) - \sigma_M(t_i))^{j-1}}{\sum_{j=0}^{\infty} b_j^{i-1}(\sigma_M(t) - \sigma_M(t_i))^j}, \tag{B7}$$

[27] The explicit expressions for b_j and b_j^{i-1} in (B7) are presented in an earlier version of this paper, which is available from the authors and omitted in this version to save space.

$$d^{i-1}(t) = \hat{d}^{i-1} e^{\int_t^{t_i}\left(\varphi_1(s)Q^{i-1}(s)+\frac{1}{2}\varphi_2(s)\right)ds} + \int_t^{t_i} e^{\int_t^{s}\left(\varphi_1(k)Q^{i-1}(k)+\frac{1}{2}\varphi_2(k)\right)dk}\left(\varphi_4(s)Q^{i-1}(s)+\varphi_5\right)ds, \tag{B8}$$

and

$$c^{i-1}(t) = \hat{c}^{i-1} + \int_t^{t_i}\left(\frac{1}{2}\varphi_1(s)(d^{i-1}(s))^2 + \frac{1}{2}\sigma_M^2(s)Q^{i-1}(s) + \varphi_4(s)d^{i-1}(s) + \varphi_6\right)ds. \tag{B9}$$

Appendix C. Proofs of Propositions 1 to 3

Proof of Proposition 1: By Liptser and Shiryaev (2001, theorems 12.5 and 12.7), the sigma algebra generated by $(S_0, \hat{Z}_1, \hat{Z}_3)$ and the one generated by the stock price S and the signal process ν are equivalent, and $\hat{Z}_1$ and $\hat{Z}_3$ are Wiener processes. Therefore, a trading strategy is adapted to the original filtration generated by the stock price process S and the signal process ν if and only if it is adapted to the filtration generated by $(S_0, \hat{Z}_1, \hat{Z}_3)$. The dynamics in (8) follow from theorem 10.5.1 of Kallianpur (1980). Thus, the optimal trading strategy that depends only on (W_t, M_t, t) is also optimal given the original filtration. Next we show that the function $J(W, M, t)$ in Proposition 1 is indeed the value function, and the stated trading strategy is optimal.

First, the expressions for c, d, and Q follow from Appendix B. Next, applying Itô's lemma to the stochastic process $L_t \equiv J(W_t, M_t, t)$, we have that the drift of L_t is always nonpositive for any admissible trading strategy and equal to zero for the candidate optimal trading strategy since the candidate value function in Proposition 1 satisfies the HJB equation (13). In addition, the stochastic integral of L_t is a martingale. Therefore, L_t is a super martingale for all admissible trading strategies and a martingale for the candidate optimal trading strategy. Thus, we have for all admissible trading strategies

$$J(W, M, 0) = L_0 \geq E[L_T] = E[u(W_T) \mid W_0 = W, M_0 = M]$$

with equality for the candidate optimal strategy. This shows the optimality of the stated trading strategy and that $J(W_t, M_t, t)$ is the value function.

The expression for the optimal trading strategy follows from the fact that θ_t^* maximizes the left-hand side of the HJB equation (13), that is,

$$\theta_t^* = -\frac{(\mu_0 + \mu_1 M_t - r)J_W + \sigma_S \sigma_M J_{WM}}{\sigma_S^2 J_{WW}}. \quad \text{Q.E.D.}$$

Proof of Proposition 2: We prove this by backward induction. In the last period N, the proof is the same as that for Proposition 1. For period $N-1$, after first taking the expectation of the value function J^N with respect to the next observation y_{t_N}, one gets the value function $J^{N-1}(W_{t_N^-}, M_{t_N^-}, t_N^-)$ at t_N^-, and then the proof for Proposition 1 applies for this period. Continuing this procedure completes the proof for Proposition 2. Q.E.D.

