SIGNAL EXTRACTION AND RATIONAL INATTENTION

YULEI LUO and ERIC R. YOUNG*

In this paper we examine the implications of two theories of informational frictions, signal extraction (SE) and rational inattention (RI), for optimal decisions and economic dynamics within the linear-quadratic-Gaussian (LQG) setting. We first show that if the variance of the noise and channel capacity (or marginal information cost) is fixed exogenously in the SE and RI problems, respectively, the two environments lead to different policy and equilibrium asset pricing implications. Second, we find that if the signal-to-noise ratio and capacity in the SE and RI problems are fixed, respectively, the two theories generate the same policy implications in the univariate case, but different policy implications in the multivariate case. We also show that our results do not depend on the presence of correlation between fundamental and noise shocks. We then discuss the applications to macroeconomic models of permanent income and price-setting. (JEL C61, D81, E21)

I. INTRODUCTION

The objective of this paper is to compare two theories of information frictions. In the first theory, agents observe the state of the world with exogenously imposed noise; we refer to this environment as the signal extraction (SE) model.[1] The second theory permits agents to design the distribution of noise terms by focusing limited attention on certain variables at the expense of others; following Sims (2003) we refer to this model as the rational inattention (RI) model.[2] In both models, agents respond to changes in the true underlying state slowly because it takes time for them to learn exactly what the new state is. The key difference is that the noise distribution under RI is an outcome of optimal choice, and thus can adapt to changing circumstances in the economy, whereas under SE the noise distribution is technologically fixed. Our goal is to explore this crucial distinction.

We explore the distinction between SE and RI within a linear-quadratic-Gaussian (LQG) setting. We first study a univariate case for which the variance-covariance matrix of the noise shocks can be solved in closed form. First, we note that (at least within the LQG framework) any variance-covariance structure for the

Luo: School of Economics and Finance, The University of Hong Kong, Hong Kong, Hong Kong. Phone 852 2859 1042, Fax 852 2548 1152. E-mail yluo@econ.hku.hk

Young: Department of Economics, University of Virginia, Charlottesville, VA 22904. Phone 434 924 3811, Fax 434 982 2904. E-mail ey2d@virginia.edu

1. Muth (1960) first applied the SE structure to an economic model (namely the permanent income model of Friedman 1957); later SE problems were extended to models with multiple agents by Townsend (1983) and Sargent (1991). More recent studies include Morris and Shin (2002), Wang (2004), and Angeletos and La'O (2010).

2. Luo (2008) was an early contribution to the RI literature; other contributions include Adam (2005), Kasa (2006), Luo (2010), Luo and Young (2010a), Maćkowiak and Wiederholt (2009), Melosi (2009), Reis (2010), Van Nieuwerburgh and Veldkamp (2010), Paciello and Wiederholt (2013), and Kim, Ko, and Yun (2012).

ABBREVIATIONS

LQG: Linear-Quadratic-Gaussian
RBC: Real Business Cycle
RI: Rational Inattention
SE: Signal Extraction
SNR: Signal-to-Noise Ratio
SOE: Small Open Economy

---

noise shocks derived from the RI problem can be exogenously imposed in the SE environment, so there is a sense in which the two environments are observationally equivalent.

The next result we find is that if the variance of the noise itself is fixed, we can use a policy experiment to distinguish SE from RI. Suppose that the variance of an exogenous shock is scaled up because of a change in policy. In the SE problem where the distribution of noise is exogenous, an increase in the variance of the shock will lead to a different solution for the conditional variance and Kalman gain; consequently, the change in policy will lead to a change in the model's dynamic behavior and the agent's payoffs. In contrast, in the RI problem, if channel capacity is fixed, a change in the variance of the exogenous shock will lead to the same change in the conditional variance of the state and the variance of the noise, but will not affect the Kalman gain. That is, inattentive agents with fixed capacity will behave as if they face noise shocks whose nature changes systematically as the dynamic properties of the economy change with policy. Furthermore, we show that if we assume that the marginal cost of information is fixed, capacity will be elastic with respect to a change in policy; consequently, the Kalman gain in this case will also adjust in response to the policy change. The losses of agents due to imperfect information depend on the value of the Kalman gain. Therefore, in general, SE and RI imply different policy recommendations.

In contrast, we also find that in the univariate case, if the ratio of the variance of the exogenous shock to that of the noise (i.e., the signal-to-noise ratio, SNR) is fixed, the SE and RI problems are observationally equivalent in the sense that they lead to the same dynamics of the model economy when the ratio of the conditional variance to that of the noise in the SE problem equals a particular function of the channel capacity in the RI problem. Our results do not depend on the presence or absence of correlation between fundamental and noise shocks. Specifically, in the presence of the correlation, a change in the variance of the exogenous shock does not change the dynamic behavior of the model in the fixed capacity case, whereas it changes the model's dynamics in the fixed information-processing cost case in which both the variance of noise and the Kalman gain are affected by the interactions between the correlation, the variance of the fundamental shock, and the conditional variance.

We then move on to study the multivariate case. In this case, given channel capacity the conditional variance-covariance matrix can be obtained by solving a semidefinite programming problem in which the inattentive agent minimizes the expected loss due to information-processing constraints. After computing the optimal steady-state conditional variance-covariance matrix, we can recover the variance-covariance matrix of the noise vector and then determine the Kalman gain. In this case, we show that SE and RI generate different dynamics after a change in the variance of the exogenous shock even if the SNR is fixed.

However, when modeling the multivariate SE problem, it is difficult to specify the process of the vector of noises ex ante without prior knowledge about the states. Ad hoc assumptions on the nature of the noise might be inconsistent with the underlying efficiency conditions (equalization of the marginal utility of additional capacity across variables).[3] Therefore, RI provides a micro-founded method for specifying the stochastic properties of the noises. It is worth noting that in the multivariate RI problem, the agent's preference, budget constraint, and information-processing constraints jointly determine the values of the conditional variance of the state, the variance of the noise, and the Kalman gain, whereas in the multivariate SE problem given the variance of the noise, the propagation equation updating the conditional variance based on the budget constraint is used to determine the conditional variance and then the Kalman gain.

Finally, we present two canonical examples to illustrate our results. First, we study a permanent income model to explore the different consumption and asset pricing implications of SE and RI. We argue that the RI model with fixed information-processing cost does a better job at replicating the different consumption behavior in emerging and developed small open economies (SOEs), and thus helps us distinguish RI from SE.[4] We also discuss a simplified model of price setting in the vein of Maćkowiak and Wiederholt (2009) that optimal attention allocation between the aggregate and firm-specific shocks can also help distinguish RI from SE.

---

3. See Melosi (2009) for an empirical investigation of this issue.

4. Aguiar and Gopinath (2007) document that the relative volatility of consumption growth to income growth in emerging countries is significantly greater than that in developed countries.

In that model, Paciello and Wiederholt (2013) show that optimal monetary policy under SE and RI differs precisely because the monetary authority can manipulate the attention decision; in other words, it *matters* whether agents live in an SE or RI world.

The remainder of the paper is organized as follows. Section II examines optimal decisions and economic dynamics in an LQG setting with SE. Section III presents the RI version of the model and compares different implications of RI and SE for dynamics and policy within the LQG setting. Section IV presents applications to models of permanent income and price setting. Section V briefly discusses the different implications of SE and RI for consumption in a continuous-time setting. Section VI concludes.

II. SIGNAL EXTRACTION IN AN LQG MODEL

A. *Full-Information Rational Expectations LQG Model*

Consider the following LQG model:

$$(1) \quad v(s_0) = \max_{\{c_t, s_{t+1}\}}$$
$$E_0\left[\sum_{t=0}^{\infty} \beta^t \left(s_t^T Q s_t + c_t^T R c_t + 2c_t^T W s_t\right)\right],$$

subject to

$$(2) \qquad s_{t+1} = As_t + Bc_t + \varepsilon_{t+1},$$

with s_0 known and given, where $\beta < 1$ is the discount factor, s_t is a $(n \times 1)$ state vector, c_t is a $(k \times 1)$ control vector, ε_{t+1} is an iid $(n \times 1)$ vector of Gaussian random variables with mean 0 and covariance matrix Ω, and $E_t[\cdot]$ denotes the mathematical expectation of a random variable given information processed at t. We assume that Q, R, and W are such that the objective function is jointly concave in s_t and c_t, and the usual conditions required for the optimal policy to exist are satisfied.

When the agent can fully observe the state s_t, the model is a standard linear-quadratic regulator problem. Solving the corresponding Bellman equation

$$s_t^T P s_t = \max_{c_t} \left\{ s_t^T Q s_t + c_t^T R c_t + 2c_t^T W s_t \right.$$
$$+ \beta E_t\left[\left(s_t^T A^T + c_t^T B^T + \varepsilon_{t+1}^T\right)\right.$$
$$\left.\left. \times P \times \left(As_t + Bc_t + \varepsilon_{t+1}\right)\right]\right\}$$

yields the decision rule

$$(3) \qquad c_t^* = -Fs_t,$$

and the Riccati equation is

$$(4) \quad P = Q + F^T R F - 2F^T W$$
$$+ \beta\left(A^T - F^T B^T\right) P\left(A - BF\right),$$

where

$$(5) \quad F = \left(R + \beta B^T P B\right)^{-1}\left(W + \beta B^T P A\right).$$

Iterating on the matrix Riccati equation (4) uniquely determines P, as the equation defines a contraction mapping. Using P, we can determine F in the optimal policy (5).

B. *Signal Extraction with Exogenous Noises*

Following the SE literature (e.g., Lucas 1972; Morris and Shin 2002; Muth 1960), we now assume that the agent cannot observe the true state s_t perfectly and only observes the noisy signal $s_t^* = s_t + \xi_t$ when making decisions. Here ξ_t is a $(n \times 1)$ vector of noises. The agent then estimates the state using a standard Kalman filtering equation. In the standard SE problem, the stochastic property of the noise ξ_t is given *exogenously*. Specifically, assume that ξ_t is an iid Gaussian innovation with mean 0 and variance-covariance matrix Λ.[5] We point out here that the agent may not have perfect information even about the endogenous part of the state vector s_t.

Under the LQG assumption, the certainty equivalence principle holds when the agent cannot observe s_t perfectly, so the decision rule under imperfect information can be written as

$$(6) \qquad c_t^* = -F\widehat{s}_t,$$

where $\widehat{s}_t = E[s_t|\mathcal{I}_t]$ is the perceived state and $\mathcal{I}_t = \left\{s_t^*, s_{t-1}^*, \cdots, s_0^*\right\}$ is the information set including perceived signals until time t.

