Signal Extraction and Rational Inattention*

Yulei Luo†
University of Hong Kong

Eric R. Young‡
University of Virginia

October 3, 2010

Abstract

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 are fixed exogenously in the SE and RI problems, respectively, the two environments lead to different policy and welfare implications. We also find that if the signal-to-noise ratio in the SE problem is fixed, the two theories generate the same policy implications in the univariate case, but different policy implications in the multivariate case. These results are robust to the presence of correlation between structural shocks and noise shocks, and the presence of the risk-sensitivity preference. Furthermore, in the multivariate case we show that under RI, the agent's preferences, budget constraint, and information-processing constraints jointly determine the conditional variance of the state and the Kalman gain, whereas under SE they are determined by the exogenously given variance of the noise and the budget constraint; in other words, RI provides a microfoundation for the imprecise observations and noise in the SE problem.

JEL Classification Numbers: *C61, D81, E21.*

Keywords: Signal Extraction, Rational Inattention, Exogenous and Endogenous Noises, Risk-sensitive Filtering.

*We thank Tom Sargent and Chris Sims for helpful discussions and Marios Angeletos for suggesting we pursue this topic. We are also grateful for useful suggestions and comments from Yang Lu and seminar participants at the Hong Kong University of Science and Technology. Luo thanks the General Research Fund (GRF) in Hong Kong and the HKU seed funding program for basic research for financial support. Young thanks the Bankard Fund for Political Economy at the University of Virginia for financial support. Part of this work was conducted while Luo was visiting the UC Davis Economics Department and the Kansas City Fed, whose hospitality is greatly appreciated. All errors are the responsibility of the authors.

† Corresponding author. School of Economics and Finance, The University of Hong Kong, Hong Kong. E-mail: `yluo@econ.hku.hk`.

‡ Department of Economics, University of Virginia, Charlottesville, VA 22904. E-mail: `ey2d@virginia.edu`.

1. Introduction

Muth (1960) applied classical filtering methods to solve for a stochastic process for permanent income for which Friedman (1956)'s adaptive expectations hypothesis would be an optimal estimator of permanent income. Specifically, Muth (1960) solved a single-agent dynamic signal extraction (SE) problem, in which an economic agent was modelled as facing exogenous signals and noises which had to be disentangled; and shows that the exponentially weighted average of past observations of a random walk plus noise process is optimal in the sense that it minimizes the mean squared estimation error.[1] Townsend (1983) and Sargent (1991) extended the single-agent signal extraction problem by studying multiple-agent settings in which agents extract signals from endogenous variables that are affected by other agents' signal extraction problems. Recently, there have been some papers examining the effects of heterogeneous information on economic dynamics within signal extraction settings. For example, Morris and Shin (2002) examined the welfare effects of asymmetric information in the presence of strategic complementarity. Angeletos and La'O (2009) studied how dispersed information about the underlying aggregate productivity shock contributes significant noise in the business cycle and helps explain cyclical variations in observed Solow residuals and labor wedges in the RBC setting. The key assumption in these signal extraction settings is that the stochastic properties of noises are given exogenously.

Sims (2003) first introduced rational inattention (RI) into economics within the linear-quadratic Gaussian (LQG) setting and argued that it is a plausible method for introducing sluggishness, randomness, and delay into economic models.[2] In his formulation agents have finite Shannon channel capacity, limiting their ability to process signals about the true state of the world. As a result, an impulse to the economy induces only gradual responses by individuals, as their limited capacity requires many periods to discover just how much the state has moved; one key change relative to the RE case is that consumption has a hump-shaped impulse response to income shocks.[3] Luo (2008) used this model to explore anomalies in the consumption literature, particularly the well-known "excess sensitivity" and "excess smoothness" puzzles, employing an LQ version of the standard permanent income model (as in Hall 1978 and Flavin 1981). In that model RI is equivalent to confronting the household with a noisy signal about the value of permanent income but permitting the agents to choose the

[1] See Sargent (1987) and Hansen and Sargent (2007) for textbook treatments.

[2] For the applications of RI within the (approximate) LQG setting, see Adam (2005), Luo and Young (2010), Maćkowiak and Wiederholt (2009), Melosi (2009), etc.

[3] See Sims (2003) and Luo (2008).

distribution of the noise terms, subject to their limited capacity. The key feature of the LQ-RI model is that the RI-induced noise is optimal and generated endogenously due to individuals' finite information-processing capacity.[4]

The main objective of this paper is to examine and compare the effects of SE and RI for economic dynamics, policy, and welfare within the linear-quadratic-Gaussian setting. The key difference between the two informational-friction modeling strategies, SE and RI, is that in the SE problem given the variances of the exogenous shock and noise, the propagation equation for the post-observation variance (i.e., the conditional variance of the state) and the Kalman filtering equation jointly determine the conditional variance and the Kalman gain. In contrast, in the RI problem given the level of finite capacity, we first solve for the conditional variance of the state and then use the propagation equation for the conditional variance to recover the variance of the noise.

To explore the effects of the two different setups on the dynamic behavior of the model, we first study the univariate case in which there is only one state and the models can be solved explicitly. The first 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 as they lead to different dynamic behavior, policy, and welfare implications. Specifically, we assume that the variance of the exogenous shock is scaled up due to a change in policy. In the SE problem with exogenous noises, an increase in the variance of the exogenous shock will lead to a different solution for the conditional variance and Kalman gain; consequently, the change in policy will eventually lead to a change in the model's dynamic behavior and the agent's welfare. In contrast, in the RI problem, if κ 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 has no impact on the Kalman gain. That is, inattentive agents will behave as if facing noise whose nature changes systematically as the dynamic properties of the economy change due to the change in policy. In addition, we show that the welfare losses of agents due to imperfect information also depend on the value of the Kalman gain. Therefore, SE and RI can lead to different policy and welfare implications in the LQG setting. In contrast, when extending to the risk-sensitive (RS) LQG setting, we show that SE and RI still lead to different policy and welfare implications but in this case the change in the variance of the exogenous shock will affect the RS filter gain governing the

[4] Fully non-LQ versions of the RI problem are solved and discussed in Sims (2005, 2006), Lewis (2006), and Tutino (2009). The main feature of the non-LQ RI models is that they have either very short horizons or extremely simple setups due to numerical obstacles – the state of the world is the distribution of true states and this distribution is not well-behaved (it is not generally a member of a known class of distributions and tends to have 'holes,' making it difficult to characterize with a small number of parameters).

dynamic behavior in the RI problem. 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 is equal to $1 - 1/\exp(2\kappa)$ in the RI problem in which κ is the exogenously given channel capacity. After considering correlated shocks and noise, we find that our results remain unchanged.

We then move on to study the multivariate case in which the state vector includes multiple elements. 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 welfare losses 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 will lead to different dynamic behavior and deliver different policy and welfare implication after the government implements a policy that changes the variance of the exogenous shock even if the signal-to-noise ratio 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 the underlying efficiency conditions (equalization of the marginal utility of additional capacity across variables).[5] Therefore, RI provides a useful and microfounded way to specify the stochastic properties of the noises by solving the agent's constrained optimization problem. 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 determined the conditional variance and then the Kalman gain.

The remainder of the paper is organized as follows. Section 2 examines optimal decisions and economic dynamics in an LQG setting with signal extraction. Section 3 presents the RI version of the model and compares different implications of RI and SE on the dynamic behavior, policy and welfare within the LQG setting. Section 4 presents applications to permanent income models and also discuss an extension to the risk-sensitive setting. Section 6 concludes.

[5] See Melosi (2009) for an application of this idea.

2. Signal Extraction in a LQ Gaussian Model

2.1. Full-information Rational Expectations LQ Model

Consider the following linear-quadratic-Gaussian (LQG) model:

$$v\left(s_0\right)=\max_{\{c_t,s_{t+1}\}} E_0\left[\sum_{t=0}^{\infty}\beta^t\left(s_t^TQs_t+c_t^TRc_t+2c_t^TWs_t\right)\right], \tag{2.1}$$

subject to

$$s_{t+1}=As_t+Bc_t+\varepsilon_{t+1}, \tag{2.2}$$

with s_0 known and given, where β 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 $E_t\left[\cdot\right]$ 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

$$s_t^TPs_t=\max_{c_t}\left\{s_t^TQs_t+c_t^TRc_t+2c_t^TWs_t+\beta E_t\left[\left(s_t^TA^T+c_t^TB^T+\varepsilon_{t+1}^T\right)P\left(As_t+Bc_t+\varepsilon_{t+1}\right)\right]\right\},$$

yields the decision rule

$$c_t^*=-Fs_t, \tag{2.3}$$

and the Riccati equation is

$$P=Q+F^TRF-2F^TW+\beta\left(A^T-F^TB^T\right)P\left(A-BF\right), \tag{2.4}$$

where

$$F=\left(R+\beta B^TPB\right)^{-1}\left(W+\beta B^TPA\right). \tag{2.5}$$

Iterating on the matrix Riccati equation (2.4) uniquely determines P, since the equation defines a contraction mapping. Using P, we can determine F in the optimal policy (2.5).

2.2. Signal Extraction with Exogenous Noises

Following the signal extraction literature (e.g., Muth 1960; Lucas 1972, 1973; Morris and Shin 2002), 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 signal extraction 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.[6]

Under the linear-quadratic-Gaussian assumption, the certainty equivalence principle holds when the agent cannot observe s_t perfectly, i.e., the decision rule under imperfect information can be written as

$$c_t^* = -F\widehat{s}_t, \tag{2.6}$$

where $\widehat{s}_t = E\left[s_t|\mathcal{I}_t\right]$ 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:

$$s_t|\mathcal{I}_t \sim N\left(E\left[s_t|\mathcal{I}_t\right], \Sigma_t\right),$$

where $\Sigma_t = E_t\left[\left(s_t - \widehat{s}_t\right)\left(s_t - \widehat{s}_t\right)^T\right]$, and the following Kalman filtering equation governs the behavior of $\widehat{s}_t$:

$$\widehat{s}_{t+1} = \left(1 - \theta_t\right)\left(A\widehat{s}_t + Bc_t\right) + \theta_t s_{t+1}^*, \tag{2.7}$$

where θ is the Kalman gain to be determined.[7] To derive the optimal Kalman gain, we write the error in estimating the true state, $e_{t+1} = s_{t+1} - \widehat{s}_{t+1}$, as follows:

$$e_{t+1} = \left(1 - \theta_t\right)Ae_t + \left(1 - \theta_t\right)\varepsilon_{t+1} - \theta_t\xi_{t+1}, \tag{2.8}$$

which means that

$$e_{t+1}e_{t+1}^T = \left(I - \theta_t\right)Ae_te_t^TA^T\left(1 - \theta_t\right)^T + \left(1 - \theta_t\right)\varepsilon_{t+1}\left[\left(1 - \theta_t\right)\varepsilon_{t+1}\right]^T + \theta_t\xi_{t+1}\left(\theta_t\xi_{t+1}\right)^T \tag{2.9}$$

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

[7] Muth (1960) shows 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.

Taking unconditional mean on both sides of (2.9) gives

$$\Sigma_{t+1} = (I - \theta_t) A\Sigma_t A^T (I - \theta_t)^T + (I - \theta_t) \Omega (I - \theta_t)^T + \theta_t \Lambda \theta_t^T, \tag{2.10}$$

where $\Sigma_{t+1} = E\left[e_{t+1} e_{t+1}^T\right]$. We now discuss how to use (2.10) determine the optimal value of θ_t. The optimality criterion that we choose to minimize is the sum of the variances of the estimation errors at time $t+1$:

$$\begin{aligned} J_{t+1} &= E\left[e_{t+1}^T e_{t+1}\right] \\ &= E\left[\operatorname{trace}\left(e_{t+1} e_{t+1}^T\right)\right] \\ &= \operatorname{trace}\left(\Sigma_{t+1}\right), \end{aligned} \tag{2.11}$$

where $e_{t+1} = \left[e_{t+1}(1), \cdots, e_{t+1}(n)\right]^T$. Using (2.10) and taking derivatives with respect to K yields[8]

$$-2(I - \theta_t) A\Sigma_t A^T - 2(I - \theta_t)\Omega + 2\theta_t \Lambda = 0,$$

which leads to the optimal Kalman gain

$$\theta_t = \left(\Omega + A\Sigma_t A^T\right)\left(\Omega + A\Sigma_t A^T + \Lambda\right)^{-1}. \tag{2.12}$$

Note that we can also assume that the optimal Kalman gain minimizes a weighted L^2-norm of the expected value of the estimation error e:

$$J_{t+1} = E\left[e_{t+1}^T S e_{t+1}\right],$$

where S is a positive definite user-defined weighting matrix. In this case

$$\begin{aligned} J_{t+1} &= E\left[e_{t+1}^T S e_{t+1}\right], \\ &= E\left[\operatorname{trace}\left(S e_{t+1} e_{t+1}^T\right)\right] \\ &= \operatorname{trace}\left(S\Sigma_{t+1}\right), \end{aligned} \tag{2.13}$$

[8] Note that here we use the fact that

$$\frac{\partial \operatorname{trace}\left(ABA^T\right)}{\partial A} = 2AB$$

if B is symmetric.

where we use the fact that $E\left[\text{trace}\left(SX_{t+1}\right)\right] = \text{trace}\left(SE_t\left[X_{t+1}\right]\right)$. Since S is a constant matrix, minimizing (2.13) leads to the same expression for the optimal Kalman gain as minimizing (2.11).