Proof of Proposition 3: For the unobservable case, we divide the whole investment horizon into the same N periods as for the periodically observable case. We use Q_u^i, d_u^i, and $V_u^i(t)$ to denote the Q, d, and V functions of the i^{th} period ($i = 1, 2, \ldots, N$). Because there is no observation, these functions have no jumps, and we have $Q_u^{i-1}(t_i) = Q_u^i(t_i), d_u^{i-1}(t_i) = d_u^i(t_i)$, and $V_u^{i-1}(t_i) = V_u^i(t_i)$ (recall that t_i is the end of the $(i-1)$th period and also the beginning of the i^{th} period). Moreover, Q_u^{i-1} and d_u^{i-1} follow equations (B1) and (B2) with terminal conditions $(Q_u^N(T) = 0, d_u^N(T) = 0)$ and $(Q_u^{i-1}(t_i) = Q_u^i(t_i), d_u^{i-1}(t_i) = d_u^i(t_i))$ for $i = 2, \ldots, N$.

We denote $V_p^i(t)$ as the i^{th} period V function ($i = 1, 2, \ldots, N$) for the periodically observable case. Because of the arrival of new information, we have

$$V_p^i(t_i) = \frac{V_p^{i-1}(t_i)\alpha_\varepsilon^{-2}}{V_p^{i-1}(t_i) + \alpha_\varepsilon^{-2}}. \tag{C1}$$

Now we define

$$Q_p^i(t) = \frac{Q_u^i(t)}{1 + (1-\gamma)\big(V_u^i(t) - V_p^i(t)\big)Q_u^i(t)} \tag{C2}$$

and

$$d_p^i(t) = \frac{d_u^i(t)}{1 + (1-\gamma)\big(V_u^i(t) - V_p^i(t)\big)Q_u^i(t)}. \tag{C3}$$

It can be shown that the defined $Q_p^i(t)$ and $d_p^i(t)$ are the Q and d functions of the i^{th} period for the periodically observable case by verifying that[28]

(1) they follow the same ODEs as those of the periodically observable case, and
(2) they have the same terminal conditions.

To prove (1), we only need to realize that the sole difference between the ODEs for the unobservable case and those for the periodically observable case is $V(t)$. From definitions (C2) and (C3), we can easily derive

$$Q_u^i(t) = \frac{Q_p^i(t)}{1 - (1-\gamma)\big(V_u^i(t) - V_p^i(t)\big)Q_p^i(t)}, \tag{C4}$$

$$d_u^i(t) = \frac{d_p^i(t)}{1 - (1-\gamma)\big(V_u^i(t) - V_p^i(t)\big)Q_p^i(t)}, \tag{C5}$$

[28] A more detailed proof of these claims is available from the authors.

$$\frac{dQ_u^i(t)}{dt} = \frac{\dfrac{dQ_p^i(t)}{dt} + (1-\gamma)\big(Q_p^i(t)\big)^2\left(\dfrac{dV_u^i(t)}{dt} - \dfrac{dV_p^i(t)}{dt}\right)}{\big[1-(1-\gamma)\big(V_u^i(t)-V_p^i(t)\big)Q_p^i(t)\big]^2}, \tag{C6}$$

and

$$\frac{d\left(d_u^i(t)\right)}{dt} = \frac{\dfrac{d\left(d_p^i(t)\right)}{dt}}{1-(1-\gamma)\big(V_u^i(t)-V_p^i(t)\big)Q_p^i(t)} - \frac{(1-\gamma)d_p^i(t)\left[\left(\dfrac{dV_u^i(t)}{dt} - \dfrac{dV_p^i(t)}{dt}\right)Q_p^i(t) + \big(V_u^i(t)-V_p^i(t)\big)\dfrac{dQ_p^i(t)}{dt}\right]}{\big[1-(1-\gamma)\big(V_u^i(t)-V_p^i(t)\big)Q_p^i(t)\big]^2}. \tag{C7}$$

In addition, by (9) both $V_u^i(t)$ and $V_p^i(t)$ satisfy

$$\frac{dV(t)}{dt} = -\left[\left(\frac{\mu_1}{\sigma_S}\right)^2 + \alpha_\nu^2\right]V(t)^2 + \left(2g_1 - \frac{2\rho\sigma_X\mu_1}{\sigma_S}\right)V(t) + (1-\rho^2)\sigma_X^2 \tag{C8}$$

and

$$\sigma_{M1}(t) = \frac{\mu_1}{\sigma_S}V(t) + \rho\sigma_X, \quad \sigma_{M2}(t) = (\alpha_\nu V(t))^2. \tag{C9}$$

Plugging (C4) to (C9) into the ODEs for the no-periodic-news case ((B1) and (B2)) and after straightforward but tedious algebra simplification, one can verify that $Q_p^i(t)$ and $d_p^i(t)$ indeed satisfy the ODEs for the case with periodic observation, that is, claim (1) holds.