Furthermore, we assume that in the steady state, the true state follows a normal distribution after observing the noisy signals $s_t|\mathcal{I}_t \sim N\left(E[s_t|\mathcal{I}_t], \Sigma_t\right)$, where $\Sigma_t = E_t\left[\left(s_t - \widehat{s}_t\right)\left(s_t - \widehat{s}_t\right)^T\right]$ is the conditional variance-covariance matrix, and the Kalman filtering equation governs the behavior of $\widehat{s}_t$

$$(7) \quad \widehat{s}_{t+1} = (1 - \theta)\left(A\widehat{s}_t + Bc_t\right) + \theta s_{t+1}^*,$$

---

5. Our quadratic objective function encompasses the standard tracking objective of minimizing the squared difference of the control from the target.

where θ is the steady-state Kalman gain to be determined.[6] Following the standard procedure in the Kalman filter literature, we have the steady state updating equation for Σ and Kalman gain θ:

(8)
$$\Sigma = (I - \theta)\left(A\Sigma A^T + \Omega\right)(I - \theta)^T + \theta\Lambda\theta^T$$

and

(9) $\theta = \left(\Omega + A\Sigma A^T\right)\left(\Omega + A\Sigma A^T + \Lambda\right)^{-1}.$

After computing (θ, Σ) by iteration, we can obtain a complete characterization of the dynamic system. The key assumption in the SE problem is that the variance-covariance matrix of the noise, Λ, is given. Given this Λ, Equations (8) and (9) jointly determine the steady state (θ, Σ). It is straightforward to show that we have the following alternative equations for computing the Kalman gain and the conditional variance-covariance matrix, (θ, Σ):

(10) $\Lambda^{-1} = \Sigma^{-1} - \Psi^{-1}.$

and

(11) $\theta = \Sigma\Lambda^{-1},$

where $\Psi = A\Sigma A^T + \Omega$ is the steady-state conditional variance of the state prior to observing the new signal. After obtaining the Kalman gain, Equations (6), (7), and (11) completely characterize the model's dynamic behavior.

III. RATIONAL INATTENTION IN THE LQG MODEL

Following Sims (2003), we introduce RI into the LQG model proposed in Section A by assuming agents face information-processing constraints and have only finite Shannon channel capacity to observe the state of the world. Specifically, we use the concept of entropy from information theory to characterize the uncertainty about a random variable; the reduction in entropy is thus a natural measure of information flow. Formally, entropy is defined as the expectation of the negative of the (natural) log of the density function, $-E\left[\ln\left(f\left(X\right)\right)\right]$. For example, the entropy of a discrete distribution with equal weight on two points is simply $E\left[\ln\left(f\left(X\right)\right)\right] = -0.5\ln\left(0.5\right) - 0.5\ln\left(0.5\right) = 0.69$, and the unit

of information contained in this distribution is 0.69 "nats."[7] In this case, an agent can remove all uncertainty about X if the capacity devoted to monitoring X is $\kappa = 0.69$ nats.

With finite capacity $\kappa \in (0, \infty)$, a variable s following a continuous distribution cannot be observed without error and thus the information set at time $t + 1$, $\mathcal{I}_{t+1}$, is generated by the entire history of noisy signals $\left\{s_j^*\right\}_{j=0}^{t+1}$. Following the literature, we assume the noisy signal takes the additive form $s_{t+1}^* = s_{t+1} + \xi_{t+1}$, where ξ_{t+1} is the endogenous noise caused by finite capacity. We further assume that ξ_{t+1} is an iid idiosyncratic shock and is independent of the fundamental shock. Note that the reason that the RI-induced noise is idiosyncratic is that the endogenous noise arises from the consumer's own internal information-processing constraint. Agents with finite capacity will choose a new signal $s_{t+1}^* \in \mathcal{I}_{t+1} = \left\{s_1^*, s_2^*, \cdots, s_{t+1}^*\right\}$ that reduces the uncertainty of the state variable s_{t+1} as much as possible. Formally, this idea can be described by the information constraint

(12) $\mathcal{H}\left(s_{t+1}|\mathcal{I}_t\right) - \mathcal{H}\left(s_{t+1}|\mathcal{I}_{t+1}\right) \leq \kappa,$

where κ is the investor's information channel capacity, $\mathcal{H}\left(s_{t+1}|\mathcal{I}_t\right)$ denotes the entropy of the state *prior to* observing the new signal at $t + 1$, and $\mathcal{H}\left(s_{t+1}|\mathcal{I}_{t+1}\right)$ is the entropy *after* observing the new signal. κ imposes an upper bound on the amount of information—that is, the change in the entropy—that can be transmitted in any given period. Finally, following the literature, we suppose that the prior distribution of s_{t+1} is Gaussian.

Under the LQG setting, as has been shown in Sims (2003, 2006), the true state under RI also follows a normal distribution $s_t|\mathcal{I}_t \sim N\left(E\left[s_t|\mathcal{I}_t\right], \Sigma_t\right)$, where $\Sigma_t = E_t\left[\left(s_t - \widehat{s}_t\right)\left(s_t - \widehat{s}_t\right)^T\right]$. In addition, given that the noisy signal takes the additive form $s_{t+1}^* = s_{t+1} + \xi_{t+1}$, the noise ξ_{t+1} should also be Gaussian. It is worth noting that the Gaussianity of the posterior variance of the true state and the noise is optimally determined by the LQG structure. This result is often assumed as a matter of convenience in SE models with exogenous noises, and RI can rationalize this assumption. In addition, in the steady state the agent observes an

---

6. Muth (1960) showed that the exponentially weighted average of past observations of a random walk plus a noise process is optimal in the sense that it minimizes the mean squared forecast error.

7. For alternative bases for the logarithm, the unit of information differs; with log base 2 the unit of information is the "bit" and with base 10 it is a "dit" or a "hartley."

additive noisy signal: $s_t^* = s_t + \xi_t$. Note that in the RI problem we also have the usual formula for updating the conditional variance-covariance matrix of a Gaussian distribution Σ in steady state:

(13) $$\Lambda^{-1} = \Sigma^{-1} - \Psi^{-1}.$$

Using these expressions, the Kalman gain θ can also be written as

(14) $$\theta = \Sigma \Lambda^{-1}.$$

A. The Univariate Case

The key difference between SE and RI is that under RI the agent faces the information-processing constraint

(15) $$-\ln(|\Sigma_{t+1}|) + \ln(|A^T \Sigma_t A + \Omega|) \le 2\kappa.$$

As more information about the state is better in single-agent models, this constraint will be binding.[8] Considering the univariate state case $n = 1$, (15) fully determines the value of the steady-state conditional variance Σ:

(16) $$\Sigma = \frac{\Omega}{\exp(2\kappa) - A^2},$$

which means that Σ is determined by the variance of the exogenous shock (Ω) and the exogenously given capacity (κ).[9] Given this Σ, we can use (13) to recover the variance of the endogenous noise (Λ):

(17) $$\Lambda = (\Sigma^{-1} - \Psi^{-1})^{-1},$$

where $\Psi = A^2\Sigma + \Omega$, and use (14) to find the Kalman gain (θ):

(18) $$\theta = \Sigma\Lambda^{-1} = 1 - \Sigma\Psi^{-1},$$

which reduces to $\theta = 1 - \frac{1}{\exp(2\kappa)}$ using Equations (14), (16), and (17). Note that Equations (17) and (18) also hold in the SE problem. To compare the RI and SE problems in the univariate case, we first consider the following two cases.

8. By "better" we mean that conditional on draws by nature for the true state, the expected utility of the agent increases if information about that state is improved.

9. Note that here we need to impose the restriction $\exp(2\kappa) - A^2 > 0$. If this condition fails, the state is not *stabilizable* and the unconditional variance diverges. Obviously we cannot directly impose conditions on A, as it is chosen by the agents in the model; it is also clear that we can, for a given model, easily find sufficient conditions that guarantee the restriction holds.

Case 1. Assume that Λ is fixed exogenously in the SE case.

In Case 1, it is clear that in the SE problem given Λ and Ω, we can compute Σ by solving the nonlinear equation (10). After obtaining Σ, we can use Equation (11) to determine the Kalman gain θ; thus, in this sense SE and RI have the same implications.

We now discuss how to use a policy experiment to distinguish RI from SE. Suppose that the variance of the exogenous shock, Ω, is scaled up because of a change in policy. In the SE problem with fixed Λ, Equations (17) and (18) imply that an increase in Ω will generally lead to a different solution for Σ and θ; consequently, the change in policy will lead to a change in the model's dynamics. Because Σ is a nonlinear function of Ω, the effect of changes in Ω on Σ could be complicated. In the next section, we will explore this relationship using some numerical examples in a permanent income model. In contrast, in the RI problem, if κ is *fixed*, Equations (16), (17), and (18) imply that a change in Ω will lead to the same change in Σ, Ψ, and Λ, but has no impact on θ. In other words, agents with fixed capacity will behave as if facing noise whose nature changes systematically as the dynamic properties of the economy change, that is, the change in policy does not change the model's dynamics.

Case 2. Assume that the SNR (SNR), $\Omega\Lambda^{-1}$, is fixed exogenously in the SE case.

Note that Equation (17) can be rewritten as

(19)
$$\Omega\Lambda^{-1} = \Omega\Sigma^{-1} - \left[A^2\left(\Omega\Sigma^{-1}\right)^{-1} + 1\right]^{-1}.$$

In Case 2, since the SNR is fixed, (19) can be used to solve for $\Omega\Sigma^{-1}$. Given the SNR and $\Omega\Sigma^{-1}$, we can compute

(20) $$\Sigma\Lambda^{-1} = (\Sigma\Omega^{-1})(\Omega\Lambda^{-1}).$$

Consider the same case in which Ω is scaled up, fixed SNR means that the exogenous noise should also be scaled up such that $\Omega\Lambda^{-1}$ is fixed at the same level; consequently, Equation (19) leads to the same solution for $\Omega\Sigma^{-1}$ and Equation (20) leads to the same $\Sigma\Lambda^{-1}$. The following proposition summarizes the main conclusion in this case:

PROPOSITION 1. *In Case 2 (i.e., the SNR is fixed), the SE and RI problems are observationally equivalent in the sense that they lead to the same dynamics if κ is fixed and $\Sigma\Lambda^{-1}$ in the*

SE problem is equal to $1 - 1/\exp(2\kappa)$ in the RI problem.