If iterations on Σ using (2.10) converge, in the steady state we have

$$\Sigma = (I-\theta)\left(A\Sigma A^T + \Omega\right)(I-\theta)^T + \theta\Lambda\theta^T \tag{2.14}$$

and

$$\theta = \left(\Omega + A\Sigma A^T\right)\left(\Omega + A\Sigma A^T + \Lambda\right)^{-1}. \tag{2.15}$$

(Baxter, Graham, and Wright 2010 discusses the convergence of the iteration for Kalman filtering problems with endogenous variables.) Starting from the initial condition Σ_0, we can compute the steady state (θ, Σ) by iterating on (2.14) and (2.15). After computing (θ, Σ), 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, (2.14) and (2.15) jointly determine the steady state (θ, Σ).

It is worth noting that we have the following alternative equations for computing the Kalman gain and the conditional variance-covariance matrix, (K, Σ):

$$\Sigma_{t+1} = \Psi_t - \Psi_t\left(\Psi_t + \Lambda_t\right)^{-1}\Psi_t, \tag{2.16}$$

and

$$\theta_t = \Sigma_t\Lambda_t^{-1}. \tag{2.17}$$

(See Appendix 7.1 for proof.) In the steady state, (2.16) and (2.17) reduce to

$$\Lambda^{-1} = \Sigma^{-1} - \Psi^{-1}. \tag{2.18}$$

and

$$\theta = \Sigma\Lambda^{-1}, \tag{2.19}$$

respectively. After obtaining (2.19), (4.21), (2.6), and (2.7) completely characterize the model's dynamic behavior. Note that the propagation mechanism of the model is mainly governed by the Kalman gain K.

3. Rational Inattention in the LQG Model

Following Sims (2003), we introduce rational inattention (RI) into the LQG model proposed in Section 2.1 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 log of the density function, $-E\left[\log\left(f\left(X\right)\right)\right]$. For example, the entropy of a discrete distribution with equal weight on two points is simply $E\left[\ln_2\left(f\left(X\right)\right)\right] = -0.5\ln_2\left(0.5\right) - 0.5\ln_2\left(0.5\right) = 0.69$, and the unit of information contained in this distribution is 0.69 "nats".[9] 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 a_{t+1} as much as possible. Formally, this idea can be described by the information constraint

$$\mathcal{H}\left(s_{t+1}|\mathcal{I}_t\right) - \mathcal{H}\left(s_{t+1}|\mathcal{I}_{t+1}\right) \leq \kappa, \tag{3.1}$$

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. We assume that the noise ξ_{t+1} is Gaussian.[10] Finally, following the literature, we suppose that the prior a_{t+1} is a Gaussian random variable.

Under the LQG setting, as shown in Sims (2003, 2006), the true state under RI also follows

[9] 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.'

[10] As shown in Sims (2003), within the linear-quadratic-Gaussian setting Gaussian noise is optimal.

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, in the steady state the agent observe an 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_t:

$$\Sigma_{t+1} = \Psi_t - \Psi_t\left(\Psi_t + \Lambda_t\right)^{-1}\Psi_t, \tag{3.2}$$

where $\Psi_t = A\Sigma_t A^T + \Omega$ is the conditional variance of the state *prior to* observing the new signal at $t+1$.[11] If iterations on Σ converge (which depends on both A and Σ), (3.2) reduces to

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

which can be solved for

$$\Lambda^{-1} = \Sigma^{-1} - \Psi^{-1}. \tag{3.4}$$

Using these expressions, the Kalman gain K can be rewritten as

$$\theta = \Sigma\Lambda^{-1}. \tag{3.5}$$

3.1. The Univariate Case

The key difference between signal extraction and rational inattention is that under RI the agent faces the following information-processing constraint:

$$-\log\left(|\Sigma_{t+1}|\right) + \log\left(\left|A^T\Sigma_t A + \Omega\right|\right) \leq 2\kappa. \tag{3.6}$$

Since more information about the state is better in economic models, this constraint should be binding.[12]

Consider the simplest univariate state case in which $n = 1$, (3.6) fully determines the value

[11] Equation (3.2) can also be expressed as

$$\Sigma_{t+1} = \left(I - K_t\right)\left(A\Sigma_t A^T + \Omega\right)\left(I - K_t\right)^T + K_t\Lambda_t K_t^T. \tag{3.3}$$

[12] 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.

of the steady state conditional variance Σ:

$$\Sigma = \frac{\Omega}{\exp(2\kappa) - A^2}, \tag{3.7}$$

which means that Σ is determined by the variance of the exogenous shock (Ω) and the exogenously given capacity (κ).[13] Given this Σ, we can use (3.4) to recover the variance of the endogenous noise (Λ):

$$\Lambda = \left(\Sigma^{-1} - \Psi^{-1}\right)^{-1}, \tag{3.8}$$

where $\Psi = A^2\Sigma + \Omega$, and use (3.5) to find the Kalman gain (K):

$$\theta = \Sigma\Lambda^{-1} = 1 - \Sigma\Psi^{-1}. \tag{3.9}$$

Substituting (3.7) and (3.8) into (3.9), we have

$$\theta = 1 - \frac{1}{\exp(2\kappa)}. \tag{3.10}$$

Note that (3.8) and (3.9) also hold in the SE problem. To compare the RI and SE problems in the univariate case, we first make the following assumption.[14]

Assumption 1. *Assume that Λ is fixed exogenously.*

Under Assumption 1, it is clear that in the SE problem given Λ and Ω, we can compute Σ by solving the nonlinear equation (2.18). After obtaining Σ, we can use (2.19) 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 due to a change in policy. In the SE problem with fixed Λ, Equations (3.8) and (3.9) 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. Since 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*, (3.7), (3.8), and (3.9) imply that a change in Ω will lead to the same change in Σ, Ψ, and Λ, but has

[13] Note that here we need to impose the restriction $\exp(2\kappa) - A^2 > 0$. If this condition fails, the agent cannot control the state enough and the unconditional variance diverges over time (that is, the state is not *controllable*.

[14] Sims (2003) also briefly discussed the different implications of this assumption in the SE and RI problems.

no impact on θ. In other words, agents under RI will behave as if facing noise whose nature changes systematically as the dynamic properties of the economy change, i.e., the change in policy does not change the model's dynamics.

In the above analysis, for simplicity we assume that κ remains unchanged when Ω is affected by the government policy. However, in reality 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.[15] 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 permanent income model examined in Section 4.

Assumption 2. *Assume that the signal-to-noise ratio* (SNR), $\Omega\Lambda^{-1}$, *is fixed exogenously.*

Note that Equation (3.8) can be rewritten as

$$\Omega\Lambda^{-1} = \Omega\Sigma^{-1} - \left[A^2\left(\Omega\Sigma^{-1}\right)^{-1} + 1\right]^{-1}. \tag{3.11}$$

Under Assumption 2, since the SNR is fixed, 3.11) can be used to solve for $\Omega\Sigma^{-1}$. Given the SNR and $\Omega\Sigma^{-1}$, we can compute

$$\Sigma\Lambda^{-1} = \left(\Sigma\Omega^{-1}\right)\left(\Omega\Lambda^{-1}\right). \tag{3.12}$$

Consider the same case in which Ω is scaled up, Assumption 2 means that the exogenous noise should also be scaled up such that $\Omega\Lambda^{-1}$ is fixed at the same level; consequently, (3.11) leads to the same solution for $\Omega\Sigma^{-1}$ and (3.12) leads to the same $\Sigma\Lambda^{-1}$. The following proposition summarizes the main conclusion in this case:

Proposition 3. *Under Assumption 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 when the* $\Sigma\Lambda^{-1}$ *in the SE problem is equal to* $1 - 1/\exp(2\kappa)$ *in the RI problem in which* κ *is exogenously given channel capacity.*

Proof. *The proof is straightforward by comparing (3.10) and (3.12).* ■

[15] Sims (2003) solves the RI problem assuming a fixed marginal utility of information (a fixed Lagrange multiplier on 3.6.)

3.1.1. Public and Private Signals

We now discuss different implications of SE and RI in a model with a continuum of agents in the presence of public and private signals. As in Morris and Shin (2002) and Angeletos and La'O (2009), in the SE problem we assume that agents observe two types of signals about the state (s): public and private signals. Specifically, the public signal takes the following additive form:

$$s^* = s + \eta, \tag{3.13}$$

where η is a normally distributed noise, independent of the true state s, with mean 0 and variance ω_η^2. The signal s^* is "public" in the sense that the actual realizations of the signal are common knowledge to all agents. Similarly, we assume that agent i observes the following private signal:

$$s_i^* = s + \epsilon_i, \tag{3.14}$$

where noise terms ϵ_i of the continuum population are normally distributed with mean 0 and variance ω_ϵ^2, independent of s and η, and $E[\epsilon_i \epsilon_j] = 0$ for $i \neq j$. The signal s_i^* is "private" in the sense that the signal is not observable by the others. In this SE problem, agents make optimal decisions after observing and processing all available information about the public and private signals (s^*, s_i^*).

In contrast, in the model with a continuum of agents, the RI theory predicts that agents in the model economy only observe a noisy signal of the form

$$s_i^* = s + \xi, \tag{3.15}$$

where ξ is the iid endogenous noise due to finite capacity with mean 0 and variance Λ. (3.15) shows that even if every consumer only faces the common shock, the RI-induced noise, ξ, introduces heterogeneity into the RI economy since the idiosyncratic noises are generated via individuals' own information-processing channels.

RI theory does not distinguish public signals from private signals explicitly. From the definition of the noisy signal, (3.15), we may regard s_i^* as a private signal. However, as argued in Sims (2003), although the randomness in an individual's response to common shocks will be idiosyncratic, there is likely a significant common component in the noise term. The intuition is that people's needs for coding macroeconomic information efficiently are similar, so they rely on common sources of coded information. Therefore, the common term of the idiosyncratic

error, $\overline{\xi}_t$, is a part of the error, ξ. Formally, assume that ξ_t consists of two independent noises: $\xi = \overline{\xi} + \xi^i$, where $\overline{\xi} = E^i[\xi]$ and ξ^i are the common and purely idiosyncratic components of the error generated by ζ_t, respectively. A single parameter,

$$\lambda = \frac{\text{var}\left[\overline{\xi}\right]}{\text{var}\left[\xi\right]} \in [0,1],$$

can be used to measure the common source of coded information on the aggregate component (or the relative importance of $\overline{\xi}$ vs. ξ). Therefore, if the noisy signal in the SE problem can be written as

$$s_i^* = s + \eta + \epsilon_i,$$

$\omega_\eta^2 = \text{var}\left[\overline{\xi}\right]$, and $\omega_\epsilon^2 = (1-\lambda)\,\text{var}\left[\xi\right]$, the RI and SE problems have the same Gaussian information structure.

3.1.2. Correlated Shock and Noise

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 reality, we do 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.

It is relatively 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:

$$\begin{aligned}\text{corr}\left(\varepsilon_{t+1}, \xi_{t+1}\right) &= \rho,\\ \text{cov}\left(\varepsilon_{t+1}, \xi_{t+1}\right) &= \Gamma = \rho\sqrt{\Omega}\sqrt{\Lambda},\end{aligned}$$

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

$$\theta = (\Psi + \Gamma)\left(\Psi + \Lambda + 2\Gamma\right)^{-1}, \tag{3.16}$$

and the updating formula for the conditional variance is:

$$\Sigma = \Psi - (\Psi + \Gamma)^2 (\Psi + \Lambda + 2\Gamma)^{-1}, \tag{3.17}$$

where $\Psi = \Omega + A^2\Sigma$. Just like the case without the correlation, given Λ and Γ, (3.16) and (3.17) 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. It seems reasonable to assume that given the same level of capacity the exogenous shock affects both the process (the dynamics of the economy) and the measurement (the perceived/sensed signal). Note that the presence of the correlation between shocks and noises does not affect the conditional variance Σ since $\Sigma = \frac{\Omega}{A^2 - \exp(2\kappa)}$. In the steady state, (3.17) can be rewritten as the following quadratic equation in terms of $\sqrt{\Lambda}$:

$$\left[\rho^2\Omega - (\Psi - \Sigma)\right]\Lambda + 2\rho\Sigma\sqrt{\Omega}\sqrt{\Lambda} + \Sigma\Psi = 0, \tag{3.18}$$

which can be solved for

$$\sqrt{\Lambda} = \frac{-\rho\Sigma\sqrt{\Omega} + \sqrt{\rho^2\Sigma^2\Omega - \Sigma\Psi\left[\rho^2\Omega - (\Psi - \Sigma)\right]}}{\rho^2\Omega - (\Psi - \Sigma)}. \tag{3.19}$$

It is clear from (3.19) 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 the correlated noise does not change the dynamic behavior of the model.