To prove (2), we start from the last period. According to definitions (C2) and (C3), it is trivial to see that $Q_p^N(T) = Q_u^N(T) = 0$ and $d_p^N(T) = d_u^N(T) = 0$. So, $Q_p^N(t)$ and $d_u^N(t)$ are indeed the Q and d functions for the periodically observable case.

Now suppose $Q_p^i(t)$ and $d_p^i(t)$ are the Q and d functions of the i^{th} period for the periodically observable case. So, at the beginning of the i^{th} period ($t = t_i$), we have

$$Q_u^i(t_i) = \frac{Q_p^i(t_i)}{1-(1-\gamma)\big(V_u^i(t_i)-V_p^i(t_i)\big)Q_p^i(t_i)} \tag{C10}$$

and

$$d_u^i(t_i) = \frac{d_p^i(t_i)}{1-(1-\gamma)\big(V_u^i(t_i)-V_p^i(t_i)\big)Q_p^i(t_i)} \tag{C11}$$

according to (C4) and (C5).

Next we move to the $(i-1)^{\text{th}}$ period. Considering the end of the $(i-1)^{\text{th}}$ period, we have $Q_u^{i-1}(t_i) = Q_u^i(t_i)$ and $d_u^{i-1}(t_i) = d_u^i(t_i)$ for the unobservable case and thus

$$Q_u^i(t_i) = Q_u^{i-1}(t_i) = \frac{Q_p^{i-1}(t_i)}{1-(1-\gamma)\big(V_u^{i-1}(t_i) - V_p^{i-1}(t_i)\big)Q_p^{i-1}(t_i)},$$

$$d_u^i(t_i) = d_u^{i-1}(t_i) = \frac{d_p^{i-1}(t_i)}{1-(1-\gamma)\big(V_u^{i-1}(t_i) - V_p^{i-1}(t_i)\big)Q_p^{i-1}(t_i)},$$

according to definitions of Q_p^{i-1} and d_p^{i-1}.

Then by (C10) and (C11) and noting that $V_u^{i-1}(t_i) = V_u^i(t_i)$ because the conditional volatility does not jump for the unobservable case, it can be shown that

$$Q_p^{i-1}(t_i) = \frac{Q_p^i(t_i)}{1-(1-\gamma)\big(V_p^{i-1}(t_i) - V_p^i(t_i)\big)Q_p^i(t_i)}$$

and

$$d_p^{i-1}(t_i) = \frac{d_p^i(t_i)}{1-(1-\gamma)\big(V_p^{i-1}(t_i) - V_p^i(t_i)\big)Q_p^i(t_i)}.$$

By (C1) and (17), we have

$$Q_p^{i-1}(t_i) = \frac{Q_p^i(t_i)}{1-(1-\gamma)V_M(t_i)Q_p^i(t_i)}$$

and

$$d_p^{i-1}(t_i) = \frac{d_p^i(t_i)}{1-(1-\gamma)V_M(t_i)Q_p^i(t_i)},$$

which are the same as the $(i-1)^{\text{th}}$ period terminal conditions (17) for the periodically observable case and thus claim (2) holds.

Finally, since in the first period $V_p^1(t) = V_u^1(t)$ for the same initial value of $V(0)$, we have $Q_p^1(t) = Q_u^1(t)$ and $d_p^1(t) = d_u^1(t)$ by (C4) and (C5). In addition, given the same $M(0)$, the trading strategies are the same by Proposition 2. Q.E.D.

Appendix D. Parameter Estimation Method

Duffie et al. (2000) show that the conditional characteristic function of the diffusion processes described in (1) and (2) has a closed form. Using their results, we obtain the characteristic function

$$\psi(u_s, u_X, s_{t+\tau}, X_{t+\tau} \mid s_t, X_t)$$

$$= \exp[\alpha(\tau, u_s, u_X) + \beta_s(\tau, u_s, u_X)s_t + \beta_X(\tau, u_s, u_X)X_t],$$

where $\alpha(\tau, u_s, u_X)$, $\beta_s(\tau, u_s, u_X)$, and $\beta_X(\tau, u_s, u_X)$ satisfy the following ODEs:

