Model Uncertainty, State Uncertainty, and State-space Models

Yulei Luo, Jun Nie, Eric R. Young

Abstract State-space models have been increasingly used to study macroeconomic and financial problems. A state-space representation consists of two equations, a measurement equation which links the observed variables to unobserved state variables and a transition equation describing the dynamics of the state variables. In this paper, we show that a classic linear-quadratic macroeconomic framework which incorporates two new assumptions can be analytically solved and explicitly mapped to a state-space representation. The two assumptions we consider are the model uncertainty due to concerns for model misspecification (*robustness*) and the state uncertainty due to limited information constraints (*rational inattention*). We show that the state-space representation of the observable and unobservable can be used to quantify the key parameters on the degree of model uncertainty. We provide examples on how this framework can be used to study a range of interesting questions in macroeconomics and international economics.

1 Introduction

State-space models have been broadly applied to study macroeconomic and financial problems. For example, they have been applied to model unobserved trends, to model transition from one economic structure to another, to forecasting models, to study wage-rate behaviors, to estimate expected inflation, and to model time-varying monetary reaction functions.

Yulei Luo
The University of Hong Kong e-mail: yluo@econ.hku.hk

Jun Nie
Federal Reserve Bank of Kansas City, 1 memorial drive, Kansas City, MO, 64196 e-mail: jun.nie@kc.frb.org

Eric R. Young
University of Virginia, Charlottesville, VA 22904 e-mail: ey2d@virginia.edu

A state-space model typically consists of two equations, a measurement equation which links the observed variables to unobserved state variables and a transition equation which describes the dynamics of the state variables. The Kalman filter, which provides a recursive way to compute the estimator of the unobserved component based on the observed variables, is a useful tool to analyze state-space models.

In this paper, we show that a classic linear-quadratic-Gaussian (LQG) macroeconomic framework which incorporates two new assumptions can still be analytically solved and explicitly mapped to a state-space representation.[1] The two assumptions we consider are model uncertainty due to concerns for model misspecification (*robustness*) and state uncertainty due to limited information constraints (*rational inattention*). We show that the state-space representation of the observable and unobservable can be used to quantify the key parameters by simulating the model. We provide examples on how this framework can be used to study a range of interesting questions in macroeconomics and international economics.

The remainder of the paper is organized as follows. Section 2 presents the general framework. Section 3 shows how to introduce the model uncertainty and state uncertainty to this framework. Section 4 provide several applications how to apply this framework to address a range of macroeconomic and international questions. In addition, it shows how this framework has a state-space representation. And this state-space representation can be used to quantify the key parameters in different models. Section 5 concludes.

2 Linear-quadratic-Gaussian State-space Models

The linear-quadratic-Gaussian framework has been widely used in macroeconomics. This specification leads to the optimal linear regulator problem, for which the Bellman equation can be solved easily using matrix algebra. The general setup is as follows. The objective function has a quadratic form,

$$\max_{\{x_t\}} E_0 \left[\sum_{t=0}^{\infty} \beta^t f(x_t) \right] \tag{1}$$

and the maximization is subjected to a linear constraint

$$g(x_t, y_t, y_{t+1}) = 0, \text{ for all } t \tag{2}$$

where $g(\cdot)$ is a linear function, x_t is the vector of control variables and y_t is the vector of state variables.

[1] Note that here "linear" means that the state transition equation is linear, "quadratic" means that the objective function is quadratic, and "Gaussian" means that the exogenous innovation is Gaussian.

Example 1 (*A small-open economy version of Hall's permanent income model*). Let $x_t = \{c_t, b_{t+1}\}$, $y_t = \{b_t, y_t\}$, $f(x_t) = -\frac{1}{2}(\bar{c} - c_t)^2$, $g(x_t, y_t, y_{t+1}) = Rb_t + y_t - c_t - b_{t+1}$, where $\bar{c}$ is the bliss point, c_t is consumption, R is the exogenous and constant gross world interest rate, b_t is the amount of the risk-free foreign bond held at the beginning of period t, and y_t is net income in period t and is defined as output minus investment and government spending. Then this becomes a small-open economy version of Hall's permanent income model in which a representative agent chooses the consumption to maximize his utility subject to the exogenous endowments. As the representative agent can borrow from the rest of the world at a risk-free interest rate, the resource constraint need not bind every period. If we remove this assumption, the model goes back to the permanent income model studied in Hall (1978).[2]