Proof. The proof is straightforward by comparing Equations (18) and (20).   ∎

In the above analysis, for simplicity we assume that κ remains unchanged when Ω is altered by the government policy. However, if an increase in Ω leads to higher marginal welfare losses due to imperfect observations, some capacity may be reallocated from other sources to reduce the welfare losses due to low capacity. In this case, θ will change accordingly as it is completely determined by capacity κ; consequently, the dynamic behavior of the model will also change in response to the change in Ω. We will further explore this issue in the next section and the permanent income model in Section 4.

Alternative Way to Model Limited Information-Processing Capacity. As argued by Sims (2010), instead of using fixed finite channel capacity to model limited information-processing ability, it is also reasonable to assume that the marginal cost of information processing is constant. That is, the Lagrange multiplier on Equation (15) is constant.[10] In the univariate case, if the decision rule under full information is $c_t^* = H s_t$ and the objective of the agent with finite capacity is to minimize $\sum_{t=0}^{\infty} \beta^t (c_t - c_t^*)^2$, the optimization problem reduces to

$$\min_{\Sigma_t} \sum_{t=0}^{\infty} \beta^t \left[H^2 \Sigma_t + \lambda \ln \left(\frac{A^2 \Sigma_{t-1} + \Omega}{\Sigma_t} \right) \right],$$

where Σ_t is the conditional variance at t, λ is the Lagrange multiplier corresponding to Equation (15), and we impose the restriction that $\beta A = 1$ and $H = A - 1$ for simplicity. Solving this problem yields the optimal steady-state conditional variance:

(21)
$$\Sigma = \frac{-(\Omega H - \lambda A) + \sqrt{(\Omega H - \lambda A)^2 + 4\lambda \Omega A^2}}{2 H A^2} > 0.$$

It is straightforward to show that as λ goes to 0, $\Sigma = 0$; and as λ goes to ∞, $\Sigma = \infty$. Comparing Equations (21) with (16), it is clear that

the two modeling strategies are *observationally equivalent* in the sense that they lead to the same conditional variance if the following equality holds:

(22)
$$\kappa = \frac{1}{2} \ln \left(A^2 + \frac{2HA^2}{\left(\begin{array}{c} -[H - A(\lambda/\Omega)] \\ + \sqrt{\begin{array}{c} [H - A(\lambda/\Omega)]^2 \\ + 4A^2(\lambda/\Omega) \end{array}} \end{array} \right)} \right).$$

In this case, the Kalman gain is

(23)
$$\theta = 1 - \left\{ A^2 + \frac{2HA^2}{\left(\begin{array}{c} -[H - A(\lambda/\Omega)] \\ + \sqrt{\begin{array}{c} [H - A(\lambda/\Omega)]^2 \\ + 4A^2(\lambda/\Omega) \end{array}} \end{array} \right)} \right\}^{-1}.$$

It is obvious that κ converges to its lower limit $\underline{\kappa} = \ln(A)$ as λ goes to ∞; and it converges to ∞ as λ goes to 0.[11] In other words, using the RI modeling strategy, the agent is allowed to adjust the optimal level of capacity in such a way that the marginal cost of information processing for the problem at hand remains constant. Note that this result is consistent with the concept of "elastic" capacity proposed in Kahneman (1973).

Furthermore, it is clear from Equation (22) that if the cost of information processing (λ) is fixed, an increase in fundamental uncertainty (Ω) will lead to higher capacity (κ) devoted to monitoring the evolution of the state. We now consider the same policy experiment discussed above: Ω is scaled up because of a change in policy. If we adopt the assumption that λ is *fixed*, Equation (21) means that there is a less change in Σ because $\frac{\partial \ln \Sigma}{\partial \ln \Omega} < 1$. Note that in the fixed κ case, $\frac{\partial \ln \Sigma}{\partial \ln \Omega} = 1$. Consequently, a change in Ω will change θ and the model's dynamics

---

10. Formally, the assumption is that κ is a choice variable and the utility cost function is $\mu\kappa$ for some constant μ.

11. We require here that $H \neq 0$; that is, the state must be *detectable*. If the state is not detectable there is no point in allocating attention to monitoring it.

if the inattentive agent is facing fixed marginal cost of information. Therefore, different ways of modeling RI may lead to different policy implications.[12]

Extension to Correlated Shocks and Noises. In the above analysis, we assumed that the exogenous fundamental shock and noise are uncorrelated. We now discuss how *correlated* shocks and noises affect the implications of SE and RI for the model's dynamic behavior. In real systems, we may observe correlated shocks and noises. For example, if the system is an airplane and winds are buffeting the plane, the random gusts of wind affect both the process (the airplane dynamics) and the measurement (the sensed wind speed) if people use an anemometer to measure wind speed as an input to the Kalman filter. Please see Simon (2006) for a discussion. To use an economic example, suppose there is a correlation between levels and volatility, and also that producing national income data is costly; then a transition from low to high volatility would be associated with low output levels and, at the margin, less accurate measurements.

It is straightforward to introduce correlated shocks and noises into the SE problem. Specifically, we consider the case in which the process shock (ε) and the noise (ξ) are correlated as follows:

$$\text{corr}\,(\varepsilon_{t+1}, \xi_{t+1}) = \rho,$$

$$\text{cov}\,(\varepsilon_{t+1}, \xi_{t+1}) = \Gamma = \rho\sqrt{\Omega}\sqrt{\Lambda},$$

where ρ is the correlation coefficient between ε_{t+1} and ξ_{t+1}, $\Omega = \text{var}\left[\varepsilon_{t+1}\right]$ and $\Lambda = \text{var}\left[\xi_{t+1}\right]$. Under SE, Λ is given exogenously and the correlation just introduces another exogenous stochastic dimension on the noise. As shown in Simon (2006), in this case the optimal Kalman gain can be written as

$$(24) \qquad \theta = (\Psi + \Gamma)\,(\Psi + \Lambda + 2\Gamma)^{-1},$$

and the updating formula for the conditional variance is

$$(25) \quad \Sigma = \Psi - (\Psi + \Gamma)^2\,(\Psi + \Lambda + 2\Gamma)^{-1},$$

where $\Psi = \Omega + A^2\Sigma$. Just like the case without the correlation, given Λ and Γ, Equations (24) and (25) jointly determine the steady state (θ, Σ).

In the RI problem, the correlation generalizes the assumption in Sims (2003) on the uncorrelated RI-induced noise. In the steady state, Equation (25) can be rewritten as the following quadratic equation in terms of $\sqrt{\Lambda}$: $\left[\rho^2\Omega - (\Psi - \Sigma)\right]\Lambda + 2\rho\Sigma\sqrt{\Omega}\sqrt{\Lambda} + \Sigma\Psi = 0$, which can be solved for

$$(26)$$

$$\Lambda = \left(\frac{-\rho\Sigma\sqrt{\Omega} + \sqrt{\rho^2\Sigma^2\Omega - \Sigma\Psi\left[\rho^2\Omega - (\Psi - \Sigma)\right]}}{\rho^2\Omega - (\Psi - \Sigma)}\right)^2.$$

It is clear from (26) that if κ is fixed, the change in Ω will lead to the same change in Σ, Ψ, and Λ, but has *no* effect on the Kalman gain $\theta = \Sigma\Lambda^{-1}$. That is, the presence of correlated noise does not change the dynamic behavior of the model.

Consider the RI problem with a fixed information-processing cost (λ). From Equations (21) and (26), it is clear that in the presence of correlated noise ($\rho > 0$), there is a less change in Σ when there is a change in Ω because $\frac{\partial \ln \Sigma}{\partial \ln \Omega} < 1$, and consequently, the change in Ω will also change θ because Λ depends on the interactions between Ω and Σ.

B. The Multivariate Case

In the multivariate RI problem, it is more difficult to determine the steady-state conditional variance-covariance matrix Σ because it cannot be computed analytically. Here we follow Sims (2003) and calculate the expected loss due to imperfect observations under RI. Specifically, we assume that the value functions under full information and imperfect information can be written as

$$v\,(s_t) = s_t^T P s_t \text{ and } \widehat{v}\,(\widehat{s}_t) = \widehat{s}_t^T \widehat{P}\widehat{s}_t,$$

respectively.[13] We can compute the optimal Σ by minimizing the expected welfare loss due to RI,

$$(27) \qquad E_t\left[v\,(s_t) - \widehat{v}\,(\widehat{s}_t)\right],$$

subject to information-processing constraints. Note that to solve this problem numerically, we need to use a two-stage procedure.[14] First, under the LQG assumption, the certainty equivalence principle applies and the decision rule under imperfect information,

$$(28) \qquad c_t^* = -F\widehat{s}_t,$$

is independent of Σ or Λ. We then use this decision rule to determine $\widehat{v}\,(\widehat{s}_t)$ which depends on Σ and Λ. Applying the welfare criterion proposed in Equation (27), we can solve for optimal steady-state Σ and Λ.

Solving the problem posed in Equation (27) is equivalent to solving the semidefinite programming problem

$$(29) \qquad \max_{\Sigma} \{\text{trace}\,(-Z\Sigma)\}$$

subject to

$$(30) \quad -\ln(|\Sigma|) + \ln(|A^T\Sigma A + \Omega|) \le 2\kappa,$$

$$(31) \qquad\qquad A^T\Sigma A + \Omega \succeq \Sigma,$$

where $Z = F^T R F - 2F^T W + \beta(F^T B^T P B F + F^T B^T P A + A^T P B F)$ (see Appendix S1, Supporting Information, for the derivation). If the positive-definiteness constraint on $A^T\Sigma A + \Omega - \Sigma$, Equation (31), does not bind, the first-order condition for Σ can be written as

$$(32) \qquad \Sigma^{-1} = \left(G\Sigma G^T + G_0\right)^{-1} - \frac{Z}{\lambda},$$

where $G = \left(A^T\right)^{-1} A$ and $G_0 = \left(A^T\right)^{-1}\Omega A^{-1}$. We can then use standard methods to solve Equation (32). When applied to a permanent income model in the next section, we first solve this equation and then check whether in fact Equation (31) is satisfied by the optimal solution of Σ. If so, the problem is solved.[15]

After computing the optimal steady-state Σ, we can then use Equation (13) to determine the steady-state Λ and Equation (14) to determine the Kalman gain θ. Therefore, the key difference between SE and RI is that in the SE problem we need to specify the process of the noise first, whereas in the RI problem we need to first specify the value of channel capacity that determines the steady-state conditional

variance of the state by solving the semidefinite programming problem proposed in Equation (29) subject to Equations (30) and (31).[16] Theoretically, it is clear that after solving an RI problem, we can always reconstruct an SE problem using the resulting endogenous noise due to RI as the input, and the two models are observationally equivalent in this sense. However, it is difficult to specify the process of the vector of noises ex ante when modeling the multivariate SE problem.[17]

When modeling the multivariate RI problem we only need to set a value for channel capacity and then compute optimal conditional variance-covariance matrices of the state and the variance-covariance matrices of the noise vector by solving the constrained semidefinite minimization problem (29). Therefore, in the multivariate RI problem, the agent's preference, budget constraint, and information-processing constraints jointly determine the values of Σ, Λ, and θ, whereas in the multivariate SE problem given Λ, Equation (13) that is used to determine Σ and θ only depends on the budget constraint. If the noise in SE is specified exogenously, it may violate the optimality conditions for RI; for example, Melosi (2009) showed that a particular estimated SE model does not equate the marginal utility of attention across states, implying that the variance-covariance matrix of the noise would not be consistent with *any* channel capacity. Of course, obtaining the marginal utility of attention requires solving the RI problem, so it will be difficult to specify ex ante an SE problem consistent with RI.[18]

We now consider the different policy effects of RI and SE in the multivariate case. We first assume that initially the SE and RI problems have the same Kalman gain that generates the same dynamic behavior. Suppose that the variance-covariance matrix of the exogenous shock, Ω, is scaled up because of a change in policy.[19] In the SE problem with fixed Λ,

14. Sims (2010) also applied this principle to solve a tracking problem with information constraints.

15. If this constraint does not bind, in principle we can apply the logic of the reverse water-filling problem to solve for Σ.