$$\begin{bmatrix} \dot{\beta}_s \\ \dot{\beta}_X \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ \mu_1 & g_1 \end{bmatrix} \begin{bmatrix} \beta_s \\ \beta_X \end{bmatrix}$$

and

$$\dot{\alpha} = \begin{bmatrix} \mu_0 - \frac{1}{2}\sigma_s^2 & g_0 \end{bmatrix} \begin{bmatrix} \beta_s \\ \beta_X \end{bmatrix} + \frac{1}{2}[\beta_s \beta_X] \begin{bmatrix} \sigma_s^2 & \rho\sigma_s\sigma_X \\ \rho\sigma_s\sigma_X & \sigma_X^2 \end{bmatrix} \begin{bmatrix} \beta_s \\ \beta_X \end{bmatrix},$$

with boundary conditions $\alpha(0, u_s, u_X) = 0$, $\beta_s(0, u_s, u_X) = u_s$, and $\beta_X(0, u_s, u_X) = u_X$. The above ODEs have the closed-form solutions

$$\beta_s(\tau, u_s, u_X) = u_s,$$

$$\beta_X(\tau, u_s, u_X) = \left(u_X + \frac{\mu_1}{g_1}u_s\right)e^{g_1\tau} - \frac{\mu_1}{g_1}u_s,$$

and

$$\begin{aligned} \alpha(\tau, u_s, u_X) = {} & \frac{1}{4g_1}\sigma_X^2 \left(u_X + \frac{\mu_1}{g_1}u_s\right)^2 \left(e^{2g_1\tau} - 1\right) \\ & + \frac{1}{g_1}\left(g_0 + \rho\sigma_s\sigma_X u_s - \frac{\mu_1}{g_1}\sigma_X^2 u_s\right)\left(u_X + \frac{\mu_1}{g_1}u_s\right)\left(e^{g_1\tau} - 1\right) \\ & + \left(\left(\frac{1}{2}\sigma_s^2 + \frac{1}{2}\frac{\mu_1^2}{g_1^2}\sigma_X^2 - \frac{\mu_1}{g_1}\rho\sigma_S\sigma_X\right)u_s^2 + \left(\mu_0 - \frac{1}{2}\sigma_s^2 - \frac{\mu_1 g_0}{g_1}\right)u_s\right)\tau. \end{aligned}$$

It can be easily verified that, conditional on s_t and X_t, the continuously compounded return $R_{t+1} = s_{t+1} - s_t$ and X_{t+1} are jointly Gaussian, and the first and second moments are

$$M_1 = E_t(R_{t+1}) = \left.\frac{\partial\psi}{\partial u_s}\right|_{u_s=0,u_X=0} = \left(\mu_0 - \frac{1}{2}\sigma_s^2 - \frac{\mu_1 g_0}{g_1}\right)$$

$$+ \frac{\mu_1}{g_1}(e^{g_1}-1)\left(X_t + \frac{g_0}{g_1}\right),$$

$$M_2 = E_t(X_{t+1}) = \left.\frac{\partial\psi}{\partial u_X}\right|_{u_s=0,u_X=0} = \frac{g_0}{g_1}(e^{g_1}-1) + e^{g_1}X_t,$$

$$\sigma_1^2 = \mathrm{var}_t(R_{t+1}) = \left.\frac{\partial^2\psi}{\partial u_s^2}\right|_{u_s=0,u_X=0} - (E_t(R_{t+1}))^2$$

$$= \frac{1}{2g_1}(e^{2g_1}-1)\frac{\mu_1^2}{g_1^2}\sigma_X^2 + \frac{2\mu_1}{g_1^2}(e^{g_1}-1)\left(\rho\sigma_s\sigma_X - \frac{\mu_1}{g_1}\sigma_X^2\right)$$

$$+ \sigma_s^2 + \frac{\mu_1^2}{g_1^2}\sigma_X^2 - \frac{2\mu_1}{g_1}\rho\sigma_s\sigma_X,$$

$$\sigma_2^2 = \mathrm{var}_t(X_{t+1}) = \left.\frac{\partial^2\psi}{\partial u_X^2}\right|_{u_s=0,u_X=0} - (E_t(X_{t+1}))^2 = \frac{1}{2g_1}(e^{2g_1}-1)\sigma_X^2,$$

and

$$\rho_{12}\sigma_1\sigma_2 = \mathrm{cov}_t(R_{t+1}, X_{t+1}) = \left.\frac{\partial^2\psi}{\partial u_s \partial u_X}\right|_{u_s=0,u_X=0} - E_t(R_{t+1})E_t(X_{t+1})$$

$$= \frac{\mu_1}{2g_1^2}(e^{g_1}-1)^2\sigma_X^2 + \frac{1}{g_1}(e^{g_1}-1)\rho\sigma_s\sigma_X.$$