Example 2 (*Barro's tax-smoothing model*). Barro (1979) proposed a simple rational expectations (RE) tax-smoothing model with only noncontingent debt in which the government spreads the burden of raising distortionary income taxes over time in order to minimize their welfare losses to address these questions.[3] This tax-smoothing hypothesis has been widely used (to address various fiscal policies) and tested. The model also falls well into this linear-quadratic framework.[4] Specifically, let $x_t = \{\tau_t, B_{t+1}\}$, $y_t = \{Y_t, G_t\}$, $f(x_t) = -\frac{1}{2}\tau_t^2$, $g(x_t, y_t, y_{t+1}) = RB_t + G_t - \tau_t Y_t - B_{t+1}$, where $E_0[\cdot]$ is the government's expectation conditional on its available and processed information set at time 0, β is the government's subjective discount factor, τ_t is the tax rate, B_t is the amount of government debt, G_t is government spending, Y_t is real GDP, and R is the gross interest rate. Here we assume that the welfare costs of taxation are proportional to the square of the tax rate.[5]

In general, the number of the state variables in these models can be more than one. But in order to facilitate the introduction of robustness we reduce the above multivariate model with a general exogeneous process to a univariate model with iid innovations that can be solved in closed-form. Specifically, following Luo and Young (2010) and Luo, Nie, and Young (2011a), we rewrite the model described by (1) and (2) as

$$\max_{\{z_t, s_{t+1}\}_{t=0}^{\infty}} \left\{ E_0 \left[\sum_{t=0}^{\infty} \beta^t f(z_t) \right] \right\} \tag{3}$$

subject to

$$s_{t+1} = Rs_t - z_t + \zeta_{t+1}, \tag{4}$$

[2] We take a small-open economy version of Hall's model as we'll use it to address some small-open economy issues in later sectors.

[3] It is worth noting that the tax-smoothing hypothesis (TSH) model is an analogy with the permanent income hypothesis (PIH) model in which consumers smooth consumption over time; tax rates respond to permanent changes in the public budgetary burden rather than transitory ones.

[4] For example, see Huang and Lin (1993), Ghosh (1995), and Cashin et al (2001).

[5] Following Barro (1979), Sargent (1987), Bohn (1989), and Huang and Lin (1993), we only need to impose the restriction, $f'(\tau) > 0$ and $f''(\tau) > 0$, on the loss function, $f(\tau)$.

where both z_t and s_t are single variables, and ζ_{t+1} is the Gaussian innovation to the state transition equation with mean 0 and variance ω_ζ^2.

For instance, for Example 1, the mapping is

$$z_t = c_t,$$
$$s_t = b_t + \frac{1}{R}\sum_{j=0}^{\infty} R^{-j} E_t\left[y_{t+j}\right],$$
$$\zeta_{t+1} = \frac{1}{R}\sum_{j=t+1}^{\infty}\left(\frac{1}{R}\right)^{j-(t+1)}\left(E_{t+1}-E_t\right)\left[y_j\right].$$

And for Example 2, the mapping is

$$z_t = \tau_t,$$
$$s_t = E_t\left[b_t + \frac{1}{(1+n)\widetilde{R}}\sum_{j=0}^{\infty}\left(\frac{1}{\widetilde{R}}\right)^j g_{t+j}\right],$$
$$\zeta_{t+1} = \sum_{j=0}^{\infty}\left(\frac{1}{\widetilde{R}}\right)^{j+1}\left(E_{t+1}-E_t\right)\left[g_{t+1+j}\right],$$

where $\widetilde{R} = R/(1+n)$ is the effective interest rate faced by the government, n is the GDP growth rate, b_t and g_t are government debt and government spending as a ratio of GDP.[6]

Finally, the recursive representation of the above problem is as follows.

$$v(s_t) = \max_{z_t}\left\{f(z_t) + \beta E_t\left[v(s_{t+1})\right]\right\} \tag{5}$$

subject to:

$$s_{t+1} = Rs_t - z_t + \zeta_{t+1}, \tag{6}$$

given s_0.