16. Note that the basic idea of solving the multivariate RI problem is the same as that in the univariate model and thus the key difference between SE and RI problems remains unchanged.

17. This problem will be particularly difficult for non-LQG problems, as the distribution of the noise shocks will generally be impossible to specify analytically.

18. The robustness of the result in Melosi (2009) is unknown, but it clearly depends on model assumptions to some degree.

19. That is, all elements in the variance-covariance matrix are scaled up.

Equations (17) and (18) imply that a change of Ω will lead to a different solution for Σ and θ, that is, the change in policy will lead to a change in the model's dynamics. In contrast, in the multivariate RI problem, as shown in Equations (29)–(31), a change in Ω will have complicated effects on Σ, Λ, and θ. In other words, in the multivariate case a change in policy will affect the model's behavior in both SE and RI problems. (Note that in the univariate case the change in policy does not change the model's dynamics.)

We next consider the effects of RI and SE in Case 2 (i.e., the SNR, $\Omega\Lambda^{-1}$, is fixed in the SE problem). As before, we assume that initially the SE and RI problems have the same Kalman gain. To illustrate how a change in Ω affects the Kalman gain in RI and SE problems, we multiply Σ on both sides of Equation (17):

$$(33) \quad \Sigma\Lambda^{-1} = I - \left[A\Sigma A^T\Sigma^{-1} \right.$$
$$\left. + \left(\Omega\Lambda^{-1}\right)\left(\Lambda\Sigma^{-1}\right) \right]^{-1},$$

where I is the identity matrix and we use the fact that $\Omega\Sigma^{-1} = \left(\Omega\Lambda^{-1}\right)\left(\Lambda\Sigma^{-1}\right)$. In Case 2, the policy has the same impact on Ω and Λ to keep the SNR fixed. Equation (33) clearly shows that if the policy changes Σ and then $A\Sigma A^T\Sigma^{-1}$, it will affect $\theta = \Sigma\Lambda^{-1}$ even in Case 2. Multiplying Ω on both sides of (17) gives

$$(34) \quad \Omega\Lambda^{-1} = \Omega\Sigma^{-1} - \left(A\Sigma A^T\Omega^{-1} + I\right)^{-1},$$

which means that a change in Ω will lead to different Σ given that $\Omega\Lambda^{-1}$ is fixed. Note that in the univariate case, $A\Sigma A^T\Sigma^{-1} = A^2$, which means that the policy has no impact on θ, and the SE and RI problems cannot be distinguished by the policy in Case 2 that the SNR, $\Omega\Lambda^{-1}$, is fixed.

IV. APPLICATIONS TO MODELS OF PERMANENT INCOME AND PRICE SETTING

In this section, we consider the effects of SE and RI for consumption dynamics and their policy and welfare implications in an otherwise standard permanent income model. As in the previous section we first consider applications to the univariate case and then discuss applications to the multivariate case.

We are not the first to examine this model. Sims (2003) examined how RI affects consumption dynamics when the agent only has limited capacity when processing information. Luo (2008) showed that the RI permanent income can be solved explicitly even if the income process is not iid, and then examined how RI can resolve the well-known excess smoothness and excess sensitivity puzzles; that model admits a reduction to a single state variable.[20] To explore the attention allocation problem explicitly, we also consider a simplified price-setting model used in Woodford (2001) and Maćkowiak and Wiederholt (2009), and argue that this model can also help distinguish RI from SE.

A. The Univariate Permanent Income Model

Optimal Consumption. Following Luo (2008), we have the subsequent univariate version of the standard permanent income model (Flavin 1981; Hall 1978) in which households solve the dynamic consumption-savings problem

$$(35) \quad v(s_0) = \max_{\{c_t\}} E_0\left[\sum_{t=0}^{\infty} \beta^t u(c_t) \right]$$

subject to

$$(36) \quad s_{t+1} = Rs_t - c_t + \zeta_{t+1},$$

where $u(c_t) = -\frac{1}{2}\left(\overline{c} - c_t\right)^2$ is the period utility function, $\overline{c} > 0$ is the bliss point, c_t is consumption,

$$(37) \quad s_t = w_t + \frac{1}{R}\sum_{j=0}^{\infty} R^{-j}E_t\left[y_{t+j}\right]$$

is permanent income (the expected present value of lifetime resources), consisting of financial wealth (w_t) plus human wealth (i.e., the expected discounted present value of current and future labor income, y),

(38)

$$\zeta_{t+1} = \frac{1}{R}\sum_{j=t+1}^{\infty}\left(\frac{1}{R}\right)^{j-(t+1)}\left(E_{t+1} - E_t\right)\left[y_j\right]$$

is the time $(t+1)$ innovation to permanent income with mean 0 and variance ω_ζ^2, w_t is cash-on-hand (or market resources), y_t is a general income process with Gaussian white noise innovations, β is the discount factor, and R is the constant gross interest rate at which the

---

20. The excess smoothness puzzle states that consumption responds too little to permanent changes in income. The excess sensitivity puzzle states that current consumption responds to changes in income that were anticipated in earlier periods.

consumer can borrow and lend freely.[21] We assume y follows an AR(1) process with persistence coefficient $\rho \in [0, 1]$, $y_{t+1} = \rho y_t + \varepsilon_{t+1}$, where $\varepsilon_{t+1} \sim N\left(0, \omega^2\right)$, $s_t = w_t + y_t/(R - \rho)$ and $\zeta_{t+1} = \varepsilon_{t+1}/(R - \rho)$.[22] For the rest of the paper we will restrict attention to points where $c_t < \overline{c}$, so that utility is increasing and concave; following the literature we impose the restriction that $\beta R = 1$, because it implies a stationary path for consumption. This specification follows that in Hall (1978) and Flavin (1981) and implies that optimal consumption is determined solely by permanent income:

$$(39) \qquad c_t = (R - 1) s_t.$$

Within this LQG setting, the certainty equivalence principle holds and introducing SE or RI lead to the following new consumption function:

$$(40) \qquad c_t = (R - 1) \widehat{s}_t,$$

where $\widehat{s}_t = E_t [s_t]$ is the perceived state and is governed by the following Kalman filtering equation

$$(41) \qquad \widehat{s}_{t+1} = (1 - \theta)(R\widehat{s}_t - c_t) + \theta(s_{t+1} + \xi_{t+1}),$$

where θ is the Kalman gain, and given $s_0 \sim N(\widehat{s}_0, \Sigma_0)$. As shown in Luo (2008), combining Equations (36), (40), with (41) yields the following expression for the change in consumption:

$$(42) \quad \Delta c_t = (R - 1)\left[\frac{\theta \zeta_t}{1 - (1 - \theta)R \cdot L} + \theta\left(\xi_t - \frac{\theta R \xi_{t-1}}{1 - (1 - \theta)R \cdot L}\right)\right],$$

where L is the lag operator. We require $(1 - \theta)R^2 < 1$, the model equivalent of the stabilizability condition stated before (this condition implies $(1 - \theta)R < 1$ since $R > 1$). This MA(∞) process shows that the dynamic behavior of the model is strongly influenced by the Kalman gain θ. Using the explicit expression for consumption growth (42), we can compute the key stochastic properties of consumption

21. Note that in this case, the flow budget constraint is $w_{t+1} = Rw_t - c_t + y_t$.

22. Here we assume that disposable income follows a persistent trend-stationary AR(1) process; Luo, Nie, and Young (2012) estimated the process for small-open economies. It is well known that given the length and structure of the data on real income, it is difficult to distinguish persistent trend-stationary AR(1), unit root, and difference-stationary (DS) processes for real income. We focus on the AR(1) case in this paper.

process: the volatility of consumption growth, the persistence of consumption growth, and the correlation between consumption growth and income shocks.[23] All these moments depend on the Kalman gain. In other words, SE and RI lead to different consumption processes if and only if the resulting θ differs.

It is worth noting that the representative-agent RI model can be regarded as an extreme case of an RI model with a continuum of agents. For example, consider two PIH models: one with a representative consumer and one with a continuum of consumers endowed with identical channel capacity. While the effect of RI on the reaction of consumption to income shocks is the same in the two PIH environments, the volatility of aggregate consumption depends on the specifics of the aggregative mechanism. As argued in Sims (2003), although the randomness in an agent's response to aggregate shocks will be idiosyncratic because it arises from the agent's own information-processing channel, there is likely a significant common component. Therefore, the common term of the noise lies between 0 and the part of the RI-induced noise, ξ_t. If we assume that ξ_t consists of two independent noises: $\xi_t = \overline{\xi}_t + \xi_t^i$, where $\overline{\xi}_t = E^i[\xi_t]$ and ξ_t^i are the common and idiosyncratic components of the error generated by ζ_t, respectively. A single parameter, $\lambda = \frac{\text{var}(\overline{\xi}_t)}{\text{var}(\xi_t)} \in [0, 1]$, can be used to measure the common source of coded information on the aggregate component (or the relative importance of $\overline{\xi}_t$ vs. ξ_t). (Sims 2003 argued that the presence of the common noise can help fit the model with the VAR evidence better.) Given the existing theories on this issue, we cannot pin down the value of λ. Therefore, the $\lambda = 1$ case can be viewed as a special case of the model with a continuum of consumers which is equivalent with the representative-agent model in which we do not need to discuss the aggregation issue.