Define

$$a_{11} = \mu_0 - \frac{1}{2}\sigma_s^2 - \frac{\mu_1 g_0}{g_1} + \frac{\mu_1 g_0}{g_1^2}(e^{g_1}-1), \quad a_{12} = \frac{\mu_1}{g_1}(e^{g_1}-1),$$

$$b_{11} = \frac{g_0}{g_1}(e^{g_1}-1), \quad \text{and} \quad b_{12} = e^{g_1}.$$

Let $\Theta \equiv (a_{11}, a_{12}, b_{11}, b_{12}, \sigma_1, \sigma_2, \rho_{12})$. The log likelihood function can then be calculated as

$$
\begin{aligned}
£(\Theta) &= \log f(X_1;\Theta) + \sum_{t=2}^{t=T} \log f(r_t, X_t \mid X_{t-1};\Theta) \\
&= -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\frac{\sigma_2^2}{1-b_{12}^2} - \frac{(X_1 - b_{11}/(1-b_{12}))^2}{2\sigma_2^2/\left(1-b_{12}^2\right)} \\
&\quad - \frac{T-1}{2}\left(\log 2\pi + \log\sigma_1^2 + \log\sigma_2^2 + \log\left(1-\rho_{12}^2\right)\right) \\
&\quad - \frac{1}{2(1-\rho_{12}^2)}\sum_{t=2}^{t=T}\left\{\frac{(R_t - a_{11} - a_{12}X_{t-1})^2}{\sigma_1^2} + \frac{(X_t - b_{11} - b_{12}X_{t-1})^2}{\sigma_2^2}\right. \\
&\quad \left. + \frac{2\rho_{12}(R_t - a_{11} - a_{12}X_{t-1})(X_t - b_{11} - b_{12}X_{t-1})}{\sigma_1\sigma_2}\right\}.
\end{aligned}
$$

We then use the maximum likelihood estimation method to estimate the parameters.

REFERENCES

Bacchetta, Philippe, and Eric van Wincoop, 2005, Rational inattention: A solution to the forward discount and predictability puzzles, Working paper, University of Virginia.

Balduzzi, Pierluigi, Edwin J. Elton, and T. Clifton Green, 1997, Economic news and the yield curve: Evidence from the U.S. Treasury market, Working paper, New York University.

Basin, Michael V., 2003, On optimal filtering for polynomial system states, *Journal of Dynamic Systems, Measurement, and Control* 125, 123–125.

Brennan, Michael J., 1998, The role of learning in dynamic portfolio decisions, *European Finance Review* 1, 295–306.

Brennan, Michael, and Yihong Xia, 2002, Tay's as good as cay, Working paper, Anderson School Business, University of California, Los Angeles.

Campbell John, Andrew Lo, and Craig MacKinlay, 1997, *The Econometrics of Financial Markets* (Princeton University Press, Princeton, New Jersey).

Chen, Guanrong, ed., 1993, *Approximate Kalman Filtering* (World Science Publishing Company, Teaneck, New Jersey).

Clark, Martin, 1978, The design of robust approximations to the stochastic differential equations of nonlinear filtering, in J. K. Skwirzynski, ed.: *Communication Systems and Random Process Theory* (Sijthoff and Noordhoff, Alphen aan den Rijn, The Netherlands).

Cready, William M., and Patricia G. Mynatt, 1991, The information content of annual reports: A price and trading response analysis, *The Accounting Review* 66, 291–312.

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

Detemple, Jérôme, 1991, Further results on asset pricing with incomplete information, *Journal of Economic Dynamics and Control* 15, 425–453.

Detemple, Jérôme, Rene Garcia, and Marcel Rindisbacher, 2003, A Monte Carlo method for optimal portfolios, *Journal of Finance* 58, 401–446.