3.2. The Multivariate Case

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

$$v(s_t) = s_t^T P s_t$$

and

$$\widehat{v}(\widehat{s}_t) = \widehat{s}_t^T \widehat{P} \widehat{s}_t$$

respectively.[16] We can compute the optimal Σ by minimizing the expected welfare loss due to RI,

$$E_t \left[v\left(s_t\right) - \widehat{v}\left(\widehat{s}_t\right) \right], \tag{3.20}$$

subject to information-processing constraints. Note that to solve this problem numerically, we need to use a two-stage procedure.[17] First, under the linear-quadratic-Gaussian assumption, the certainty equivalence principle applies and the decision rule under imperfect information,

$$c_t^* = -F\widehat{s}_t, \tag{3.21}$$

is independent of Σ or Λ. We then use this decision rule to determine $\widehat{v}\left(\widehat{s}_t\right)$ which depends on Σ and Λ. Applying the welfare criterion proposed in (3.20), we can solve for optimal steady state Σ and Λ.[18]

Solving the problem proposed in (3.20) is equivalent to solving solve the following semidefinite programming problem:

$$\max_{\Sigma} \left\{ \text{trace}\left(-Z\Sigma\right) \right\} \tag{3.22}$$

subject to

$$-\log\left(|\Sigma|\right) + \log\left(\left|A^T\Sigma A + \Omega\right|\right) \leq 2\kappa, \tag{3.23}$$

$$A^T\Sigma A + \Omega \succeq \Sigma, \tag{3.24}$$

where

$$Z = F^T R F - 2F^T W + \beta\left(F^T B^T P B F + F^T B^T P A + A^T P B F\right).$$

(see Appendix 7.2 for the derivation.) If the positive-definiteness constraint about $A^T\Sigma A + \Omega - \Sigma$, (3.24), does not bind, the first-order condition for Σ can be written as follows:

$$Z = \lambda\left[A\left(A\Sigma A^T + \Omega\right)^{-1} A^T - \Sigma^{-1}\right], \tag{3.25}$$

which can be reduced to:

$$\Sigma^{-1} = \left(G\Sigma G^T + G_0\right)^{-1} - \frac{Z}{\lambda}, \tag{3.26}$$

[16] See also Maćkowiak and Wiederholt (2009).

[17] Sims (2010) also applied this principle solve a tracking problem with information constraints.

[18] Matejka and Sims (2010) show that there exist discrete solutions to the RI problem that may dominate the Gaussian one.

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 (3.26). When applied to a permanent income model in the next section, we first solve this equation and then check whether in fact (3.24) is satisfied by the optimal solution of Σ. If so, the problem is solved.

After computing the optimal steady state Σ, we can then use (3.4) to determine the steady state Λ and (3.5) 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 (3.22) subject to the information-processing constraints.[19] Theoretically, it is clear that after solving an RI problem, we can always reconstruct a 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.[20] Note that 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. 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, (3.4) 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) shows that an estimated SE model does not equate the marginal utility of additional attention across variables, implying that the variance-covariance matrix of the noise would not be consistent with *any* channel capacity.

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 due to a change in policy.[21] In the SE problem with fixed Λ (i.e., under Assumption 1), Equations (3.8) and (3.9) imply that a change of Ω will lead to a different solution for Σ and θ, i.e., the change in policy will lead to a change in the model's dynamics.

[19] 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.

[20] This problem will be particularly difficult for non-LQ Gaussian problems, since the distribution of the noise shocks may be impossible to specify in closed-form.

[21] That is, all elements in the variance-covariance matrix are scaled up.

In contrast, in the multivariate RI problem, as shown in (3.22)-(3.24), 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 under Assumption 2 (i.e., the SNR, $\Omega\Lambda^{-1}$, is fixed). 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 under Assumption 2, we multiply Σ on both sides of (3.8):

$$\Sigma\Lambda^{-1} = I - \left[A\Sigma A^T\Sigma^{-1} + \left(\Omega\Lambda^{-1}\right)\left(\Lambda\Sigma^{-1}\right)\right]^{-1}, \tag{3.27}$$

where I is the identity matrix and we use the fact that $\Omega\Sigma^{-1} = \left(\Omega\Lambda^{-1}\right)\left(\Lambda\Sigma^{-1}\right)$. Under Assumption 2, the policy has the same impact on Ω and Λ to keep the SNR fixed. (3.27) clearly shows that if the policy changes Σ and then $A\Sigma A^T\Sigma^{-1}$, it will affect $\theta = \Sigma\Lambda^{-1}$ even under Assumption 2. Multiplying Ω on both sides of (3.8) gives

$$\Omega\Lambda^{-1} = \Omega\Sigma^{-1} - \left(A\Sigma A^T\Omega^{-1} + I\right)^{-1}, \tag{3.28}$$

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 under Assumption 2 that the SNR, $\Omega\Lambda^{-1}$, is fixed.

4. Applications to the Permanent Income Model

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.[22]

[22] 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 examines how RI can resolve the excess smoothness puzzle and the excess sensitivity puzzle. In another research line, Reis (2006) analyzed optimal consumption decisions of agents facing costs of planning and showed that inattentiveness due to costly planning could be an alternative explanation for the two puzzles. The key feature of Reis' model is that the existence of decision costs induces agents to only infrequently update their decisions. The two modeling strategies have distinct mechanisms of individuals' slow adjustment in consumption but lead to similar aggregate consumption dynamics under some restrictions.

4.1. The Univariate Case

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

$$v(s_0) = \max_{\{c_t\}} E_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t) \right] \tag{4.1}$$

subject to

$$s_{t+1} = Rs_t - c_t + \zeta_{t+1}, \tag{4.2}$$

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

$$s_t = w_t + \frac{1}{R} \sum_{j=0}^{\infty} R^{-j} E_t [y_{t+j}] \tag{4.3}$$

is permanent income, i.e., the expected present value of lifetime resources, consisting of financial wealth (w_t) plus human wealth,

$$\zeta_{t+1} = \frac{1}{R} \sum_{j=t+1}^{\infty} \left(\frac{1}{R}\right)^{j-(t+1)} (E_{t+1} - E_t)[y_j]; \tag{4.4}$$

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 consumer can borrow and lend freely. Note that when y follows an AR(1) process with the persistence coefficient $\rho \in [0,1]$, $y_{t+1} = \rho y_t + \varepsilon_{t+1}$, where $\varepsilon_{t+1} \sim N\left(0, \omega^2\right)$, $\zeta_{t+1} = \varepsilon_{t+1}/(R-\rho)$. For the rest of the paper we will restrict attention to points where $c_t < \overline{c}$, so that utility is increasing and concave.[23] This specification follows that in Hall (1978) and Flavin (1981) and implies that optimal consumption is determined solely by permanent income:

$$c_t = (R-1)\, s_t. \tag{4.5}$$

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

$$c_t = (R-1)\, \widehat{s}_t, \tag{4.6}$$

[23] Here following the literature we impose the restriction that $\beta R = 1$.

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

$$\widehat{s}_{t+1} = (1-\theta)(R\widehat{s}_t - c_t) + \theta\left(s_{t+1} + \xi_{t+1}\right), \tag{4.7}$$

where θ is the Kalman gain, and given $s_0 \sim N\left(\widehat{s}_0, \sigma^2\right)$. As shown in Luo (2008), combining (4.21), (4.6), with (4.7) yields the following expression for the change in consumption:

$$\Delta c_t = (R-1)\left[\frac{\theta\zeta_t}{1-(1-\theta)R\cdot L} + \theta\left(\xi_t - \frac{\theta R\xi_{t-1}}{1-(1-\theta)R\cdot L}\right)\right], \tag{4.8}$$

where L is the lag operator. We require $(1-\theta)R^2 < 1$, the model equivalent of the controllability 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 mainly governed by the Kalman gain θ. Note that using the explicit expression for consumption growth (4.8), we can compute the key stochastic properties of consumption process: the volatility of consumption growth, the persistence of consumption growth, and the correlation between consumption growth and income shocks.[24] And 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.

4.1.1. Policy Implications under SE and RI

In this univariate permanent income model, substituting $A = R$ into Equation (3.11),

$$\omega_\zeta^2\Lambda^{-1} = \omega_\zeta^2\Sigma^{-1} - \left[A^2\left(\omega_\zeta^2\Sigma^{-1}\right)^{-1} + 1\right]^{-1}, \tag{4.9}$$

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 equation about (π, θ):

$$\pi = \theta\left(\frac{1}{1-\theta} - R^2\right), \tag{4.10}$$

which means that

$$\theta = \frac{-(1+\pi) + \sqrt{(1+\pi)^2 + 4R^2(\pi + R^2)}}{2R^2}, \tag{4.11}$$

where we omit the negative values of θ as both Σ and Λ must be positive. Figure 7.1 below 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. Note that $\theta = \Sigma\Lambda^{-1}$ is just the Kalman gain.

[24] See Luo (2008) for a discussion on the effects of RI on consumption dynamics.

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

$$\Sigma = \frac{\Omega}{\exp(2\kappa) - R^2}, \tag{4.12}$$

$$\Lambda = \left(\Sigma^{-1} - \Psi^{-1}\right)^{-1}, \tag{4.13}$$

where $\Psi = R^2\Sigma + \Omega$. Using (4.12) and (4.13), the Kalman filter gain under RI can be written as

$$\theta = \Sigma\Lambda^{-1} = 1 - \frac{1}{\exp(2\kappa)}. \tag{4.14}$$

Comparing (4.11) with (4.14), it is clear that the signal to noise ratio (π) and the level of channel capacity (κ) have one-to-one correspondence. Figure 7.2 shows that 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. This result is consistent with the general conclusion we obtained using Assumption 2 in the previous section.

Using the same expression for μ, (4.11), we can examine how Assumption 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 counter-cyclical policies (i.e., stabilization policies), the signal to noise ratio $\pi = \frac{\omega_\zeta^2}{\Lambda} = 2$. In this case, $\theta = 0.79$. After the government implements these policies, the variance of the shock to permanent income will be reduced from ω_ζ^2 to $0.5\omega_\zeta^2$. Since Λ is fixed under Assumption 1, $\pi = \frac{\omega_\zeta^2}{\Lambda}$ will fall from 2 to 1; consequently, $\theta = 0.74$. We now assume that the RI and SE problems are observationally equivalent in the sense that they lead to the same $\theta = 0.79$ *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.[25] Consequently, the Kalman filter gain, $\theta = \Sigma\Lambda^{-1} = 0.79$, remains unchanged. In other words, the stabilization policies have different implications for consumption dynamics in the SE and RI models. Figure 7.3 plots the different implications of SE and RI for consumption dynamics after implementing the stabilization policies.

[25] A proof is straightforward from Expressions (4.12) and (4.13).

4.1.2. Welfare Effects of Imperfect Information under SE and RI

Given the restriction that $\beta R = 1$, the value function for the RI or SE models is

$$\widehat{v}\left(\widehat{s}_0\right) = -\frac{(R-1)\,R}{2}\widehat{s}_0^2 + R\overline{c}\widehat{s}_0 - \frac{1}{2}R\left(\frac{1}{R-1}\overline{c}^2 + \omega_\eta^2\right), \tag{4.15}$$

where

$$\omega_\eta^2 = \text{var}\left[\eta_{t+1}\right] = \frac{\theta}{1-(1-\theta)\,R^2}\omega_\zeta^2 > \omega_\zeta^2, \tag{4.16}$$

and

$$\eta_{t+1} = \theta\left[\left(\frac{\zeta_{t+1}}{1-(1-\theta)R\cdot L}\right) + \left(\xi_{t+1} - \frac{\theta R\xi_t}{1-(1-\theta)R\cdot L}\right)\right]. \tag{4.17}$$

With the Kalman gain $\theta < 1$, $\omega_\eta^2 > \omega_\zeta^2$ and it then follows that ω_η^2 is decreasing in θ.

Proposition 1. $\frac{\partial\omega_\eta^2}{\partial\theta} < 0$.

Proof. By simple calculation we obtain

$$\frac{\partial\omega_\eta^2}{\partial\theta} = \frac{\left(1-R^2\right)\omega_\zeta^2}{\left[1-(1-\theta)\,R^2\right]^2} < 0 \tag{4.18}$$

because $R > 1$ and $1-(1-\theta)\,R^2 > 0$. ■

The value function (4.15), together with (4.16) and (4.18) clearly show that imperfect observations due to finite capacity lead to more uncertainty about the state, which thus increases welfare losses. More importantly, they also show that after implementing the stabilization policy discussed in the preceding subsection, SE and RI will lead to different welfare implications of imperfect observations. The reason is that the policy will lead to different Kalman gain in the SE and RI problem and thus affects the welfare losses due to imperfect information. Using the same example discussed above, when ω_ζ^2 is reduced to $0.5\omega_\zeta^2$, $\theta = 0.79$ under RI, where it is equal to 0.74 under SE. That is, given the same initial conditions, the stabilization policy will reduce the welfare of agents under SE, whereas it has no impact on agents with RI. The intuition is that the stochastic property of the noise in the RI problem changes accordingly in response to the change in the policy.

Since imperfect information about the state cannot help in decision making, we can use an alternative welfare criterion, the expected welfare gap between the unconstrained value function and the constrained value function conditional on the processed information at the current period, to evaluate different welfare implications of SE and RI. Note that the unconstrained

value function in the full-information case can be written as

$$v(s_t) = -\frac{(R-1)R}{2}s_t^2 + R\overline{c}s_t - \frac{1}{2}R\left(\frac{1}{R-1}\overline{c}^2 + \omega_\zeta^2\right). \tag{4.19}$$

The expected welfare loss due to imperfect information, $\Delta(\Sigma, \theta)$, can thus be written as follows:

$$\begin{aligned} \Delta(\Sigma, \theta) &= E_t\left[v(s_t) - \widehat{v}(\widehat{s}_t)\right] \\ &= -\frac{(R-1)R}{2}\operatorname{var}_t[s_t] - \frac{1}{2}R\left(\omega_\zeta^2 - \omega_\eta^2\right) \\ &= -\frac{(R-1)R}{2}\Sigma + \frac{1}{2}R\left[\frac{\theta}{1-(1-\theta)R^2} - 1\right]\omega_\zeta^2, \end{aligned} \tag{4.20}$$

where the expectation, $E_t[\cdot]$, is conditional on processed information at time t and Σ is the conditional variance of the state. (See Appendix 7.3 for the derivation.) From (4.20), we can clearly see that imperfect information affects the welfare losses via two channels:

1. The post-observation variance, i.e., the conditional variance of the state, Σ. The first term in (4.20) means that the conditional variance will reduce the welfare loss. The intuition behind this result is that $\widehat{s}_t^2$ in the constrained value function is in the time-t information set, while s_t^2 in the constrained value function is not in the information set; consequently, $E_t\left[s_t^2\right] > \widehat{s}_t^2$. However, when R is close to 1, this term is close to 0 and thus has little effect on the welfare losses.