Policy Implications under SE and RI. In this univariate permanent income model, substituting $A = R$ into Equation (19),

$$(43) \qquad \omega_\zeta^2 \Lambda^{-1} = \omega_\zeta^2 \Sigma^{-1} - \left[R^2\left(\omega_\zeta^2 \Sigma^{-1}\right)^{-1} + 1\right]^{-1},$$

23. See Luo (2008) for a discussion on the effects of RI on consumption dynamics.
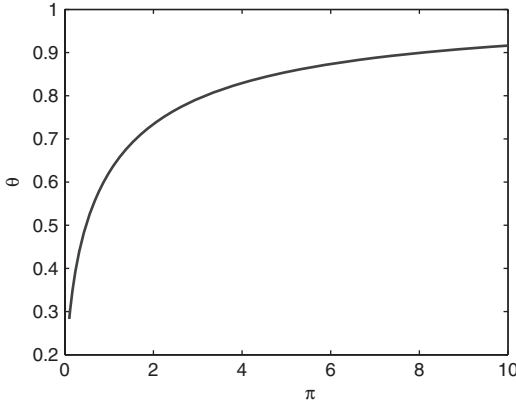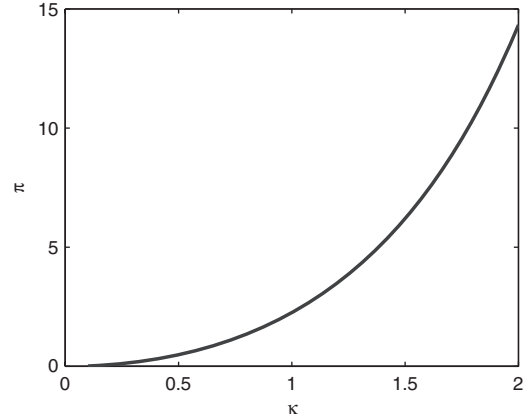
FIGURE 1

Relationship between π and θ

FIGURE 2

Relationship between κ and π

where $\omega_\zeta^2 \Sigma^{-1} = \left(\omega_\zeta^2 \Lambda^{-1}\right)\left(\Lambda \Sigma^{-1}\right)$, and denote $\theta = \Sigma \Lambda^{-1}$ and $\pi = \omega_\zeta^2 \Lambda^{-1}$, we obtain the following relation between π and θ:

$$(44) \qquad \pi = \theta \left(\frac{1}{1-\theta} - R^2\right).$$

Solving for θ yields

$$(45) \qquad \theta = \frac{\left(\begin{array}{c} -\left(1 + \pi - R^2\right) \\ +\sqrt{\left(1 + \pi - R^2\right)^2 + 4R^2\pi} \end{array}\right)}{2R^2},$$

where we omit the negative root of θ as both Σ and Λ must be positive. Figure 1 illustrates the relationship between π and θ given $R = 1.02$ and $\pi \in [0.1, 10]$. It clearly shows that θ is an increasing function of π, the signal-to-noise ratio. Also, as $\pi \to \infty$, $\theta \to 1$.

In the RI version of the permanent income model, we have

$$(46) \qquad \Sigma = \frac{\Omega}{\exp(2\kappa) - R^2},$$

$$(47) \qquad \Lambda = \left(\Sigma^{-1} - \Psi^{-1}\right)^{-1},$$

where $\Psi = R^2 \Sigma + \Omega$. Using (46) and (47), the Kalman filter gain under RI can be written as

$$(48) \qquad \theta = \Sigma \Lambda^{-1} = 1 - \frac{1}{\exp(2\kappa)}.$$

Comparing Equations (45) with (48), it is clear that the signal-to-noise ratio (π) and the level of channel capacity (κ) have one-to-one correspondence. Figure 2 shows the relationship between κ and π when the SE and RI

problems are observationally equivalent in the sense that they lead to the same consumption dynamics governed by the Kalman gain θ, that is, $\pi = (1 - \exp(-2\kappa))\left(\exp(2\kappa) - R^2\right)$. This result is consistent with the general conclusion we obtained using Case 2 in the previous section.

Using the same expression for θ, (45), we can examine how Case 1 can be used to distinguish SE and RI when implementing a change in government policy. Specifically, in the SE problem, we assume that before the government implements stabilization policies, the SNR $\pi = \omega_\zeta^2/\Lambda = 2$. In this case, $\theta = 0.73$. After the government implements these policies, the variance of the shock to permanent income will be reduced from ω_ζ^2 to $0.5\omega_\zeta^2$. As Λ is fixed in Case 1, π will fall from 2 to 1; consequently, $\theta = 0.62$. We now assume that the RI and SE problems are observationally equivalent in the sense that they lead to the same $\theta = 0.73$ before implementing the stabilization policies. After implementing these policies, ω_ζ^2 will be scaled down to $0.5\omega_\zeta^2$, and the RI theory predicts that both Σ and Λ will be scaled down to 0.5Σ and 0.5Λ, respectively.[24] Consequently, the Kalman filter gain, $\theta = \Sigma \Lambda^{-1} = 0.73$, remains unchanged. In other words, stabilization policy has different implications for consumption dynamics in the SE and RI models.

Alternatively, if we assume that the cost of information processing (λ) is fixed, the optimal

24. A proof is straightforward from Expressions (46) and (47).

conditional variance equals

(49)

$$\Sigma = \frac{\begin{pmatrix} -\left[\Omega\left(R-1\right)-\lambda R\right] \\ +\sqrt{\left[\Omega\left(R-1\right)-\lambda R\right]^2 + 4\lambda\Omega R^2} \end{pmatrix}}{2\left(R-1\right)R^2}.$$

Comparing Equations (49) with (46), it is clear that the two modeling strategies are *observationally equivalent* in the sense that they lead to the same conditional variance if the following equality holds:

(50)

$$\kappa = \frac{1}{2}\ln\left(R^2 + \frac{2\left(R-1\right)R^2}{\begin{pmatrix} -\left[(R-1)-R\left(\lambda/\omega_\zeta^2\right)\right] \\ +\sqrt{\begin{array}{c}\left[(R-1)-R\left(\lambda/\omega_\zeta^2\right)\right]^2 \\ +4R^2\left(\lambda/\omega_\zeta^2\right)\end{array}} \end{pmatrix}}\right).$$

In this case, the Kalman gain is

(51)

$$\theta = 1 - \left\{R^2 + \frac{2\left(R-1\right)R^2}{\begin{pmatrix} -\left[(R-1)-R\left(\lambda/\omega_\zeta^2\right)\right] \\ +\sqrt{\begin{array}{c}\left[(R-1)-R\left(\lambda/\omega_\zeta^2\right)\right]^2 \\ +4R^2\left(\lambda/\omega_\zeta^2\right)\end{array}} \end{pmatrix}}\right\}^{-1}.$$

After implementing these policies, ω_ζ^2 is scaled down to $0.5\omega_\zeta^2$ (i.e., the economy switches to a more stable environment) and the fixed λ theory predicts that the Kalman filter gain, $\theta = \Sigma\Lambda^{-1}$, is reduced. For example, before the government implements stabilization policies, we have $\lambda/\omega_\zeta^2 = 0.000135$ and $\theta = 0.79$. After the policy, we can easily calculate that $\theta = 0.68$ using Equation (51). Figure 3 plots the different implications of SE and RI for consumption dynamics after implementing the stabilization policy: consumption growth falls more (less) under SE than RI when κ is fixed (when λ is fixed), since the Kalman gain decreases (increases). The intuition behind this result is simple. In the fixed λ case some capacity will be reallocated to other sources because a reduction in macroeconomic uncertainty leads to smaller welfare losses due to RI.

FIGURE 3

Consumption Dynamics under SE and RI after Policy
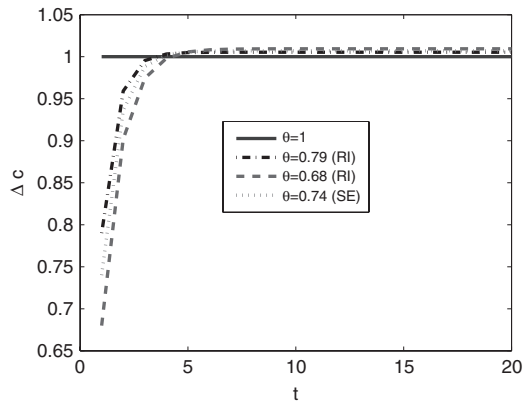
Implications for Consumption Volatility in Emerging and Developed Countries. Aguiar and Gopinath (2007) consider two groups of small economies (emerging and developed economies), and find that consumption is more volatile than income at business cycle frequencies for emerging markets, as compared to a ratio of less than one for developed markets. They show that a SOE real business cycle (RBC) model driven primarily by shocks to trend growth can explain well this regularity about the relative volatility of consumption to income in emerging markets. However, by using a long time series data over 1900–2005, García-Cicco, Pancrazi, and Uribe (2010) estimate an RBC model driven by the same shocks considered in Aguiar and Gopinath (2007), and find that the model does a poor job explaining the observed business cycle fluctuations in Argentina and Mexico along a number of dimensions including the relative volatility of consumption to income. In this section, we briefly show that the RI model with a fixed information-processing cost can explain the observed difference in consumption volatility in emerging and developed countries.