Detemple, Jérôme, and Richard Kihlstrom, 1987, Acquisition d'information dans un modèle intertemporel en temps continu, *L'Actualite Économique* 63, 118–137.

Duffie, Darrell, and David Lando, 2001, Term structures of credit spreads with incomplete accounting information, *Econometrica* 69, 633–664.

Duffie, Darrell, Jun Pan, and Kenneth Singleton, 2000, Transform analysis and asset pricing for affine jump-diffusions, *Econometrica* 68, 1343–1376.

Fama, Eugene, and Kenneth French, 1988, Dividend yields and expected stock returns, *Journal of Financial Economics* 22, 3–27.

Fiske, Susan, and Shelley Taylor, 1991, *Social Cognition*, 2nd ed. (McGraw-Hill, New York).

Fleming, Wendell H., and Raymond W. Rishel, 1975, *Deterministic and Stochastic Optimal Control* (Springer, New York).

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

Hahn, Jaehoon, and Hangyong Lee, 2001, On the estimation of the consumption-wealth ratio: Cointegrating parameter instability and its implications for stock return forecasting, Working paper, Columbia University.

Hodrick, Robert, 1992, Dividend yields and expected stock returns: Alternative procedures for inference and measurement, *Review of Financial Studies* 5, 357–386.

Hong, Harrison, Walter Torous, and Rossen Valkanov, 2002, Do industries lead the stock market? Inattention, delayed reaction and cross-asset return predictability, Working paper, Princeton University.

Inoue, Atsushi, and Lutz Kilian, 2004, In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews* 23, 371–402.

James, Matthew R., Vikram Krishnamurthy, and Francois Le Gland, 1996, Time discretization of continuous-time filters and smoothers for HMM parameter estimation, *IEEE Transactions on Information Theory* 42, 593–605.

Judd, Kenneth L., 1998, *Numerical Methods in Economics* (MIT Press, Cambridge, Massachusetts).

Kahneman, Daniel, 1973, *Attention and Effort* (Prenctice Hall, Englewood Cliffs, New Jersey).

Kallianpur, Gopinath, 1980, *Stochastic Filtering Theory* (Springer-Verlag, New York).

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

Keim, Donald, and Robert Stambaugh, 1986, Predicting returns in stock and bond markets, *Journal of Financial Economics* 17, 357–390.

Lee, Bong-Soo, 1992, Causal relations among stock returns, interest rates, real activity, and inflation, *Journal of Finance* 47, 1591–1603.

Lettau, Martin, and Sydney Ludvigson, 2001, Consumption, aggregate wealth, and expected stock returns, *Journal of Finance* 56, 815–849.

Liptser, Robert, and Albert Shiryaev, 2001, *Statistics of Random Processes II: Applications* (Springer-Verlag, Berlin).

Liu, Hong, 2004, Optimal consumption and investment with transaction costs and multiple risky assets, *Journal of Finance* 59, 289–338.

Liu, Hong, and Mark Loewenstein, 2002, Optimal portfolio selection with transaction costs and finite horizons, *Review of Financial Studies* 15, 805–835.

McQueen, Grant, and V. Vance Roley, 1993, Stock prices, news, and business conditions, *Review of Financial Studies* 6, 683–707.

Merton, Robert C., 1992, *Continuous-Time Finance* (Blackwell, Cambridge, Massachusetts).

Moscarini, Giuseppe, 2004, Limited information capacity as a source of inertia, *Journal of Economic Dynamics and Control* 28, 2003–2035.

Patelis, Alex, 1997, Stock return predictability and the role of monetary policy, *Journal of Finance* 52, 1951–1972.

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563–602.

Shoji, Isao, and Tohru Ozaki, 1998, Estimation for nonlinear stochastic differential equations by a local linearization method, *Stochastic Analysis and Applications* 16, 733–752.

Sims, Christopher, 2003, Implications of rational inattention, *Journal of Monetary Economics* 50, 665–690.

Singleton, Kenneth, 2001, Estimation of affine asset pricing models using the empirical characteristic function, *Journal of Econometrics* 102, 111–141.

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

Woodruff, Catherine S., and Andrew J. Senchack Jr., 1988, Intradaily price-volume adjustments of NYSE stocks to unexpected earnings, *Journal of Finance* 43, 467–491.

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