3 Incorporating Model Uncertainty and State Uncertainty

In this section we show how to incorporate model uncertainty and state uncertainty into the framework presented in the previous section.

[6] n is assumed to be constant.

3.1 Introducing Model Uncertainty

We focus on the model uncertainty due to a concern for model misspecification (robustness). Hansen and Sargent (1995, 2007a) first introduce robustness (a concern for model misspecification) into economic models. In robust control problems, agents are concerned about the possibility that their model is misspecified in a manner that is difficult to detect statistically; consequently, they choose their decisions as if the subjective distribution over shocks was chosen by a malevolent nature in order to minimize their expected utility (that is, the solution to a robust decision-maker's problem is the equilibrium of a max-min game between the decision-maker and nature). Specifically, a robustness version of the model represented by (5) and (6) are

$$v(s_t) = \max_{z_t} \min_{\nu_t} \left\{ f(z_t) + \beta \left[\vartheta \nu_t^2 + E_t \left[v(s_{t+1}) \right] \right] \right\} \tag{7}$$

subject to the distorted transition equation (i.e., the worst-case model):

$$s_{t+1} = Rs_t - z_t + \zeta_{t+1} + \omega_\zeta \nu_t, \tag{8}$$

where ν_t distorts the mean of the innovation and $\vartheta > 0$ controls how bad the error can be.[7]

3.2 Introducing State Uncertainty

In this section we introduce state uncertainty into the model we see in the previous section. It will be seen that state uncertainty will further amplify the effect due to model uncertainty.[8] We consider the model with imperfect state observation (*state uncertainty*) due to finite information-processing capacity (rational inattention or RI). Sims (2003) first introduced RI into economics and argued that it is a plausible method for introducing sluggishness, randomness, and delay into economic models. 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.

Under RI, consumers in the economy face both the usual flow budget constraint and information-processing constraint due to finite Shannon capacity first introduced by Sims (2003). As argued by Sims (2003, 2006), individuals with finite

---

[7] Formally, this setup is a game between the decision-maker and a malevolent nature that chooses the distortion process ν_t. $\vartheta \geq 0$ is a penalty parameter that restricts attention to a limited class of distortion processes; it can be mapped into an entropy condition that implies agents choose rules that are robust against processes which are close to the trusted one. In a later section we will apply an error detection approach to calibrate ϑ.

[8] This will be clearer when we go to the applications in later sections.

channel capacity cannot observe the state variables perfectly; consequently, they react to exogenous shocks incompletely and gradually. They need to choose the posterior distribution of the true state after observing the corresponding signal. This choice is in addition to the usual consumption choice that agents make in their utility maximization problem.[9]

Following Sims (2003), the consumer's information-processing constraint can be characterized by the following inequality:

$$\mathscr{H}\left(s_{t+1}|\mathscr{I}_t\right)-\mathscr{H}\left(s_{t+1}|\mathscr{I}_{t+1}\right)\leq\kappa, \tag{9}$$

where κ is the consumer's channel capacity, $\mathscr{H}\left(s_{t+1}|\mathscr{I}_t\right)$ denotes the entropy of the state prior to observing the new signal at $t+1$, and $\mathscr{H}\left(s_{t+1}|\mathscr{I}_{t+1}\right)$ is the entropy after observing the new signal.[10] The concept of *entropy* is from information theory, and it characterizes the uncertainty in a random variable. The right-hand side of (9), being the reduction in entropy, measures the amount of information in the new signal received at $t+1$. Hence, as a whole, (9) means that the reduction in the uncertainty about the state variable gained from observing a new signal is bounded from above by κ. Since the *ex post* distribution of s_t is a normal distribution, $N\left(\widehat{s}_t,\sigma_t^2\right)$, (9) can be reduced to

$$\log|\psi_t^2|-\log|\sigma_{t+1}^2|\leq 2\kappa \tag{10}$$

where $\widehat{s}_t$ is the conditional mean of the true state, and $\sigma_{t+1}^2=\mathrm{var}\left[s_{t+1}|\mathscr{I}_{t+1}\right]$ and $\psi_t^2=\mathrm{var}\left[s_{t+1}|\mathscr{I}_t\right]$ are the posterior variance and prior variance of the state variable, respectively. To obtain (10), we use the fact that the entropy of a Gaussian random variable is equal to half of its logarithm variance plus a constant term.