The PIH model presented in Section A can be regarded as an SOE model in which the constant interest rate is given exogenously and there are a continuum of consumers in the model economy. Using Equation (42) and assuming that all idiosyncratic errors are canceled out after aggregation, the relative volatility of consumption growth to income growth can be written

as

$$(52) \quad \mu \equiv \frac{\text{sd}(\Delta c_t)}{\text{sd}(\Delta y_t)} = \left(\frac{R-1}{R-\rho}\sqrt{\frac{1+\rho}{2}}\right)$$
$$\times \sqrt{\frac{\theta^2}{1-((1-\theta)R)^2}},$$

where $\text{sd}(\cdot)$ denotes standard deviation. It is straightforward to show that the relative consumption volatility is decreasing with the degree of imperfect state observations because $\partial\mu/\partial\theta < 0$. As shown in Luo, Nie, and Young (2012), if there is no imperfect-state-observation assumption (i.e., $\theta = 1$), the model cannot generate the empirical relative consumption volatility. For example, if $R = 1.04$, the full information model predicts that μ in emerging and developed economies would be 0.28 and 0.24, respectively. In contrast, in the data, the corresponding μ values are 1.35 and 0.98, respectively.[25] In the RI model with fixed capacity, θ is uniquely determined by fixed capacity κ and thus has no impact on the cross-country comparison if emerging and developed countries have the same average amount of channel capacity. In contrast, if we adopt the fixed information-processing cost assumption, Equations (51) and (52) can have the potential to generate the observed difference in consumption volatility in emerging and developed countries because θ is an increasing function of income uncertainty and income uncertainty in emerging countries is much higher than that in developed countries ($\text{sd}(\Delta y)/\text{mean}(y)$ is 3.82 in emerging countries, while it is 2.07 in developed countries). Intuitively, in developed countries consumers pay less attention to macroeconomic conditions because the fundamental uncertainty is low; consequently, the aggregate consumption process in these countries is more stable relative to the income process.[26]

General Equilibrium Asset Pricing Implications.
The PIH model presented in Section A is usually regarded as a partial equilibrium model. However, as noted in Hansen (1987) and Cochrane

25. See Table 1 in Luo, Nie, and Young (2012) for the estimated income processes in both emerging and developed countries.

26. It is worth noting that this prediction can also be used to distinguish SE and RI by examining the time-series behavior of an economy across "regime changes," such as the observed U.S. Great Moderation in which the volatility of output dropped after 1984. This rationale was also used in Sargent (1976) to discuss the observational equivalence between Classical and Keynesian models.

(2005), it is not a partial equilibrium result—it is a general equilibrium model with a linear production technology and an exogenous income process. Given the expression of optimal consumption in terms of the state variables derived from the PIH model with imperfect-state-observation, we can price assets by treating the process of aggregate consumption that solves the model as though it was an endowment process. In this setup, equilibrium prices are shadow prices that leave the agent content with that endowment process.

In the model setting specified in Section A, w can be regarded as capital. R can be regarded as the return on the linear technology and is not yet the interest rate (the equilibrium rate of return on one-period claims to consumption). As proposed in Cochrane (2005) and used in Luo and Young (2010b), after finding optimal consumption as in Equation (40), we can price one-period claims using this equilibrium consumption stream. Denoting the risk-free rate by R^f, we have the following Euler equation:

$$\frac{1}{R^f} \equiv E_t\left[\beta\frac{u'(c_{t+1})}{u'(c_t)}\right] = \beta E_t\left[\frac{\bar{c}-c_{t+1}}{\bar{c}-c_t}\right]$$
$$= \beta = \frac{1}{R},$$

where $E_t[\cdot]$ is the consumer's expectation operator conditional on his/her processed information at time t. We can now use the basic pricing equation, $p = E[mx]$,[27] to compute the price of the stream of aggregate consumption (treated as the stream of endowments) as[28]

$$(53) \quad p_t = E_t\left[\sum_{j=1}^{\infty}\left(m_{t,t+j}c_{t+j}\right)\right]$$
$$= \underbrace{\frac{1}{R-1}c_t}_{p_t^{rn}} - \underbrace{\frac{1}{\bar{c}-c_t}\Xi}_{p_t^{ra}},$$

where $m_{t,t+j} \equiv \beta^j\frac{u'(c_{t+j})}{u'(c_t)}$ is the stochastic discount factor, and $\Xi \equiv \sum_{j=1}^{\infty}\left(\beta^j\text{var}_t[c_{t+j}]\right) = \frac{(2-\theta)R}{1-R^2(1-\theta)}\omega_\zeta^2$. Denoting the risk-neutral component by p_t^{rn} and the risk-adjusted component by

27. Note that we know $E[mx]$ after solving the PIH model given the state variables and can use them to determine the asset price p.

28. For the details of the derivation, see Luo and Young (2010b).

p_t^{ra}, we have

(54)
$$p_t^{rn} = \frac{1}{R-1} c_t$$

and

(55)
$$p_t^{ra} = \frac{1}{\bar{c} - c_t} \frac{(2-\theta)R}{1 - R^2(1-\theta)} \omega_\zeta^2.$$

Equation (53) yields the following implications. The first term in Equation (53) is the risk-neutral component denoted by p_t^{rn}. This term can be regarded as the value of a perpetuity paying c_t. The second term is the risk-adjusted component, p_t^{ra}; it lowers the asset price relative to the risk-neutral level because $c_t \leq \bar{c}$ and it is decreasing with the degree of attention (θ).

From Equations (53) and (55), it is clear that the Kalman gain (θ) also plays a key role in determining the general equilibrium asset prices under both SE and RI. Consider the same policy experiment we discussed in the last section. If the economy switches to a more stable environment as a result of stabilization policy (i.e., ω_ζ^2 is scaled down to $0.5\omega_\zeta^2$), the fixed λ assumption predicts that the Kalman filter gain, $\theta = \Sigma\Lambda^{-1}$, is reduced. Given that $\lambda/\omega_\zeta^2 = 0.000135$ and $\theta = 0.79$, we can easily calculate that $\theta = 0.68$ using Equation (51). Equation (55) clearly shows that asset price falls less (more) under SE than RI when κ is fixed (when λ is fixed), since the Kalman gain decreases (increases).

B. The Multivariate Permanent Income Model

In this section, we solve for optimal steady-state Σ and Λ in a parametric multivariate RI permanent income model and then illustrate the differences between RI and SE problems. This example is similar to that discussed in Sims (2003) and considers multiple income shocks with different stochastic properties. Specifically, we assume that the original budget constraint is as follows

(56)
$$w_{t+1} = Rw_t - c_t + y_{t+1},$$

where w_t is the amount of cash-in-hand, and the income process y_t have two persistent components (x and z) and one transitory component ($\varepsilon_{y,t}$):

$$y_t = \bar{y} + x_t + z_t + \varepsilon_{y,t},$$
$$x_t = 0.99x_{t-1} + \varepsilon_{x,t},$$
$$z_t = 0.95z_{t-1} + \varepsilon_{z,t},$$

with

(57)
$$\Omega = \text{var}\begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \\ \varepsilon_{z,t} \end{bmatrix}$$

$$= 10^{-3}\begin{bmatrix} 0.9 & 0 & 0 \\ 0 & 0.009 & 0 \\ 0 & 0 & 0.27 \end{bmatrix},$$

where x_t is the most persistent and smooth component and $\varepsilon_{y,t}$ is the most transitory and volatile component. For the quadratic utility function $u(c_t) = -\frac{1}{2}(c_t - \bar{c})^2$, using the first welfare criterion (27) provided in Section B and setting $\beta = 0.95$, we can compute that

(58)
$$\Sigma = 10^{-3}\begin{bmatrix} 0.1399 & -0.0737 & -0.0110 \\ -0.0737 & 0.1596 & -0.1820 \\ -0.0110 & -0.1820 & 0.5555 \end{bmatrix},$$

when capacity $\kappa = 2.2$ bits, which can be used to compute the variance of the noise Λ using $\Lambda^{-1} = \Sigma^{-1} - \Psi^{-1}$, and then compute the Kalman gain according to $\theta = \Sigma\Lambda^{-1}$. It is clear from Equation (58) that owing to the low capacity devoted to monitoring the state, the post-observation variances (i.e., the conditional variances) of both the x and z components are greater than the corresponding innovation variances in Equation (57). More importantly, the conditional variance of the slow-moving x component is 18 times larger than its corresponding innovation variance, whereas that of the fast-moving z component is only 2 times larger than its innovation variance.[29] The intuition behind this result is that the optimizing agent devotes much less capacity to monitoring the slow-moving component, which leads to greater effects on the conditional variance term. Figure 4 plots the impulse responses of consumption to the income shocks and noises. It shows that consumption reacts to the income shocks gradually and with delay, and reacts to the corresponding noises promptly. In addition, we can see that the response of consumption to the slow-moving x component is much more damped than that to the fast-moving z component. It is also worth noting that as the agent only cares about the trace of $Z\Sigma$ and the symmetric matrix Z is negative semidefinite, the

29. Alternatively, we can also see that the conditional variance of the x component is about three times smaller than its corresponding unconditional variance (0.4523), whereas that of the z component is about five times smaller than its corresponding unconditional variance (2.7692).

FIGURE 4

Responses of Consumption to Income Shocks and Noises
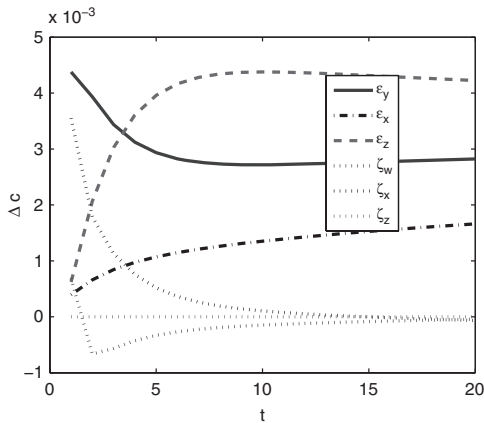
agent with low capacity will choose to make the post-observations of the states be negatively correlated. This correlation conserves capacity by permitting some information about each state to be transmitted using a single nat.

When we relax the information-processing capacity and increase κ to 2.8 nats, the conditional covariance matrix becomes

(59)

$$\Sigma = 10^{-3} \begin{bmatrix} 0.0787 & -0.0419 & 0.0153 \\ -0.0419 & 0.1172 & -0.1926 \\ 0.0153 & -0.1926 & 0.5170 \end{bmatrix}.$$

Comparing Equations (58) with (59), we can see that relaxing information-processing capacity has the largest impact on the conditional variance of the endogenous state variable w: the post-observation variance of w is reduced to about half the initial value. The intuition behind this result is that the endogenous variable plays the most important role in the welfare losses due to RI. To see this clearly, the matrix Z is displayed here:

(60)

$$Z = 10^{-2} \begin{bmatrix} -0.0204 & -0.6732 & -0.2769 \\ -0.6732 & -22.2156 & -9.1363 \\ -0.2769 & -9.1363 & -3.7573 \end{bmatrix}.$$

While w per unit has less of an effect on welfare, it is proportionally much larger than either of the other two state variables. It is also clear that as the information constraint is relaxed the agent chooses to allocate more capacity to monitoring the slow-moving component x than to monitoring the z component.