2. The innovation to the level of perceived permanent income, η_{t+1}, is more volatile than that to the level of actual permanent income, ζ_{t+1}. That is, $\omega_\eta^2 > \omega_\zeta^2$. Therefore, the second term in (4.20) means that imperfect information will reduce the welfare loss by increasing the volatility of the innovation to the perceived state. Note that the unconstrained value function is determined by the dynamics of the actual state s_t:

$$s_{t+1} = Rs_t - c_t + \zeta_{t+1}, \tag{4.21}$$

whereas the constrained value function is determined by the dynamics of $\widehat{s}_t$:

$$\widehat{s}_{t+1} = R\widehat{s}_t - c_t + \eta_{t+1}. \tag{4.22}$$

As will be shown later in (4.24), the second channel will dominate the first channel and thus imperfect information always leads to welfare losses.

To evaluate the different effects of SE and RI on the welfare losses under government stabilization policies, we divide ω_ζ^2 on both sides of (4.20) and obtain

$$\frac{\Delta(\Sigma,\theta)}{\omega_\zeta^2} = -\frac{(R-1)R}{2}\frac{\Sigma}{\omega_\zeta^2} + \frac{1}{2}R\left[\frac{\theta}{1-(1-\theta)R^2} - 1\right],$$

which can be rewritten as

$$\widetilde{\Delta}(\pi,K) = \frac{\Delta(\Sigma,\theta)}{\omega_\zeta^2} = -\frac{(R-1)R}{2}\frac{\theta}{\pi} + \frac{1}{2}R\left[\frac{\theta}{1-(1-\theta)R^2} - 1\right], \tag{4.23}$$

where we use the fact that

$$\frac{\Sigma}{\omega_\zeta^2} = \frac{\Sigma}{\Lambda}\frac{\Lambda}{\omega_\zeta^2} = \frac{\theta}{\pi},$$

where $\pi = \omega_\zeta^2/\Lambda$ is defined as the signal-to-noise ratio. Using (4.23), we can examine how stabilization policies affect the welfare losses under SE and RI. Specifically, after implementing the stabilization policy (i.e., ω_ζ^2 is reduced to $0.5\omega_\zeta^2$), SE and RI will lead to different welfare losses via Σ and θ. Under RI, when ω_ζ^2 is reduced to $0.5\omega_\zeta^2$, the conditional variance will fall from Σ to 0.5Σ but $\theta = 0.79$ will remain unchanged. (4.23) therefore implies that $\frac{\Delta(\Sigma,\theta)}{\omega_\zeta^2}$ will remain the same after implementing the policy as RI has no impact on the signal-to-noise ratio π. In contrast, under SE, under Assumption 1 (i.e., Λ is fixed), θ will be reduced to 0.74. (4.23) therefore implies that $\frac{\Delta(\Sigma,\theta)}{\omega_\zeta^2}$ will be reduced after implementing the policy as π will fall to 0.5π. That is, given the same initial conditions ($\theta = 0.79$), the stabilization policy will lead to smaller welfare losses in the SE model than in the RI model. Note that under RI, substituting $\Sigma = \frac{\omega_\zeta^2}{\exp(2\kappa)-R^2} = \frac{\omega_\zeta^2}{1/(1-\theta)-R^2}$, into (4.20), we can further simplify (4.20) as follows:

$$\Delta(\Sigma,K) = \frac{1}{2}\omega_\zeta^2(R-1)\frac{(1-\theta)R^2}{1-(1-\theta)R^2}, \tag{4.24}$$

which means that given ω_ζ^2, the welfare loss is decreasing with the Kalman gain, i.e., is increasing with the degree of inattention (see Appendix 7.3 for the derivation).

Following Barro (2007), we use (4.15) to compute the welfare effects of changes in channel capacity and compare them with those from proportionate changes in the initial level of the perceived state ($\widehat{s}_0$). Specifically, the relative *marginal* welfare losses (rmw) due to imperfect information *at different capacity* (κ) can be written as

$$\text{rmw} = \frac{\partial\widehat{v}/\partial\kappa}{(\partial\widehat{v}/\partial\widehat{s}_0)\,\widehat{s}_0} = -\frac{1}{-(R-1)\,\widehat{s}_0^2 + \overline{c}\widehat{s}_0}\frac{\left(1-R^2\right)\exp(-2\kappa)\,\omega_\zeta^2}{\left[1-\exp(-2\kappa)\,R^2\right]^2} > 0, \tag{4.25}$$

where $\frac{\partial \widehat{v}}{\partial \widehat{s}_0} = -(R-1) R\widehat{s}_0 + R\overline{c} > 0$, $\frac{\partial \widehat{v}}{\partial \kappa} = -\frac{1}{2} R \frac{\partial \omega_\eta^2}{\partial K} \frac{\partial K}{\partial \kappa} > 0$ is evaluated for given $\widehat{s}_0$, Expression (4.25) gives the proportionate increase in $\widehat{s}_0$ that compensates, at the margin, for an reduction in capacity κ devoted to monitoring the state, in the sense of preserving the lifetime utility. Using Expression (4.25), it is straightforward to show that

$$\frac{\partial (\text{rmw})}{\partial \kappa} < 0. \tag{4.26}$$

Denote by $f(\kappa) = \frac{\exp(-2\kappa)}{[1-\exp(-2\kappa)R^2]^2}$. It is clear that only the $f(\kappa)$ term is important for the effects of RI on rmw. Figure 7.4 illustrates how capacity κ affects $f(\kappa)$ and rmw when $R = 1.01$. It clearly shows that RI can have significant effects the relative marginal welfare losses when capacity devoted to monitoring the state is low. For example, when $\kappa = 0.2$ nats, $f = 16.6$, whereas $f = 0.45$ when $\kappa = 1$ nat.

To do quantitative welfare analysis, we need to know the level of the initial level of permanent income, $\widehat{s}_0$. For simplicity we assume that $\widehat{s}_0$ is just mean permanent income. To compute $\widehat{s}_0$, denote by γ the local coefficient of relative risk aversion, which equals

$$\gamma = \frac{E[y]}{\overline{c} - E[y]}$$

for the utility function $u(\cdot)$ evaluated at mean income $E[y]$. Here we impose $\beta = 0.9971$ such that the annual real interest rate is 2.5%. We then follow the procedure used in Hansen and Sargent (2004) and use the estimated one-factor endowment process as follows

$$y_{t+1} = 0.9992 y_t + \varepsilon_{t+1}, \tag{4.27}$$

and ε_{t+1} follows an iid process distributed as $N\left(0, 5.5819^2\right)$. Here we set the coefficient of variation of endowment, $\text{sd}[y_t]/E[y_t]$, to be 0.1, which can be used to compute the mean income level $E[y] = 1396$ and then the value of the bliss point $\overline{c}$ that generates reasonable relative risk aversion γ. For example, when the local CRRA γ is set to 1, we have $\overline{c} = 2E[y] = 2792$. Furthermore, assume that the ratio of mean financial wealth to mean labor income, $E[w]/E[y]$, is 5.[26] Since $s_t = w_t + \frac{1}{R-0.9992} y_t$ and $\zeta_t = \frac{\varepsilon_t}{R-0.9992}$ we have

$$E[s] = \left(5 + \frac{1}{R - 0.9992}\right) E[y].$$

[26] This number varies substantially for different individuals, from 2 to 20. 5 is the average wealth/income ratio in the Survey of Consumer Finances 2001.

Given this specification and set the values of κ and R, we can use (4.25) to compute the welfare effects of finite capacity quantitatively. Figure 7.5 illustrates the values of rmw *at different capacity* for given γ. We can see that rmw is decreasing with κ, i.e., the proportionate increase in $\widehat{s}_0$ that compensates for an reduction in κ in the sense of preserving the lifetime utility is increasing with the degree of inattention.[27] For example, given $\gamma = 1$, when $\kappa = 0.2$ nats, rmw $= 2.7127 \cdot 10^{-2}\%$, whereas rmw $= 9.0105 \cdot 10^{-4}\%$ when $\kappa = 1$ nat. That is, if the agent's capacity is reduced from 1 bit to 0.2 nats, the proportionate increase in $\widehat{s}_0$ that compensates for an reduction in κ in the sense of preserving the expected utility will be increased by about 30 times. Hence, if the level of $\widehat{s}_0$ is large, the agent would have strong incentive to reallocate more capacity to monitor this state if he is allowed to adjust his capacity.

After the government implements the stabilization policies that reduces the variance of the shock from ω_ζ^2 to $0.5\omega_\zeta^2$, the economy switches to a more stable environment. If we relax the assumption that κ is fixed, some capacity will be reallocated to other sources to increase the economic efficiency because a reduction in macroeconomic uncertainty leads to less marginal welfare losses due to RI. In this case, the Kalman gain θ will fall accordingly as it is an increasing function of κ; consequently, the dynamic behavior in the RI model will also change in response to the change in ω_ζ^2.

4.2. The Multivariate Case

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

$$w_{t+1} = Rw_t - c_t + y_{t+1}, \tag{4.28}$$

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

$$y_t = \overline{y} + x_t + z_t + \varepsilon_{y,t}, \tag{4.29}$$

$$x_t = 0.99x_{t-1} + \varepsilon_{x,t}, \tag{4.30}$$

$$z_t = 0.95z_{t-1} + \varepsilon_{z,t}, \tag{4.31}$$

[27] In addition, given κ, rmw is increasing with γ. That is, agents who are more risk averse require more compensation for a reduction in capacity to maintain the initial level of expected utility.

with

$$\Omega = \text{var}\begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \\ \varepsilon_{z,t} \end{bmatrix} = 10^{-3}\begin{bmatrix} 0.9 & 0 & 0 \\ 0 & 0.009 & 0 \\ 0 & 0 & 0.27 \end{bmatrix}, \tag{4.32}$$

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 $u(c_t) = -\frac{1}{2}(c_t - \overline{c})^2$, the model economy can be characterized as the following three equations system:

$$\begin{bmatrix} w_t \\ x_t \\ z_t \end{bmatrix} = \begin{bmatrix} R & 0.99 & 0.95 \\ 0 & 0.99 & 0 \\ 0 & 0 & 0.95 \end{bmatrix}\begin{bmatrix} w_{t-1} \\ x_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}(c_{t-1} - \overline{c}) + \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \varepsilon_{y,t} \\ \varepsilon_{x,t} \\ \varepsilon_{z,t} \end{bmatrix} + \begin{bmatrix} \overline{y} - \overline{c} \\ 0 \\ 0 \end{bmatrix}, \tag{4.33}$$

where β is set to 0.95. Using the first welfare criterion (3.20) provided in Section 3.2, we can compute that

$$\Sigma = 10^{-3}\begin{bmatrix} 0.1399 & -0.0737 & -0.0110 \\ -0.0737 & 0.1596 & -0.1820 \\ -0.0110 & -0.1820 & 0.5555 \end{bmatrix}, \tag{4.34}$$

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 (4.34) that due to the low capacity devoted to monitoring the state, the post-observation variances (i.e., the conditional variances) of the x and z components are both greater than the corresponding innovation variances in (4.32). 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.[28] The intuition behind this result is that the optimizing agent devoted much less capacity to monitoring the slow-moving component, which leads to greater impacts on the conditional variance term. Figure 7.6 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-

[28] Alternatively, we can also see that the conditional variance of the x component is about 3 times smaller than its corresponding unconditional variance (0.4523), whereas that of the z component is about 5 times smaller than its corresponding unconditional variance (2.7692).

moving z component. It is also worth noting that since the agent only cares about the trace of $Z\Sigma$ and the symmetric matrix Z is negative semidefinite, the agent with low capacity will choose to make the post-observations of the states be negatively correlated.

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

$$\Sigma = 10^{-3} \begin{bmatrix} 0.0787 & -0.0419 & 0.0153 \\ -0.0419 & 0.1172 & -0.1926 \\ 0.0153 & -0.1926 & 0.5170 \end{bmatrix}. \tag{4.35}$$

Comparing (4.34) with (4.35), 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 can be reduced to about half. The intuition behind this result is that the endogenous variable plays the most important role in affecting the welfare losses due to RI. 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 (4.34) is optimal in 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 K can be computed. However, it is difficult to specify Λ without prior knowledge about the states. Ad hoc assumptions on Λ might contradict 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 (that is, inconsistent with *any* value for κ.)

5. Extension: Risk-sensitive Filtering under RI and SE

Risk-sensitivity (RS) was first introduced into the LQ-Gaussian framework by Jacobson (1973) and extended by Whittle (1981, 1990). Exploiting the recursive utility framework of Epstein and Zin (1989), Hansen and Sargent (1995) introduce discounting into the RS specification and show that the resulting decision rules are time-invariant. In the RS model agents effectively compute expectations through a distorted lens, increasing their effective risk aversion by over-

weighting negative outcomes. The resulting decision rules depend explicitly on the variance of the shocks, producing precautionary savings, but the value functions are still quadratic functions of the states.[29] In HST (1999) and Hansen and Sargent (2007), they interpret the RS preference in terms of a concern about model uncertainty (robustness or RB) and argue that RS introduces precautionary savings because RS consumers want to protect themselves against model specification errors. Furthermore, when the state cannot be observed perfectly, the classical Kalman filter that minimizes (maximizes) the expected value of a certain quadratic loss (revenue) function is no longer optimal in the risk-sensitive LQ setting. The reason is that in the risk-sensitive LQ problem the objective function is an alternative exponential-quadratic cost function; consequently, the risk-sensitive filter is more suitable than the Kalman filter (i.e., the conditional mean estimator) for estimating the imperfectly observed state. See Whittle (1990), Speyer, Fan, and Banavar (1992), and Banavar and Speyer (1998) for detailed discussions and proofs about risk-sensitive filtering.