It is straightforward to show that in the univariate case (10) has a unique steady state σ^2.[11] In that steady state the consumer behaves as if observing a noisy measurement which is $s_{t+1}^*=s_{t+1}+\xi_{t+1}$, where ξ_{t+1} is the endogenous noise and its variance $\alpha_t^2=\mathrm{var}\left[\xi_{t+1}|\mathscr{I}_t\right]$ is determined by the usual updating formula of the variance of a Gaussian distribution based on a linear observation:

$$\sigma_{t+1}^2=\psi_t^2-\psi_t^2\left(\psi_t^2+\alpha_t^2\right)^{-1}\psi_t^2. \tag{11}$$

Note that in the steady state $\sigma^2=\psi^2-\psi^2\left(\psi^2+\alpha^2\right)^{-1}\psi^2$, which can be solved as $\alpha^2=\left[\left(\sigma^2\right)^{-1}-\left(\psi^2\right)^{-1}\right]^{-1}$. Note that (11) implies that in the steady state $\sigma^2=\frac{\omega_\zeta^2}{\exp(2\kappa)-R^2}$ and $\alpha^2=\mathrm{var}\left[\xi_{t+1}\right]=\frac{\left[\omega_\zeta^2+R^2\sigma^2\right]\sigma^2}{\omega_\zeta^2+\left(R^2-1\right)\sigma^2}$.

[9] More generally, agents choose the joint distribution of consumption and current permanent income subject to restrictions about the transition from prior (the distribution before the current signal) to posterior (the distribution after the current signal). The budget constraint implies a link between the distribution of consumption and the distribution of next period permanent income.

[10] We regard κ as a technological parameter. If the base for logarithms is 2, the unit used to measure information flow is a 'bit', and for the natural logarithm e the unit is a 'nat'. 1 nat is equal to $\log_2 e\approx 1.433$ bits.

[11] Convergence requires that $\kappa>\log(R)\approx R-1$; see Luo and Young (2010) for a discussion.

We now incorporate state uncertainty due to RI into the RB model proposed in the last section. There two different ways to do it. The simpler way is to assume that the consumer only has doubts about the process for the shock to permanent income ζ_{t+1}, but trusts his regular Kalman filter hitting the endogenous noise (ξ_{t+1}) and updating the estimated state. In the next subsection, we will relax the assumption that the consumer trusts the Kalman filter equation which generates an additional dimension along which the agents in the economy desire robustness.

The RB-RI model is formulated as

$$\widehat{v}(\widehat{s}_t) = \max_{z_t} \min_{\nu_t} \left\{ f(z_t) + \beta E_t \left[\vartheta \nu_t^2 + \widehat{v}(\widehat{s}_{t+1}) \right] \right\}, \tag{12}$$

subject to the (budget) constraint

$$s_{t+1} = R s_t - z_t + \omega_\zeta \nu_t + \zeta_{t+1} \tag{13}$$

and the regular Kalman filter equation

$$\widehat{s}_{t+1} = (1-\theta)\left(R\widehat{s}_t - z_t + \omega_\zeta \nu_t\right) + \theta\left(s_{t+1} + \xi_{t+1}\right) \tag{14}$$

Notice that $f(z_t)$ is a quadratic function, so the model is in a linear-quadratic form. As to be shown in the next section, we can explicitly solve the optimal choice for control variable z_t and the worst case shock ν_t. After substituting these two solutions into the transition equations for s_t and $\widehat{s}_t$, it can easily be shown that the model has a state-space representation.

3.2.1 Robust filtering under RI

It is clear that the Kalman filter under RI, (13), is not only affected by the fundamental shock (ζ_{t+1}), but also affected by the endogenous noise (ξ_{t+1}) induced by finite capacity; these noise shocks could be another source of the demand for robustness. We therefore need to consider this demand for robustness in the RB-RI model. By adding the additional concern for robustness developed here, we are able to strengthen the effects of robustness on decisions.[12] Specifically, we assume that the agent thinks that (14) is the approximating model. Following Hansen and Sargent (2007), we surround (14) with a set of alternative models to represent a preference for robustness:

$$\widehat{s}_{t+1} = R\widehat{s}_t - z_t + \omega_\eta \nu_t + \eta_{t+1}. \tag{15}$$

where

$$\eta_{t+1} = \vartheta R(s_t - \widehat{s}_t) + \vartheta(\zeta_{t+1} + \xi_{t+1}) \tag{16}$$

and $E_t[\eta_{t+1}] = 0$ because the expectation is conditional on the perceived signals and inattentive agents cannot perceive the lagged shocks perfectly.