Note that in the RI problem (58) is optimal in the sense that it minimizes the expected welfare losses due to finite information-processing capacity by allocating fixed capacity optimally across different elements in the state vector. In contrast, in the SE problem, Λ must be specified first and then Σ and θ can be computed. However, it is difficult to specify Λ without prior knowledge about the states. Ad hoc assumptions on Λ might be inconsistent with the underlying efficiency conditions. Therefore, RI could provide a useful way to specify the stochastic properties of the noises by solving the agent's optimization problem subject to information constraints. As we have noted previously, Melosi (2009) presents an application of this idea; he notes that a particular estimated model shows that the marginal utility of information is not equated across variables and is thus inconsistent with RI (i.e., inconsistent with *any* value for κ).

C. Attention Allocation in a Price-Setting Model

In the multivariate permanent income model we discussed in the previous section, the difference between the RI and SE problems comes from the optimal allocation of capacity, and the consumption reactions to different income shocks do depend on optimal attention allocation. However, in this setting, the optimal attention/capacity allocation cannot be addressed explicitly because there exists an endogenous state variable and attention allocation between the endogenous state variable (w) and exogenous state variables (persistent components in y) is nonsymmetric and complicated. In contrast, as has been shown in Maćkowiak and Wiederholt (2009), the attention allocation problem between the exogenously given aggregate and idiosyncratic shocks has an explicit solution. In this section, we present a simplified version of their price-setting model.

Consider a firm price-setting model with monopolistic competition in which firms continuously observe noisy signals about two random variables: one aggregate variable and one firm-specific variable. Under the full-information assumption, the profit-maximizing price (in logs) of firm i, p_i, can be written as

(61) $$p_i^{\text{f}} = p + \alpha_x x + \alpha_z z_i,$$

where p is the log of the aggregate price level, x is the log of aggregate output, z_i is an idiosyncratic demand shock, and α_x and

α_z are coefficients that depend on structural parameters in the profit function (the superscript f in the price function indicates full information). All the variables on the RHS are assumed to be normally distributed. Since the sum of two normal variables is also normal, we can summarize the aggregate condition as $y = p + \alpha_x x$ such that

$$(62) \qquad p_i^{\mathrm{f}} = y + \alpha_z z_i,$$

where y and z are assumed to be Gaussian variables with mean 0 and variances σ_y^2 and σ_z^2, respectively. Under RI, the typical firm cannot observe y and z_i perfectly, so all it can observe are noisy signals y^* on the aggregate state and z^* on the idiosyncratic state:

$$(63) \qquad y^* = y + \xi_y \text{ and } z^* = z_i + \xi_z,$$

where the noises due to RI, ξ_y and ξ_z are Gaussian variables with mean 0 and variances ω_y^2 and ω_z^2, respectively. They are assumed to be independent with each other and are also not correlated across firms.[30] Given the observed independent noisy signals, the optimal price can be written as

$$(64) \qquad p_i^* = E\left[y|y^*\right] + \alpha_z E\left[z_i|z^*\right].$$

Because p_i^* is different from the full-information solution p_i^f, the firm suffers profit losses from RI. The profit loss function of the firm can be written as

$$(65) \qquad \Delta\pi = \frac{\gamma}{2}\left(p_i^f - p_i^*\right)^2,$$

where $\gamma > 0$. Given Equations (64) and (65), the attention allocation problem can be reduced to

$$(66) \quad \min \operatorname{var}\left(p_i^f|y^*, z^*\right) = \sigma_{y|y^*}^2 + \alpha_2^2 \sigma_{z|z^*}^2,$$

subject to the information-processing constraint (IPC):

$$\left[\frac{1}{2}\ln\left(2\pi\sigma_y^2\right) - \frac{1}{2}\ln\left(2\pi\sigma_{y|y^*}^2\right)\right]$$

$$+ \left[\frac{1}{2}\ln\left(2\pi\sigma_z^2\right) + \frac{1}{2}\ln\left(2\pi\sigma_{z|z^*}^2\right)\right] \le \kappa,$$

which can be reduced to $\frac{\sigma_{y|y^*}^{-2}\sigma_{z|z^*}^{-2}}{\sigma_y^{-2}\sigma_z^{-2}} \le \exp(2\kappa)$, where $\sigma_{y|y^*}^2$ and $\sigma_{z|z^*}^2$ are posterior conditional variances of y and z_i, respectively, and we use the facts that (1) for a quadratic objective

function and Gaussian state variables, it is optimal to choose the joint density of the states and noisy signals to be also Gaussian and (2) the aggregate and idiosyncratic variables are uncorrelated. After using the IPC to substitute out $\sigma_{z|z^*}^{-2}$, the optimal solution of $\sigma_{y|y^*}^{-2}$:

$$(67) \quad \frac{\sigma_{y|y^*}^{-2}}{\sigma_y^{-2}} = \exp(\kappa)\left|\frac{1}{\alpha_z}\right|\frac{\sigma_y}{\sigma_z} \in (1, \exp(2\kappa)),$$

which gives the ratio of posterior to prior precision of briefs about the aggregate condition under the optimal attention allocation.[31] Equation (67) provides several important implications for the optimal attention allocation to the aggregate condition. First, greater values of $\sigma_{y|y^*}^{-2}/\sigma_y^{-2}$ mean that the firm pays more attention to the aggregate condition. Second, given α_z, $\sigma_{y|y^*}^{-2}/\sigma_y^{-2}$ is increasing with the relative importance of the prior variances of the aggregate and firm-specific conditions. As the firm-specific shock is 10 times more volatile than the aggregate shock as calibrated in Maćkowiak and Wiederholt (2009), the firm optimally pays much more attention to the firm-specific shock. Similarly, we can obtain optimal $\frac{\sigma_{z|z^*}^{-2}}{\sigma_z^{-2}}$:

$$(68) \qquad \frac{\sigma_{z|z^*}^{-2}}{\sigma_z^{-2}} = \exp(\kappa)|\alpha_z|\left(\frac{\sigma_y}{\sigma_z}\right)^{-1}.$$

Finally, using Equations (67) and (68), we can easily recover the variances of the noises, ω_y^2 and ω_z^2,

$$(69) \qquad \omega_y^{-2} = \sigma_{y|y^*}^{-2} - \sigma_y^{-2},$$

$$(70) \qquad \omega_z^{-2} = \sigma_{z|z^*}^{-2} - \sigma_z^{-2}.$$

In the corresponding SE problem in which ω_y^2 and ω_z^2 are given exogenously and arbitrarily, using Equations (69) and (70), we can recover the posterior conditional variances, $\sigma_{y|y^*}^2$ and $\sigma_{z|z^*}^2$, which may not be optimal in the sense that they do not lead to the minimum welfare loss, $\operatorname{var}\left(p_i^f|y^*, z^*\right)$. The argument in this model is similar to that in the permanent income

---

30. Maćkowiak and Wiederholt (2009) show that independent signals are optimal.

31. Note that this ratio must be greater than 1, as otherwise the posterior variance would be higher than the prior variance, which means that the firms forget. If this ratio is greater than $\exp(2\kappa)$, it violates the information-processing constraint.

model with an endogenous state variable in the previous section.[32]

We conclude this section with a point regarding the importance of distinguishing between SE and RI. In the context of this model, Paciello and Wiederholt (2013) show how the distinction for SE and RI matters for optimal policy. As under RI the variance-covariance matrix of the noises will shift with monetary policy, the central bank will use that fact to improve allocations—specifically, they find that the *divine coincidence* of monetary policy (no trade-off between inflation and output stabilization) holds under RI but not under SE. Of course, this result does not help us distinguish between the two theories, but rather provides a motive for doing so.

V. DISCUSSION ON THE CONTINUOUS-TIME SETTING

So far we have focused on the discrete-time setting in which agents are assumed to have imperfect state information. In this section we discuss the different implications of SE and RI in the continuous-time setting. For simplicity, we use the permanent income model as an example to illustrate the key different implications of the two information frictions. In the continuous-time version of the permanent income model, the typical consumer is assumed to maximize the following lifetime utility:

$$(71) \qquad E_0 \left[\int_{t=0}^{\infty} \exp\left(-\beta t\right) u(c_t) dt \right],$$

subject to the following budget constraint:

$$(72) \qquad dw_t = \left(r w_t + y_t - c_t\right) dt,$$

where $u(c_t) = -\left(\overline{c} - c_t\right)^2/2$, and all the variables here have the same definitions as that in the last section. Labor income (y_t) is assumed to follow a continuous-time AR(1) (Ornstein-Uhlenbeck) process:

$$(73) \qquad dy_t = \rho\left(\frac{\mu}{\rho} - y_t\right) dt + \sigma d\zeta_t,$$

where the steady-state income $\overline{y} = \mu/\rho$, $\rho > 0$ governs the speed of convergence from the steady state, and ζ_t is a standard Brownian motion.

Just like the discrete-time model, to simplify the model, we adopt the same state-space reduction approach and define a new state $s_t = a_t + y_t/(r+\rho)$. Consequently, the new state transition equation can be written as:

$$(74) \qquad ds_t = (r s_t - c_t) dt + \sigma_s d\zeta_t,$$

where $\sigma_s = \sigma/(r+\rho)$. In the full-information case, it is straightforward to show that the consumption function is:

$$(75) \qquad c_t = r s_t.$$

We assume that the consumer cannot observe the state perfectly and only observes a noisy signal:

$$(76) \qquad s_t^* = s_t + \xi_t,$$

where ξ_t is a standard Brownian motion with mean 0 and variance Λ, and is independent of ζ_t. To solve the model, we rewrite Equation (76) in the following differentiation form:

$$ds_t^* = ds_t + d\xi_t.$$

In this continuous-time LQG setting, the separation principle holds and the consumption function can now be written as

$$(77) \qquad c_t = r\widehat{s_t},$$

and the perceived state $\widehat{s_t} = E_t[s_t]$ is governed by the following Kalman filtering equation:

$$(78) \qquad d\widehat{s_t} = (r\widehat{s_t} - c_t) dt + \theta d\eta_t,$$

where

$$(79) \qquad d\eta_t = r(s_t - \widehat{s_t}) dt + \sigma_s d\zeta_t + d\xi_t,$$

with mean $E[d\eta_t] = 0$ and var$(d\eta_t) = (\sigma_s^2 + \Lambda)dt$, $\theta = \Sigma\Lambda^{-1}$ is the steady-state Kalman gain, and $\Sigma(= (\sigma_s^{-2} + \Lambda^{-1})^{-1} > 0)$ is the conditional variance of s. Note that η_t is a Brownian motion with mean 0. Although the Brownian variable, ξ_t, is not observable, the innovation process, η_t, is observable because it is derived from observable processes (i.e., ds_t^* and $(r\widehat{s_t} - c_t) dt$). In this case, the path of the conditional expectation, $\widehat{s_t}$, is generated by the path of the innovation process, η_t. In summary, given Λ, Equations (77) and (78) completely characterize the dynamics of the continuous-time PIH model.