In this section we will explore the different implications of SE and RI for economic behavior in the risk-sensitive version of the LQ permanent income model discussed in the last section. The optimization problem of the agent in the full-information version of the risk-sensitive control based on recursive preferences with an exponential certainty equivalence function can be written as

$$v\left(s_t\right) = \min_{c_t} \left\{ s_t^T Q s_t + c_t^T R c_t + 2 c_t^T W s_t + \beta \mathcal{R}_t \left[v\left(s_{t+1}\right) \right] \right\}, \tag{5.1}$$

subject to the same constraint, (2.2), specified in Section 2.1. The distorted expectation operator $\mathcal{R}_t$ is defined by

$$\mathcal{R}_t \left[v\left(s_{t+1}\right) \right] = \frac{1}{\alpha} \log E_t \left[\exp\left(\alpha v\left(s_{t+1}\right) \right) \right], \tag{5.2}$$

where $\alpha > 0$ measures higher risk aversion *vis a vis* the von Neumann-Morgenstern specification; see Hansen and Sargent (1995), Backus, Routledge, and Zin (2004), and Luo and Young (2010) for discussion. However, when the state s cannot be observed perfectly, we need to know the evolution of the perceived state $\widehat{s}$ before we solve for the optimal behavior of the RS agent. Given the risk-sensitivity preference specified in (5.1), rather than minimizing the weighted quadratic sum of the squares of the estimation error (i.e., the Kalman filter) in the LQ setting, in the risk-sensitive LQ setting the agent minimizes the exponential cost criterion

[29] Formally, one can view risk-sensitive agents as ones who have non-state-separable preferences, as in Epstein and Zin (1989), but with a restricted value for the intertemporal elasticity of substitution (see Tallarini 2000).

to obtain the risk-sensitive filter. Specifically, a risk-sensitive filter can be used to solve the following minimization problem

$$\min_{\{s_t\}} \mathcal{C}_t(\alpha) = \min_{\{s_t\}} E\left[\exp\left(\alpha \mathcal{J}_t\right)\right], \tag{5.3}$$

where

$$\mathcal{J}_t = \sum_{t=1}^{N} \left(s_t - \widehat{s}_t\right)^T \left(s_t - \widehat{s}_t\right),$$

where $\alpha > 0$ (< 0) means risk-averse (risk-loving), $\widehat{s}_t$ is the perceived state that is a causal function of the measurement history. The detailed procedure to solve (5.3) has been provided in Speyer, Fan, and Banavar (1992). The following proposition summarizes the main results about the risk-sensitive filter in the permanent income model:

Proposition 1. *Given the constraint specified in (2.2), the evolution of the perceived state $\widehat{s}_t$ follows:*

$$\widehat{s}_t = A\widehat{s}_{t-1} + \theta_t\left(s_t^* - A\widehat{s}_{t-1}\right),$$

where $s_t^* = s_t + \xi_t$ is the noisy signal. The RS filter gain θ can be written as

$$\theta_t = \left(\Psi_t^{-1} + \Lambda_t^{-1}\right)^{-1} \Lambda_t^{-1}, \tag{5.4}$$

and the prior-observation variance Ψ_t is propagated according to

$$\Psi_{t+1} = A\left(\Psi_t^{-1} + \Lambda_t^{-1} - \alpha\right)^{-1} A^T + \Omega. \tag{5.5}$$

For simplicity here we focus on the univariate case.[30] In the univariate RI problem, the post-observation variance Σ_t is determined by channel capacity κ: $\Sigma = \frac{\Omega}{\exp(2\kappa) - R^2}$. Note that given the information-processing constraint, (3.6), the following equation always holds:

$$\Psi = A^2\Sigma + \Omega.$$

In the steady state, (5.5) reduces to

$$\Psi^{-1} + \Lambda^{-1} - \alpha = \Sigma^{-1},$$

[30] Within the risk-sensitive LQG setting, the multivariate RI problem predicts that the agent's preference for risk-sensitivity, budget constraint, and information-processing constraints jointly determine the values of Σ, Λ, and θ, whereas in the multivariate SE problem Λ is independent of the risk-sensitivity preference.

which can be used to determine the value of the variance of the endogenous noise:

$$\Lambda^{-1} = \Sigma^{-1} - \Psi^{-1} + \alpha \tag{5.6}$$

and the risk-sensitive filter gain:

$$\begin{aligned}
\theta &= \left(\Psi^{-1} + \Lambda^{-1}\right)^{-1} \Lambda^{-1}, \\
&= \left(\Sigma^{-1} + \alpha\right)^{-1} \left(\Sigma^{-1} - \Psi_t^{-1} + \alpha\right) \\
&= \frac{\Sigma^{-1} - \left(A^2\Sigma + \Omega\right)^{-1} + \alpha}{\Sigma^{-1} + \alpha}.
\end{aligned}$$

To evaluate the effects of risk-sensitivity and the fundamental uncertainty on the dynamics governed by the risk-sensitive filter gain θ, we rewrite the expression for θ as

$$\theta = \frac{\Omega\Sigma^{-1} - \left(R^2\Sigma\Omega^{-1} + 1\right)^{-1} + \alpha\Omega}{\Omega\Sigma^{-1} + \alpha\Omega} \tag{5.7}$$

Since $\Omega\Sigma^{-1}$ does not change when Ω changes, we have

$$\frac{\partial\theta}{\partial\alpha} = \frac{\Omega\left(R^2\lambda^{-1} + 1\right)^{-1}}{\left(\lambda + \alpha\Omega\right)^2} > 0 \tag{5.8}$$

and

$$\frac{\partial\theta}{\partial\Omega} = \frac{\alpha\left(R^2\lambda^{-1} + 1\right)^{-1}}{\left(\lambda + \alpha\Omega\right)^2} > 0 \text{ if } \alpha > 0. \tag{5.9}$$

Note that as shown in Section 3.1, in the RI-LQ setting, a change in Ω will lead to the same change in Σ, Ψ, and Λ, but has no impact on K, which means that in the LQ setting agents under RI will behave as if facing noise whose nature changes systematically as the dynamic properties of the economy change, i.e., the change in policy does not change the model's dynamics. In contrast, (5.9) clearly shows that the risk-sensitive filter gain that governs the dynamic behavior will change according to (5.7) if there is a change in Ω. Figure 7.7 illustrates how the RS filter gain varies with the degree of RS measured by $\alpha\Omega$ and channel capacity. It clearly shows that the RS gain is increasing with both the risk-sensitivity preference and the degree of attention.[31] For example, given $\kappa = 0.86$ bits, when $\alpha\Omega$ reduces from 0.4 to 0.2, the gain θ reduces from 0.74 to 0.7. Since $\theta = 1 - 1/\exp(2\kappa) = 0.67$ in the LQ model ($\alpha = 0$)

[31] Note that here we impose the restriction that $\alpha\Omega > \left(R^2\Sigma\Omega^{-1} + 1\right)^{-1} - \Omega\Sigma^{-1}$ where $\Omega\Sigma^{-1} = \exp(2\kappa) - R^2$ such that the RS filter gain is positive. This condition is related to the breakdown condition discussed in Hansen and Sargent (2007).

and is independent of changes in Ω, we can see that the dynamic behavior of the economy will change systematically in response to a change in Ω in the RS LQ setting.

In the SE problem, under Assumption 1 (i.e., the variance of the noise Λ is fixed),

$$\Sigma^{-1} - \Psi^{-1} = \Lambda^{-1} - \alpha,$$

can be used to pin down Σ and Ψ. Substituting $\Sigma^{-1} = A^2 (\Psi - \Omega)^{-1}$ into yields

$$A^2 (\Psi - \Omega)^{-1} - \Psi^{-1} = \Lambda^{-1} - \alpha,$$

which can be rewritten as

$$A^2 \left(\Psi\Omega^{-1} - 1\right)^{-1} - \Omega\Psi^{-1} = \Omega\Lambda^{-1} - \alpha\Omega. \tag{5.10}$$

Given Ω and Λ, Equation (5.10) can be used to determine the value of $\Psi\Omega^{-1}$ and Ψ. The risk-sensitive filter gain can then be solved using[32]

$$\begin{aligned} \theta &= \left(\Psi^{-1} + \Lambda^{-1}\right)^{-1} \Lambda^{-1} \\ &= \left(\Omega\Psi^{-1} + \Omega\Lambda^{-1}\right)^{-1} \left(\Omega\Lambda^{-1}\right) \end{aligned} \tag{5.11}$$

In this SE problem, before implementing the policy, we assume that $\Omega\Lambda^{-1} = 2$ and $\alpha\Omega = 0.4$. Using (5.11), we can calculate that $\theta = 0.744$ when $A = 1.02$. After implementing the stabilization policy, $\Omega\Lambda^{-1}$ and $\alpha\Omega$ will be reduced from 2 and 0.4 to 1 and 0.2, respectively. Consequently, the gain falls to 0.64. Therefore, under Assumption 1 and starting from the same state, the SE and RI problems exhibit different dynamic behavior after implementing a stabilization policy that reduces the volatility of the fundamental shock. Under Assumption 2 (i.e., if $\Omega\Lambda^{-1} = 2$ is fixed), using (5.10), we can obtain $\Omega\Psi^{-1} = 0.71$ after implementing the policy that reduces $\alpha\Omega$ from 0.4 to 0.2; consequently, θ is 0.739 which is slightly different from the value of θ before implementing the policy.

In addition, as shown in Luo and Young (2010) in the risk-sensitive LQ control problem, the interactions of RS and imperfect information can also affect decision rules, which further

[32] Note that under Assumption 2, (5.11) can be rewritten as

$$\theta = \frac{\Omega\Sigma^{-1} - \left(R^2\Sigma\Omega^{-1} + 1\right)^{-1} + \alpha\Omega}{\Omega\Sigma^{-1} + \alpha\Omega}$$

even if α is nonzero.

changes the dynamic properties of the model economy. In other words, the RS preference affects both the dynamics of the perceived state and the functional form of optimal decision rules. Specifically, we can formulate an imperfect information version of risk-sensitive control based on recursive preferences with an exponential certainty equivalence function as follows:

$$\widehat{v}\left(\widehat{s}_t\right) = \max_{c_t}\left\{-\frac{1}{2}\left(c_t - \overline{c}\right)^2 + \beta\mathcal{R}_t\left[\widehat{v}\left(\widehat{s}_{t+1}\right)\right]\right\} \tag{5.12}$$

subject to the budget constraint, (4.21), and the filter equation, (4.7). Note that combining with yields the following constraint facing the risk-sensitive inattentive agent:

$$\widehat{s}_{t+1} = R\widehat{s}_t - c_t + \eta_{t+1}. \tag{5.13}$$

The distorted expectation operator is now given by

$$\mathcal{R}_t\left[\widehat{v}\left(\widehat{s}_{t+1}\right)\right] = -\frac{1}{\alpha}\log E_t\left[\exp\left(-\alpha\widehat{v}\left(\widehat{s}_{t+1}\right)\right)\right],$$

where $s_0|\,\mathcal{I}_0 \sim N\left(\widehat{s}_0, \Sigma\right)$, $\widehat{s}_t = E_t\left[s_t\right]$ is the perceived state variable, and η_{t+1} is defined in (4.17). The following proposition summarizes the solution to the RS model when $\beta R = 1$:

Proposition 2. *Given the risk-sensitive filter gain θ and the degree of risk-sensitivity α, the consumption function of a risk-sensitive consumer under RI*

$$c_t = \frac{R-1}{1-\Pi}\widehat{s}_t - \frac{\Pi\overline{c}}{1-\Pi}, \tag{5.14}$$

where

$$\Pi = R\alpha\omega_\eta^2 > 0, \tag{5.15}$$

$$\omega_\eta^2 = \text{var}\left[\eta_{t+1}\right] = \frac{\theta}{1-(1-\theta)R^2}\omega_\zeta^2, \tag{5.16}$$

and $\eta_{t+1} = \theta\left[\frac{\zeta_{t+1}}{1-(1-\theta)RL} + \left(\xi_{t+1} - \frac{\theta R\xi_t}{1-(1-\theta)RL}\right)\right]$.

Proof. See Appendix 7.4. ■

It is clear from (5.14) that the marginal propensity of consumption out of the perceived state $(\widehat{s}_t)$ is affected by the RS preference via both the risk-sensitive control and risk-sensitive filtering. Specifically, the RS preference α affects the form of the consumption function via Π. Second, it also affects ω_η^2 that depends on the risk-sensitive Kalman gain, θ. Using the

same experiment as above, given $\kappa = 0.86$ bits, when $\alpha\Omega$ reduces from 0.4 to 0.2, the RS filter gain θ reduces from 0.74 to 0.7 in the RI problem, whereas it reduces from 0.74 to 0.64 in the SE problem under Assumption 1.[33] Figure 7.8 illuastates different impulse responses of consumption to income shocks under RI and SE after implementing the stabilization policy. It is clear that in the presence of the RS preference, the policy will lead to different consumption dynamics under SE and RI. (Note that before implementing the policy RI and SE generate the same consumption dynamics.)

6. Conclusions

In this paper we have explored the implications of two informational frictions theories, signal extraction and rational inattention, for economic behavior, policy, and welfare within the linear-quadratic-Gaussian (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 signal-to-noise ratio (SNR) in the SE problem is fixed, 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 an efficient way to specify the nature of the Kalman gain that governs the model's dynamics.