We now move on to discuss the RI case in which we assume that the consumer only has finite channel capacity when processing relevant information about the state. As emphasized in

---

32. Menkulasi (2010) presents a model of the Great Moderation under RI and shows that RI implies that the required decline in volatility for the shocks needed to replicate observed GDP volatility declines is smaller, but does not provide evidence of the actual volatility declines.

Sims (1998) and discussed in Kasa (2006) and Reis (2010), the observation equation, (76), is not suitable to model RI owing to finite capacity because this specification means that in any finite interval, arbitrarily large amounts of information can be passed through the consumer's channel. Specifically, Sims (1998) shows that when the true state (s) and the noise (ξ) are generated by nicely behaved stochastic differential equations driven by Wiener processes, then if the noise has at least as many derivatives as the true state, noisy observations can transmit infinite amounts of information in finite time, while if the noise has fewer derivatives than the true state, noisy observations over a finite interval can transmit only finite amounts of information. Moscarini (2004) solves this problem by assuming that the agent samples the process at discrete intervals.[33] Kasa (2006) proposes an alternative way to solve this problem. He assumes that the observation equation is written as:

$$(80) \qquad ds_t^* = s_t dt + d\xi_t,$$

where the current signal affects the instantaneous rate of change in the noisy observation. In this specification, the Kalman filtering equation can be written as

$$(81) \qquad d\widehat{s}_t = (r\widehat{s}_t - c_t)\, dt + \theta d\widetilde{\eta}_t,$$

where

$$(82) \qquad d\widetilde{\eta}_t = (s_t - \widehat{s}_t)\, dt + d\xi,$$

with mean $E\left[d\widetilde{\eta}_t\right] = 0$ and $\mathrm{var}\left(d\widetilde{\eta}_t\right) = \Lambda dt$. Here $\widetilde{\eta}_t$ is also an observable Brownian motion with mean 0. It is clear from Equations (79) and (82) that the two specifications of imperfect information observations, (76) and (80), lead to different consumption behavior.

In this RI case, of course, the variance of the noise, Λ, is determined endogenously by the following continuous-time version of the information-processing constraint specified in Section A:

$$(83) \qquad \int \ln\left(\frac{dP_{ss^*}}{d\left(P_s \times P_{s^*}\right)}\right) dP_{ss^*} \le \kappa t,$$

where P_{ss^*} is the joint probability measure on s and s^*, P_s and P_{s^*} are the marginal probability measures of s and s^*, respectively, and κ is finite capacity. As shown in Duncan (1970) and Reis (2010), Equation (83) reduces to $\Sigma\Lambda^{-1} \le \kappa$,

which determines the value of Λ. Comparing Equations (79) with (82), it is clear that we must distinguish the two noisy signal specifications when modeling RI due to finite capacity; that is, it is not the case that all SE environments can be derived from an underlying RI problem.

VI. CONCLUSIONS

In this paper we have explored the implications of two informational friction theories, SE and RI, for economic behavior, policy, and welfare within the LQG setting. First, we showed that if the variance of the noise itself is fixed exogenously, the two theories can be distinguished as they lead to different dynamics and welfare after implementing government policies. Second, we showed that if the SNR is fixed in the SE problem and the channel capacity is fixed in the RI problem, SE and RI are observationally equivalent in the sense that they lead to the same dynamics even after implementing policies in the univariate case, whereas they generate different policy and welfare implications in the multivariate case. Furthermore, in the multivariate case we showed that under RI the agent's preference, budget constraint, and information-processing constraints jointly determine the stochastic properties of the post-observation variance and endogenous noise; hence, RI provides a micro-founded way to specify the nature of the Kalman gain that governs the model's dynamics.

Distinguishing between SE and RI has policy relevance beyond the simple examples we consider here. As we noted earlier, Paciello and Wiederholt (2013) study optimal monetary policy in a model that nests both SE and RI. They find that the policy can differ across the two types of models—the key is that under RI the policymaker can affect the attention allocation between different shocks. In particular, they find that the divine coincidence of no trade-off between stabilizing prices and output holds under RI, but not under SE.

Extending our results to compare environments outside the LQG setup will be challenging. Solving the SE problem is straightforward, although computation of the conditional expectations may be difficult depending on the distributional assumptions.[34] The RI problem is difficult to solve, however, because the optimal

---

33. Moscarini (2004) argues that although the state transition and predictions (actions) happen continuously, observations only occur at discrete intervals.

34. See Veldkamp (2011) for a textbook treatment on information choice.

joint distribution of states and controls is typically not of a known form and not easy to approximate; the optimality of discrete solutions discussed in Matejka and Sims (2010) make it difficult to characterize the distribution in terms of a small number of parameters. It seems therefore unlikely that SE and RI environments will generally be observationally equivalent.

REFERENCES

Adam, K. "Optimal Monetary Policy in the Presence of Imperfect Common Knowledge." *Journal of Monetary Economics*, 54(2), 2005, 267–301.

Aguiar, M., and G. Gopinath. "Emerging Market Business Cycles: The Cycle Is the Trend." *Journal of Political Economy*, 115(1), 2007, 69–102.

Angeletos, G.-M., and J. La'O. "Noisy Business Cycles." *NBER Macroeconomics Annual 2009*, 24, 2010, 319–78.

Cochrane, J. *Asset Pricing*. Princeton, NJ: Princeton University Press, 2005.

Duncan, T. E. "On the Calculation of Mutual Information." *SIAM Journal of Applied Mathematics*, 19, 1970, 215–20.

Flavin, M. A. "The Adjustment of Consumption to Changing Expectations About Future Income." *Journal of Political Economy*, 89(5), 1981, 974–1009.

García-Cicco, J., R. Pancrazi, and M. Uribe. "Real Business Cycles in Emerging Countries?" *American Economic Review*, 100(5), 2010, 2510–31.

Hall, R. E. "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence." *Journal of Political Economy*, 86(6), 1978, 971–87.

Hansen, L. P. "Calculating Asset Prices in Three Example Economies," in *Advances in Econometrics, Fifth World Congress*, edited by T. F. Bewley. Cambridge: Cambridge University Press, 1987.

Hansen, L. P., and T. J. Sargent. *Robustness*. Princeton, NJ: Princeton University Press, 2007.

Kahneman, D. *Attention and Effort*, Englewood Cliffs, NJ: Prentice Hall Press, 1973.

Kasa, K. "Robustness and Information Processing." *Review of Economic Dynamics*, 9(1), 2006, 1–33.

Kim, J., E. Ko, and T. Yun. "The Role of Bounded Rationality in Macro-Finance Affine Term-Structure Models." Unpublished, 2012.

Lucas, R. E., Jr. "Expectations and the Neutrality of Money." *Journal of Economic Theory*, 4(1), 1972, 103–24.

Luo, Y. "Consumption Dynamics under Information Processing Constraints." *Review of Economic Dynamics*, 11(2), 2008, 366–85.

———. "Rational Inattention, Long-run Consumption Risk, and Portfolio Choice." *Review of Economic Dynamics*, 13(4), 2010, 843–60.

Luo, Y., and E. R. Young. "Risk-sensitive Consumption and Savings under Rational Inattention." *American Economic Journal: Macroeconomics*, 2(4), 2010a, 281–325.

———. "Asset Pricing under Information-Processing Constraints." *Economics Letters*, 107(1), 2010b, 26–29.

Luo, Y., J. Nie, and E. R. Young. "Robustness, Information-Processing Constraints, and the Current Account in Small Open Economies." *Journal of International Economics*, 88(1), 2012, 104–20.

Maćkowiak, B., and M. Wiederholt. "Optimal Sticky Prices under Rational Inattention." *American Economic Review*, 99(3), 2009, 769–803.

Matejka, F., and C. A. Sims. "Discrete Actions in Information-Constrained Tracking Problems." Unpublished, 2010.

Melosi, L. "A Likelihood Analysis of Models with Information Frictions," Penn Institute for Economic Research Working Paper 09-009, 2009.

Menkulasi, J. "Rational Inattention and Changes in Macroeconomic Volatility." Unpublished, University of Maryland, 2010.

Morris, S., and H. Song Shin. "The Social Value of Public Information." *American Economic Review*, 92(5), 2002, 1521–34.

Moscarini, G. "Limited Information Capacity as a Source of Inertia." *Journal of Economic Dynamics and Control*, 28(10), 2004, 2003–35.

Muth, J. F. "Optimal Properties of Exponentially Weighted Forecasts." *Journal of the American Statistical Association*, 55(290), 1960, 299–306.

Paciello, L., and M. Wiederholt. "Exogenous Information, Endogenous Information, and Optimal Monetary Policy." *Review of Economic Studies*, 2013. doi: 10.1093/restud/rdt024

Reis, R. "When Should Policymakers Make Announcements?" Unpublished, 2010.

Sargent, T. J. "The Observational Equivalence of Natural and Unnatural Rate Theories of Macroeconomics." *Journal of Political Economy*, 84(3), 1976, 631–40.

———. "Equilibrium with Signal Extraction from Endogenous Variables." *Journal of Economic Dynamics and Control*, 15(2), 1991, 245–73.

Simon, D. *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches*. Hoboken, NJ: John Wiley & Sons, 2006.

Sims, C. A. "Stickiness." *Carnegie-Rochester Conference Series on Public Policy*, 49, 1998, 317–56.

———. "Implications of Rational Inattention." *Journal of Monetary Economics*, 50(3), 2003, 665–90.

———. "Rational Inattention: Beyond the Linear-Quadratic Case." *American Economic Review*, 96(2), 2006, 158–63.

———. "Rational Inattention and Monetary Economics," in *Handbook of Monetary Economics*, Vol. 3, Chapter 4, edited by B. M. Friedman and M. Woodford. Amsterdam: Elsevier, 2010, 155–81.

Townsend, R. M. "Forecasting the Forecasts of Others." *Journal of Political Economy*, 91(4), 1983, 546–88.

Van Nieuwerburgh, S., and L. Veldkamp. "Information Acquisition and Under-Diversification." *Review of Economic Studies*, 77(2), 2010, 779–805.

Veldkamp, L. *Information Choice in Macroeconomics and Finance*. Princeton, NJ: Princeton University Press, 2011.

Wang, N. "Precautionary Saving and Partially Observed Income." *Journal of Monetary Economics*, 51(8), 2004, 1645–81.

Woodford, M. "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford. Princeton, NJ: Princeton University Press, 2001.