[33] Note that under Assumption 2, $\theta = 0.739$ in the SE problem is also significantly different from $\theta = 0.64$ obtained in the RI problem.

7. Appendix

7.1. Deriving Equations (2.16) and (2.17)

Proof. Using (2.14), we have

$$\begin{aligned}
\Sigma_{t+1} &= (I-\theta_t)\,\Psi_t\,(I-\theta_t)^T + \theta_t\Lambda_t\theta_t^T \\
&= \left[I-\Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\right]\Psi_t\left[I-\Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\right]' + \Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\Lambda_t\left[\Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\right]' \\
&= \left[I-\Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\right]\Psi_t\left[I-\left(\left(\Psi_t+\Lambda_t\right)^{-1}\right)'\Psi_t'\right] + \Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\Lambda_t\left[\left(\Psi_t+\Lambda_t\right)^{-1}\right]'\Psi_t' \\
&= \Psi_t - \Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\Psi_t - \Psi_t\left[\left(\Psi_t+\Lambda_t\right)^{-1}\right]'\Psi_t' + \Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\Psi_t\left[\left(\Psi_t+\Lambda_t\right)^{-1}\right]'\Psi_t' \\
&+ \Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\Lambda\left[\left(\Psi_t+\Lambda_t\right)^{-1}\right]'\Psi_t' \\
&= \Psi_t - \Psi_t\left(\Psi_t+\Lambda_t\right)^{-1}\Psi_t\ .
\end{aligned}$$

■

Proof. Using (2.15), we have

$$\begin{aligned}
\theta &= \Psi\left(\Psi+\Lambda\right)^{-1} \\
&= I - \left(\Lambda^{-1}+\Psi^{-1}\right)^{-1}\Psi^{-1} \\
&= I - \Sigma\Psi^{-1} \\
&= \Sigma\Lambda^{-1},
\end{aligned}$$

where the second equality uses the fact that $\left(\Psi+\Lambda\right)^{-1} = \Psi^{-1} - \Psi^{-1}\left(\Lambda^{-1}+\Psi^{-1}\right)^{-1}\Psi^{-1}$ and the third and fourth equalities use the fact that $\Lambda^{-1} = \Sigma^{-1} - \Psi^{-1}$. ■

7.2. Solving for the Steady State Conditional Variance-Covariance Matrix

Assume that the value functions under full information and imperfect information can be written as

$$v\left(s_t\right) = s_t^T P s_t$$

and

$$\widehat{v}\left(\widehat{s}_t\right) = \widehat{s}_t^T \widehat{P}\widehat{s}_t$$

respectively. Note that the two value functions satisfy the following Bellman equations

$$v\left(s_t\right)=s_t^T\left(Q+F^TRF-2F^TW\right)s_t+\beta E_t\left(v\left(s_{t+1}^*\right)\right), \tag{7.1}$$

$$\widehat{v}\left(\widehat{s}_t\right)=E_t\left[s_t^TQs_t+\widehat{s}_t^TF^TRF\widehat{s}_t-2\widehat{s}_t^TF^TWs_t\right]+\beta E_t\left[\widehat{v}\left(\widehat{s}_{t+1}\right)\right], \tag{7.2}$$

where

$$s_{t+1}^*=As_t-BFs_t+\varepsilon_{t+1}$$

is the value of s_{t+1} when the agent can observe s_t perfectly. The agent thus chooses the steady state conditional covariance matrix Σ to minimize the expected welfare loss due to imperfect observations:

$$E_t\left[v\left(s_t\right)\right]-\widehat{v}\left(\widehat{s}_t\right)=E_t\left[s_t^TPs_t\right]-\widehat{s}_t^T\widehat{P}\widehat{s}_t. \tag{7.3}$$

Substituting the two Bellman equations into this objective function gives

$$\begin{aligned}
&E_t\left[s_t^TPs_t\right]-\widehat{s}_t^T\widehat{P}\widehat{s}_t\\
&=E_t\left[s_t^T\left(Q+F^TRF-2F^TW\right)s_t\right]+\beta E_t\left[s_{t+1}^{*T}Ps_{t+1}^*\right]\\
&-E_t\left[s_t^TQs_t+\widehat{s}_t^TF^TRF\widehat{s}_t-2\widehat{s}_t^TF^TWs_t\right]-\beta E_t\left[\widehat{s}_{t+1}^T\widehat{P}\widehat{s}_{t+1}\right]\\
&=E_t\left[s_t^T\left(Q+F^TRF-2F^TW\right)s_t\right]-E_t\left[s_t^TQs_t+\widehat{s}_t^TF^TRF\widehat{s}_t-2\widehat{s}_t^TF^TWs_t\right]\\
&+\beta E_t\left[s_{t+1}^{*T}Ps_{t+1}^*-\widehat{s}_{t+1}^T\widehat{P}\widehat{s}_{t+1}\right]\\
&=E_t\left[s_t^T\left(Q+F^TRF-2F^TW\right)s_t\right]-E_t\left[s_t^TQs_t+\widehat{s}_t^TF^TRF\widehat{s}_t-2\widehat{s}_t^TF^TWs_t\right]\\
&+\beta E_t\left[s_{t+1}^{*T}Ps_{t+1}^*-\widehat{s}_{t+1}^T\widehat{P}\widehat{s}_{t+1}\right]\\
&=\operatorname{trace}\left(\left(F^TRF-2F^TW\right)\Sigma\right)+\beta E_t\left[s_{t+1}^{*T}Ps_{t+1}^*-\widehat{s}_{t+1}^T\widehat{P}\widehat{s}_{t+1}\right]\\
&=\operatorname{trace}\left(\left(F^TRF-2F^TW\right)\Sigma\right)+\beta E_t\left[s_{t+1}^{*T}Ps_{t+1}^*-s_{t+1}^TPs_{t+1}+s_{t+1}^TPs_{t+1}-\widehat{s}_{t+1}^T\widehat{P}\widehat{s}_{t+1}\right]
\end{aligned}$$

Given the LQ setting, $E_t\left[s_t^TPs_t\right]-\widehat{s}_t^T\widehat{P}\widehat{s}_t$ is a constant. Assume that this constant is M. Then, we have

$$M=E_t\left[s_t^TPs_t\right]-\widehat{s}_t^T\widehat{P}\widehat{s}_t=\operatorname{trace}\left(\left(F^TRF-2F^TW\right)\Sigma\right)+\beta E_t\left[s_{t+1}^{*T}Ps_{t+1}^*-s_{t+1}^TPs_{t+1}\right]+\beta M,$$

which means that

$$(1-\beta)M=\operatorname{trace}\left(\left(F^TRF-2F^TW\right)\Sigma\right)+\beta E_t\left[s_{t+1}^{*T}Ps_{t+1}^*-s_{t+1}^TPs_{t+1}\right].$$

Using the definitions of s_{t+1}^* and s_{t+1}, we obtain

$$
\begin{aligned}
(1-\beta)\,M &= \text{trace}\left(\left(F^T RF - 2F^T W\right)\Sigma\right) + \beta E_t \begin{bmatrix} \left(s_t^T - \widehat{s}_t^T\right) F^T B^T PBF\left(s_t - \widehat{s}_t\right) \\ +2\left(s_t^T - \widehat{s}_t^T\right) F^T B^T P\left(As_t - BF\widehat{s}_t + \varepsilon_{t+1}\right) \end{bmatrix} \\
&= \text{trace}\left(\left[F^T RF - 2F^T W + \beta\left(F^T B^T PBF + F^T B^T PA + A^T PBF\right)\right]\Sigma\right) \\
&= \text{trace}\left(Z\Sigma\right),
\end{aligned}
$$

where Z is a constant matrix.

7.3. Deriving the Conditional Welfare Gap

Given (4.15) and (4.19), we have

$$
\begin{aligned}
\Delta\left(\Sigma, K\right) &= E_t\left[v\left(s_t\right) - \widehat{v}\left(\widehat{s}_t\right)\right] \qquad (7.4) \\
&= -\frac{(R-1)\,R}{2} E_t\left[s_t^2\right] + R\overline{c}E_t\left[s_t\right] - \frac{1}{2}R\left(\frac{1}{R-1}\overline{c}^2 + \omega_\zeta^2\right) \\
&\quad - \left\{-\frac{(R-1)\,R}{2}\widehat{s}_t^2 + R\overline{c}\widehat{s}_t - \frac{1}{2}R\left(\frac{1}{R-1}\overline{c}^2 + \omega_\eta^2\right)\right\} \\
&= -\frac{(R-1)\,R}{2}\left(\text{var}_t\left[s_t\right] + \widehat{s}_t^2\right) - \frac{1}{2}R\omega_\zeta^2 - \left[-\frac{(R-1)\,R}{2}\widehat{s}_t^2 - \frac{1}{2}R\omega_\eta^2\right] \\
&= -\frac{(R-1)\,R}{2}\text{var}_t\left[s_t\right] - \frac{1}{2}R\left(\omega_\zeta^2 - \omega_\eta^2\right) \\
&= -\frac{(R-1)\,R}{2}\Sigma + \frac{1}{2}R\left[\frac{K}{1-(1-K)\,R^2} - 1\right]\omega_\zeta^2,
\end{aligned}
$$

which is just (4.20) in the main text. The above expression can be further simplified as follows:

$$
\begin{aligned}
\Delta\left(\Sigma, K\right) &= -\frac{(R-1)\,R}{2}\Sigma - \frac{1}{2}R\left(\omega_\zeta^2 - \omega_\eta^2\right) \\
&= -\frac{(R-1)\,R}{2}\Sigma + \frac{1}{2}R\left[\frac{K}{1-(1-K)\,R^2} - 1\right]\omega_\zeta^2 \\
&= -\frac{(R-1)\,R}{2}\frac{\omega_\zeta^2}{1/\left(1-K\right) - R^2} + \frac{1}{2}R\left[\frac{K}{1-(1-K)\,R^2} - 1\right]\omega_\zeta^2 \\
&= -\frac{1}{2}R\omega_\zeta^2\left[\frac{R-1}{1/\left(1-K\right) - R^2} + \left(1 - \frac{K}{1-(1-K)\,R^2}\right)\right] \\
&= \frac{1}{2}\omega_\zeta^2\left(R-1\right)\frac{(1-K)\,R^2}{1-(1-K)\,R^2},
\end{aligned}
$$

where we use that the fact that $\Sigma = \frac{\omega_\zeta^2}{1/(1-K)-R^2}$ and $\omega_\eta^2 = \frac{K}{1-(1-K)R^2}\omega_\zeta^2$.

7.4. Solving the Risk-sensitive Model with Imperfect Information

To solve the Bellman equation (5.12) subject to (5.13), we conjecture that

$$v\left(\widehat{s}_t\right) = -C - B\widehat{s}_t - A\widehat{s}_t^2, \tag{7.5}$$

where A, B, and C are constants to be determined. We can then evaluate $E_t\left[\exp\left(-\alpha v\left(\widehat{s}_{t+1}\right)\right)\right]$ to obtain

$$\begin{aligned}
&E_t\left[\exp\left(-\alpha v\left(\widehat{s}_{t+1}\right)\right)\right] \\
&= E_t\left[\exp\left(\alpha A\widehat{s}_{t+1}^2 + \alpha B\widehat{s}_{t+1} + \alpha C\right)\right] \\
&= E_t\left[\exp\left(\alpha A\left(R\widehat{s}_t - c_t\right)^2 + \alpha B\left(R\widehat{s}_t - c_t\right) + \left[2\alpha A\left(R\widehat{s}_t - c_t\right) + \alpha B\right]\eta_{t+1} + \alpha A\eta_{t+1}^2 + \alpha C\right)\right] \\
&= (1-2c)^{-1/2}\exp\left(a + \frac{b^2}{2\left(1-2c\right)}\right),
\end{aligned}$$

where

$$\begin{aligned}
a &= \alpha A\left(R\widehat{s}_t - c_t\right)^2 + \alpha B\left(R\widehat{s}_t - c_t\right) + \alpha C, \\
b &= \left[2\alpha A\left(R\widehat{s}_t - c_t\right) + \alpha B\right]\omega_\eta, \\
c &= \alpha A\omega_\eta^2.
\end{aligned}$$

Thus, the distorted expectations operator can be written as

$$\begin{aligned}
\mathcal{R}_t\left[v\left(\widehat{s}_{t+1}\right)\right] &= -\frac{1}{\alpha}\left\{-\frac{1}{2}\log\left(1-2c\right) + a + \frac{b^2}{2\left(1-2c\right)}\right\} \\
&= \frac{1}{2\alpha}\log\left(1-2\alpha A\omega_\eta^2\right) - \frac{A}{1-2\alpha A\omega_\eta^2}\left(R\widehat{s}_t - c_t\right)^2 - \frac{B}{1-2\alpha A\omega_\eta^2}\left(R\widehat{s}_t - c_t\right) - \left[C + \frac{\alpha B^2\omega_\eta^2}{2\left(1-2\alpha A\omega_\zeta^2\right)}\right].
\end{aligned} \tag{7.6}$$

Maximizing the RHS of (5.1) with respect to c_t yields the first-order condition

$$-\left(c_t - \overline{c}\right) + \frac{2\beta A}{1-2\alpha A\omega_\eta^2}\left(R\widehat{s}_t - c_t\right) + \frac{B\beta}{1-2\alpha A\omega_\eta^2} = 0,$$

which means that

$$c_t = \frac{2A\beta R}{1-2\alpha A\omega_\eta^2 + 2A\beta}\widehat{s}_t + \frac{\overline{c}\left(1-2\alpha A\omega_\eta^2\right) + B\beta}{1-2\alpha A\omega_\eta^2 + 2A\beta}. \tag{7.7}$$

Substituting (7.7) and (7.5) into (5.12), and collecting and matching terms, the constant coefficients turn out to be

$$A = \frac{\beta R^2 - 1}{2\beta - 2\alpha\omega_\eta^2}, \tag{7.8}$$

$$B = \frac{\left(\beta R^2 - 1\right)\overline{c}}{(R-1)\left(\alpha\omega_\eta^2 - \beta\right)}, \tag{7.9}$$

$$C = \frac{R\left(\beta R^2 - 1\right)}{2\left(\beta R - R\alpha\omega_\eta^2\right)(R-1)^2}\left((R-1)\,\omega_\eta^2 + \overline{c}^2\right). \tag{7.10}$$

Substituting (7.8) and (7.9) into (7.7) yields the consumption function (5.14) in the text.

References

[1] Adam, Klaus (2005), "Optimal Monetary Policy in the Presence of Imperfect Common Knowledge," *Journal of Monetary Economics* 54(2), 267-301.

[2] Angeletos, George-Marios and Jennifer La'O (2009), "Noisy Business Cycles," forthcoming, *NBER Macroeconomics Annual 2009.*

[3] Backus, David K., Bryan R. Routledge, and Stanley E. Zin (2004), "Exotic Preferences for Macroeconomists," *NBER Macroeconomics Annual 2004*, 319-414.

[4] Banavar, R. N. and J.L. Speyer (1998), "Properties of Risk-sensitive Filters/Estimators," *IEEE* Proceedings.

[5] Barro, Robert. (2007), "On the Welfare Costs of Consumption Uncertainty," manuscript, Harvard University.

[6] Baxter, Brad, Liam Graham, and Stephen Wright (2010), "Invertible and Non-Invertible Information Sets in Dynamic Stochastic General Equilibrium," forthcoming, *Journal of Economic Dynamics and Control.*

[7] Cover, Thomas M. and Joy A. Thomas (1991), *Elements of Information Theory*, John Wiley and Sons.

[8] Flavin, Marjorie A. (1981), "The Adjustment of Consumption to Changing Expectations About Future Income," *Journal of Political Economy,* 89(5): 974-1009.

[9] Friedman, Milton. (1957), *A Theory of the Consumption Function.* Cambridge, MA: National Bureau of Economic Research.

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

[11] Hansen, Lars Peter and Thomas J. Sargent (1995), "Discounted Linear Exponential Quadratic Gaussian Control," *IEEE Transactions on Automatic Control,* 40: 968-71.

[12] Hansen, Lars Peter and Thomas J. Sargent (2004), "'Certainty Equivalence' and 'Model Uncertainty'." http://www.federalreserve.gov/events/conferences/mmp2004/pdf/hansensargent.pdf.

[13] Hansen, Lars Peter and Thomas J. Sargent (2007), *Robustness*, Princeton University Press.

[14] Jacobson, David H. (1973), "Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relation to Deterministic Differential Games," *IEEE Transactions on Automatic Control,* 18: 124-31.

[15] Lorenzoni, Guido (2009), "A Theory of Demand Shocks," *American Economic Review* 99(5), 2050-2084.

[16] Lucas, Robert E. Jr. (1972), "Expectations and the Neutrality of Money," *Journal of Economic Theory* 4, 103-24.

[17] Lucas, Robert E. Jr. (1973), "Some International Evidence on Output-inflation Tradeoff," *American Economic Review* 63(3), 326-334.

[18] Luo, Yulei (2008), "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics,* 11, 366-385.

[19] Luo, Yulei and Eric R. Young (2010), "Risk-sensitive Consumption and Savings under Rational Inattention," *American Economic Journal: Macroeconomics.* 2 (4), 281-325.

[20] Maćkowiak, Bartosz and Mirko Wiederholt (2009), "Optimal Sticky Prices under Rational Inattention," *American Economic Review* 99(3), 769-803.

[21] Matejka, Filip and Christopher A. Sims (2010), "Discrete Actions in Information-Constrained Tracking Problems," manuscript.

[22] Melosi, Leonardo (2009), "A Likelihood Analysis of Models with Information Frictions," Penn Institute for Economic Research Working Paper 09-009.

[23] Morris, Stephen and Hyun Song Shin (2002), "The Social Value of Public Information," *American Economic Review* 92, 1521-34.

[24] Muth, John F. (1960), "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association* 55: 299-306.

[25] Reis, Ricardo (2006). "Inattentive Consumers," *Journal of Monetary Economics*, 53(8): 1761-800.

[26] Sargent, J. Thomas, (1987), *Macroeconomic Theory*, 2nd edition, Academic Press, New York.

[27] Sargent, J. Thomas, (1991), "Equilibrium with Signal Extraction From Endogenous Variables," *Journal of Economic Dynamics and Control* (15), 245-273.

[28] Simon, Dan (2006), *Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches,* John Wiley & Sons.

[29] Sims, Christopher A. (2003), "Implications of Rational Inattention," *Journal of Monetary Economics,* 50, 665-690.

[30] Sims, Christopher A. (2005), "Rational Inattention: A Research Agenda," Technical Report, Princeton University.

[31] Sims, Christopher A. (2006), "Rational Inattention: Beyond the Linear-Quadratic Case," *American Economic Review* 96(2), 158-163.

[32] Sims, Christopher A. (2010), "Rational Inattention and Monetary Economics," forthcoming in *Handbook of Monetary Policy.*

[33] Speyer, J.L, C. Fan and N. Banavar (1992), "Optimal Stochastic Estimation with Exponential Cost Criteria," Proc. 31st *IEEE* CDC, 2293-2298.

[34] Townsend, Robert. M. (1983), "Forecasting the forecasts of others," *Journal of Political Economy* 91, 546-588.

[35] Whittle, Peter (1981), "Risk-Sensitive Linear/Quadratic/Gaussian Control," *Advances in Applied Probability,* 13: 764-77.

[36] Whittle, Peter (1990). *Risk-Sensitive Optimal Control.* John Wiley and Sons.

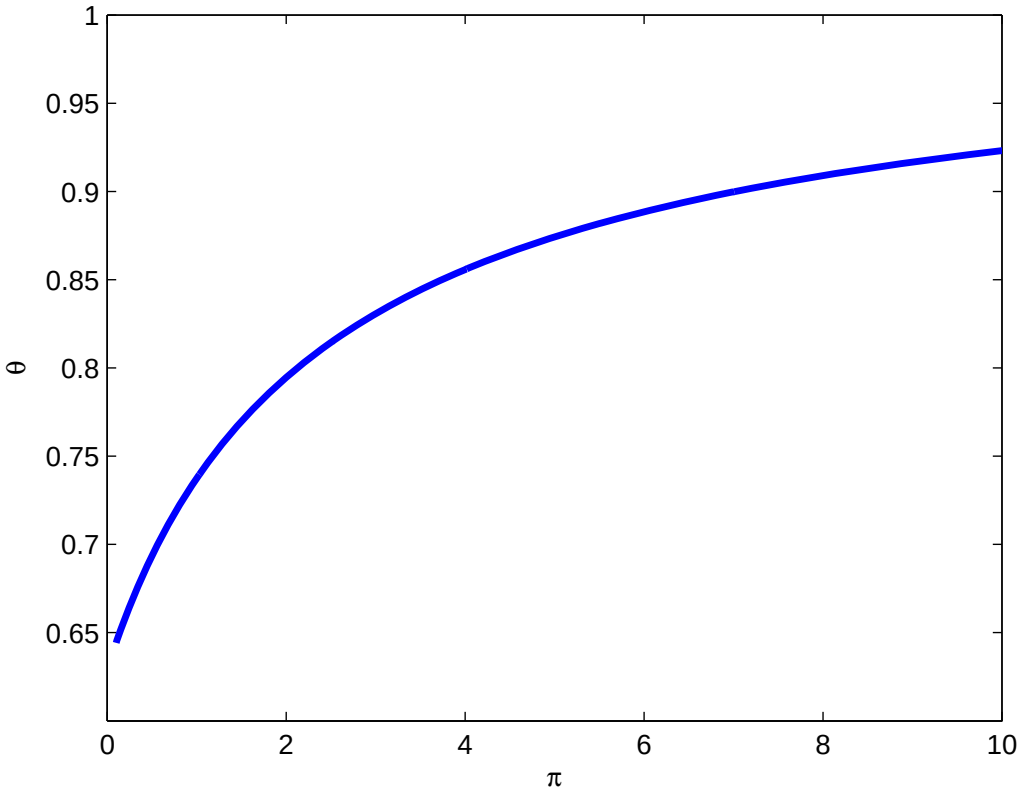

Figure 7.1: Relationship between π and μ

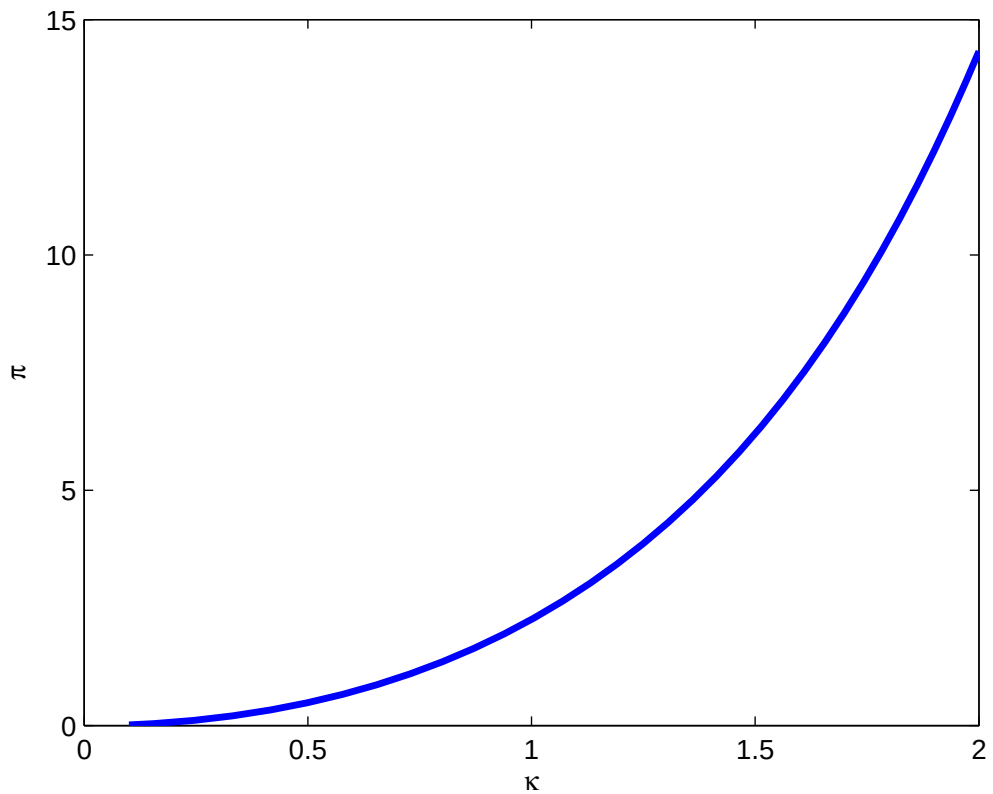

Figure 7.2: Relationship between κ and π

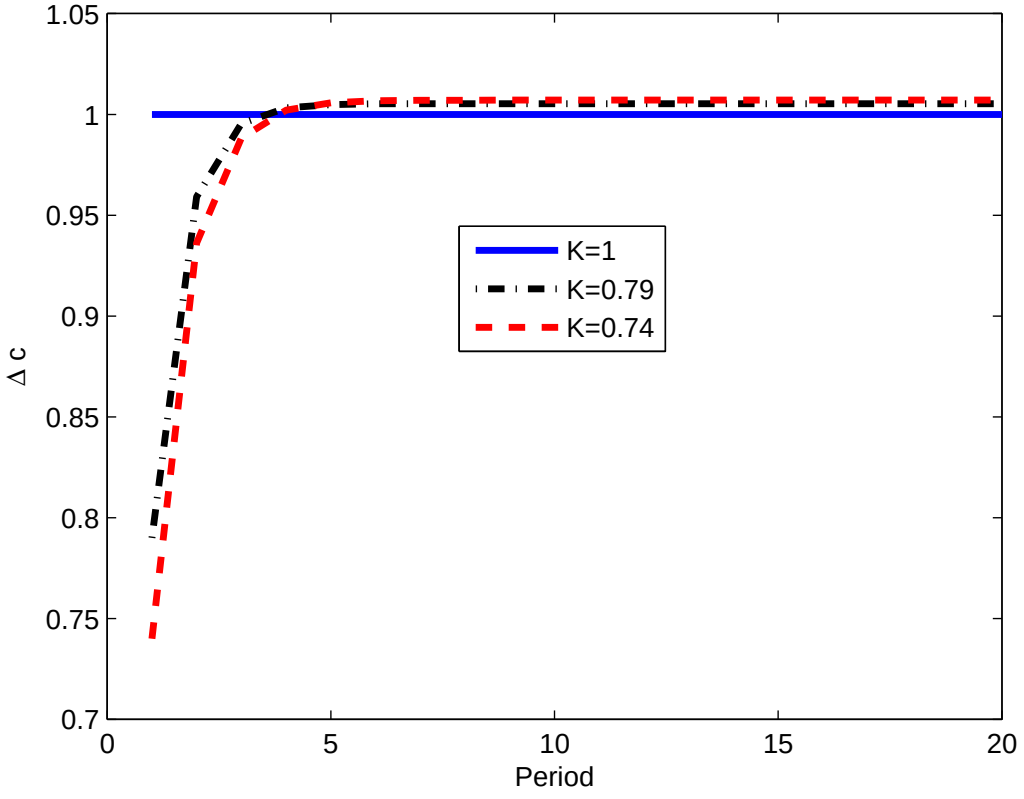

Figure 7.3: Different Consumption Dynamics under SE and RI after Implementing Policy

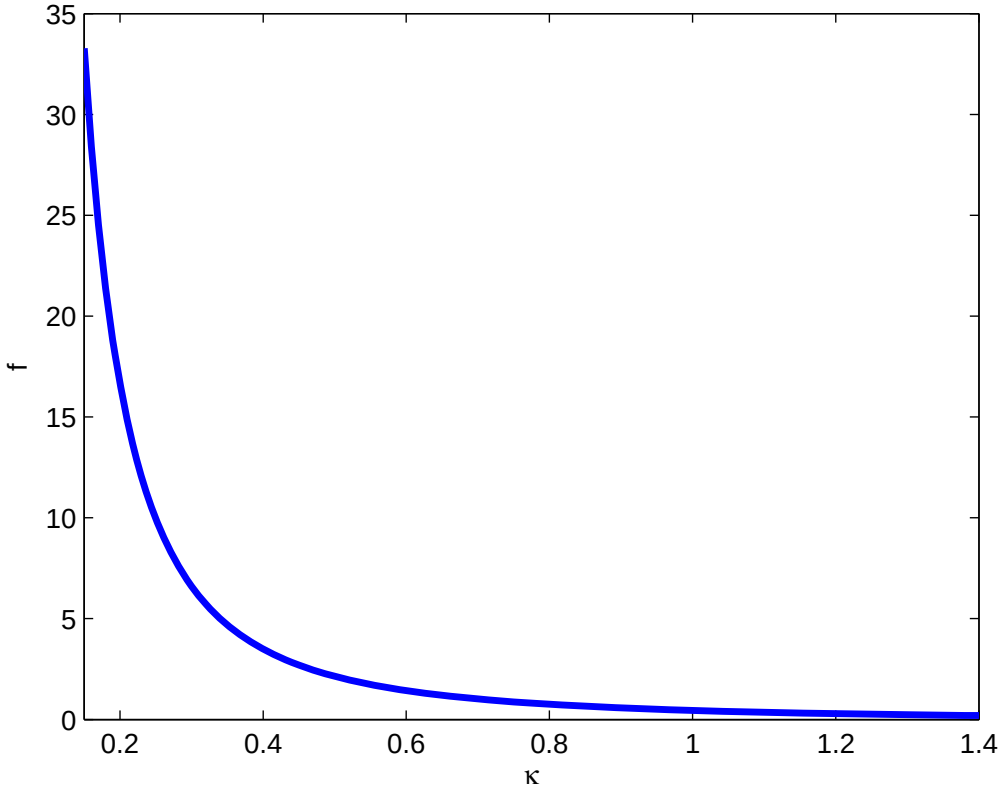

Figure 7.4: The Effects of RI on the Relative Marginal Welfare

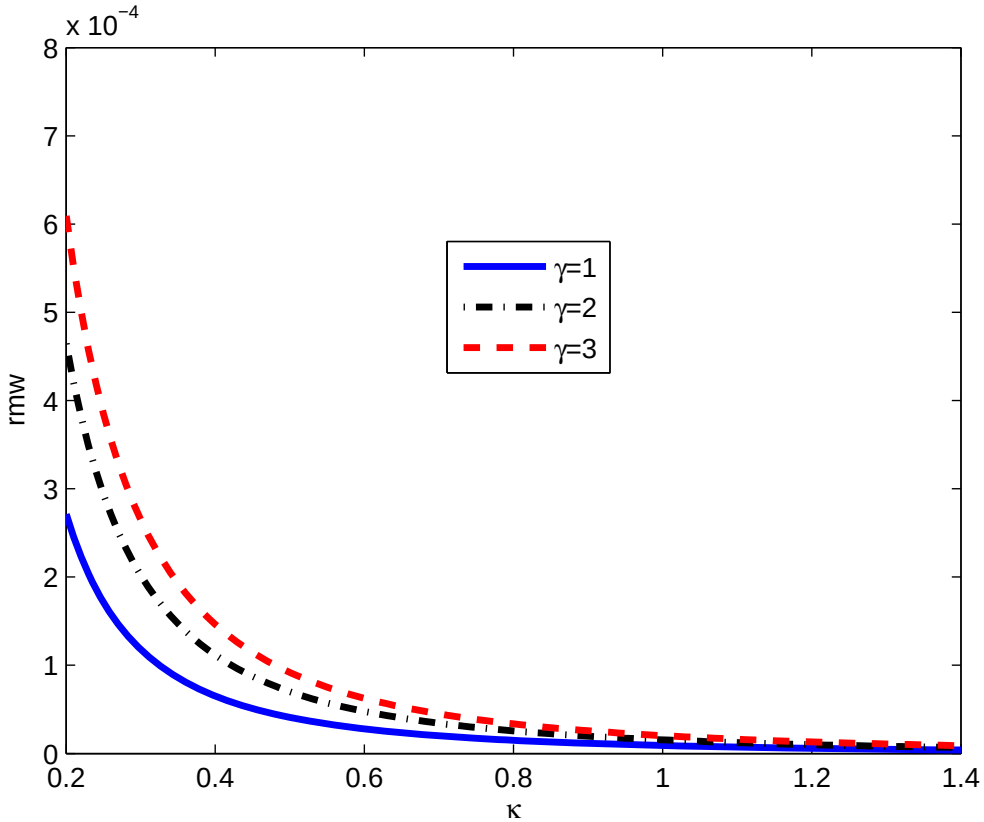

Figure 7.5: The Effects of RI on the Relative Marginal Welfare

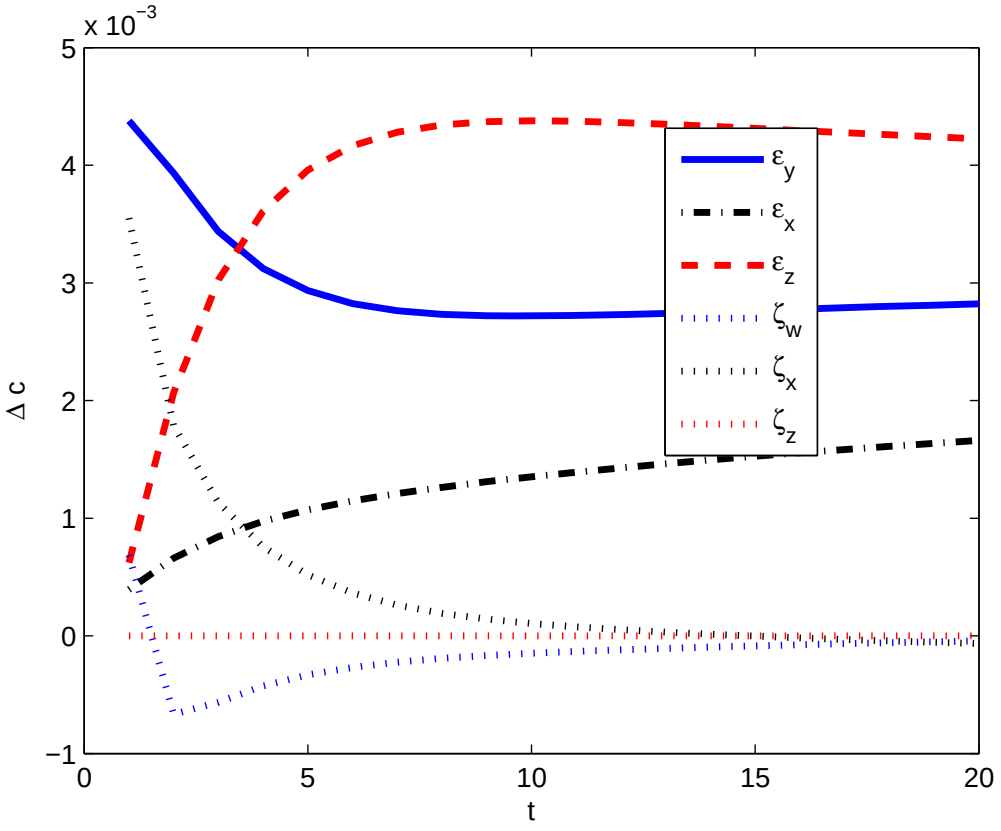

Figure 7.6: Impulse Responses of Consumption to Income Shocks and Noises

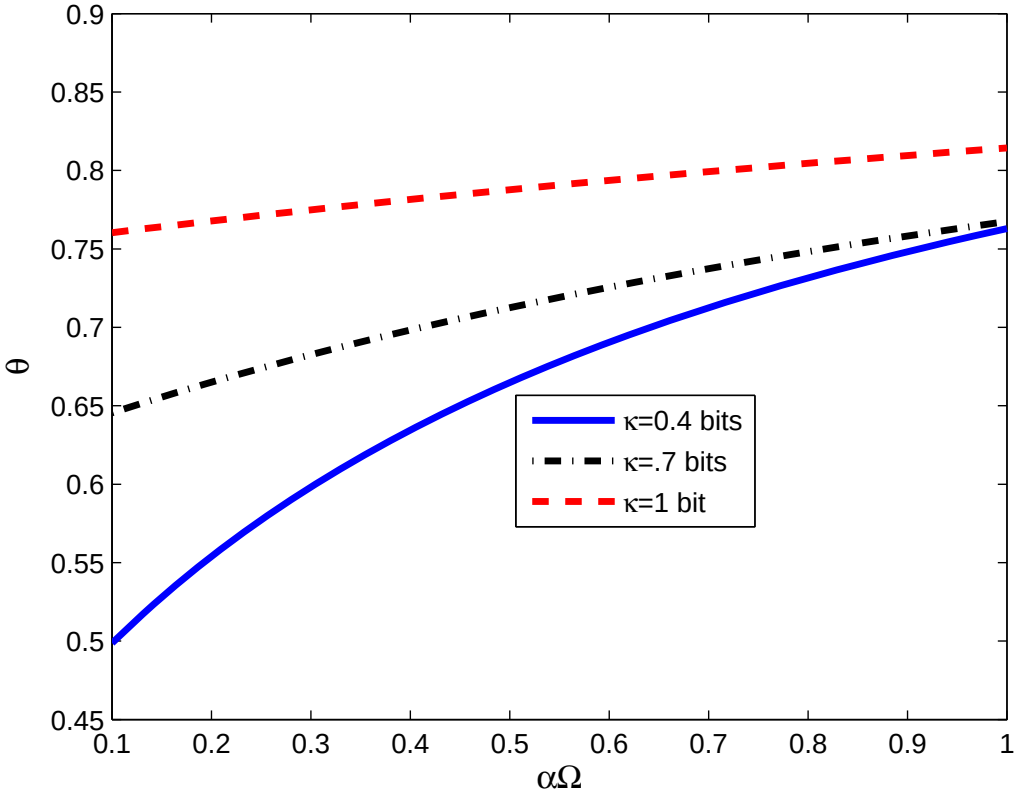

Figure 7.7: The Effects of RS Filtering and RI on the Kalman Gain

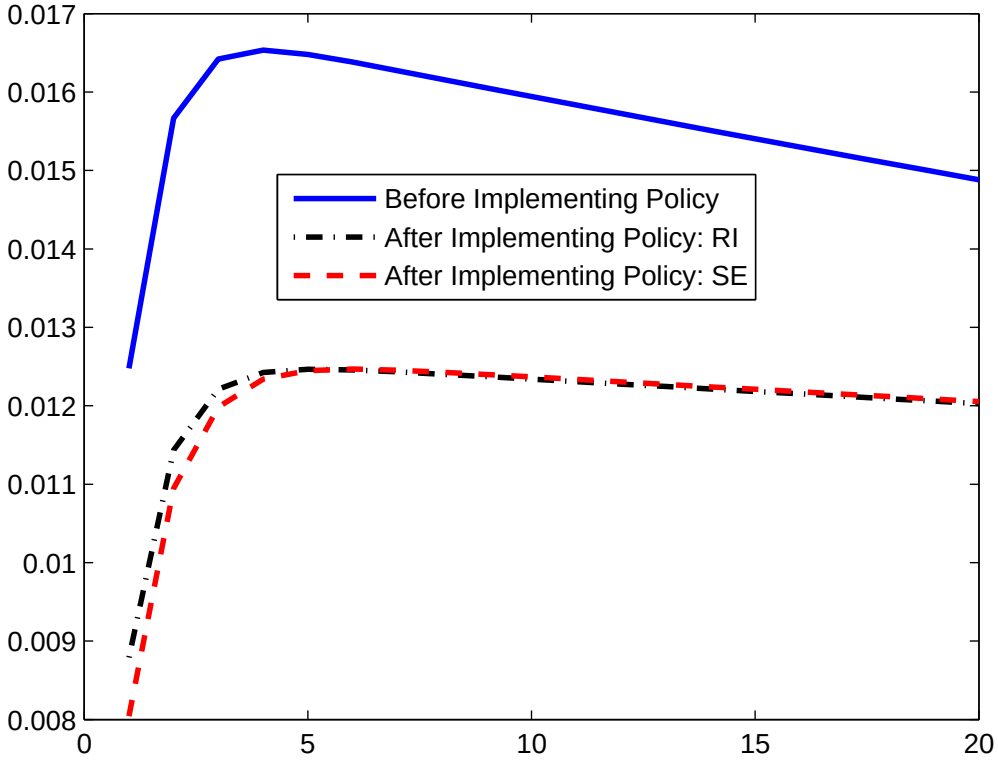

Figure 7.8: The Impulse Responses of Consumption under SE and RI