[12] Luo, Nie, and Young (2011a) use this approach to study the joint dynamics of consumption, income, and the current account.

Under RI the innovation η_{t+1}, (16), that the agent distrusts is composed of two MA(∞) processes and includes the entire history of the exogenous income shock and the endogenous noise, $\{\zeta_{t+1}, \zeta_t, \cdots, \zeta_0; \xi_{t+1}, \xi_t, \cdots, \xi_0\}$. The difference between (13)) and (15) is the third term; in (13) the coefficient on ν_t is ω_ζ while in (15) the coefficient is ω_η; note that with $\theta < 1$ and $R > 1$ it holds that $\omega_\zeta < \omega_\eta$.

The optimizing problem for this RB-RI model can be formulated as follows:

$$\widehat{v}(\widehat{s}_t) = \max_{c_t} \min_{\nu_t} \left\{ f(z_t) + \beta E_t \left[\vartheta \nu_t^2 + \widehat{v}(\widehat{s}_{t+1}) \right] \right\} \tag{17}$$

subject to (15). (17) is a standard dynamic programming problem and can be easily solved using the standard procedure.

4 Applications

This section provides several applications of the framework developed in Section 3.[13] In each application, the model can be mapped into the general framework presented in the previous section. Using these examples, we show how this framework can be analytically solved and can be explicitly mapped to a state-space representation (Section 4.1). We also show that this state-space representation plays an important role in quantifying the model uncertainty and state uncertainty (Section 4.4). These applications show how model uncertainty (RB) and state uncertainty (RI or imperfect information) alter the results from the standard framework presented in Section 2.

4.1 Explaining Current Account Dynamics

Return in to Example 1 in Section 2. The model is a small-open economy version of the permanent income model. The standard model is represented by (5) and (6), while the model incorporating model uncertainty and state uncertainty is represented by (12)-(14). (Notice that $z_t = c_t$ and $f(x_t) = -\frac{1}{2}(\overline{c} - c_t)^2$.)

As shown in Luo et al (2011a), given ϑ and θ, the consumption function under RB and RI is

$$c_t = \frac{R-1}{1-\Sigma}\widehat{s}_t - \frac{\Sigma \overline{c}}{1-\Sigma}, \tag{18}$$

the mean of the worst-case shock is

$$\omega_\eta \nu_t = \frac{(R-1)\Sigma}{1-\Sigma}\widehat{s}_t - \frac{\Sigma}{1-\Sigma}\overline{c}, \tag{19}$$

[13] These illustrations are based on the research by Luo and Young (2010) and Luo, Nie and Young (2011a, 2011b, 2011c).

where $\rho_s = \frac{1-R\Sigma}{1-\Sigma} \in (0,1)$, $\Sigma = R\omega_\eta^2 / (2\vartheta)$, $\omega_\eta^2 = \mathrm{var}[\eta_{t+1}] = \frac{\theta}{1-(1-\theta)R^2}\omega_\zeta^2$.

Substituting (19) into (13) and combining with (14), the observed s_t and unobserved $\widehat{s}_t$ are governed by the following two equations

$$s_t - \widehat{s}_t = \frac{(1-\theta)\,\zeta_t}{1-(1-\theta)R\cdot L} - \frac{\theta\xi_t}{1-(1-\theta)R\cdot L} \tag{20}$$

$$\widehat{s}_{t+1} = \rho_s \widehat{s}_t + \eta_{t+1}. \tag{21}$$

where

$$\eta_{t+1} = \theta R\,(s_t - \widehat{s}_t) + \theta\,(\zeta_{t+1} + \xi_{t+1}) \tag{22}$$

Thus, it's clear to see that (20) and (21) form a state-space representation the model in which (20) is the measurement equation that links the observed variable s_t to unobserved variable $\widehat{s}_t$ and (21) is the transition equation which describes the dynamics of $\widehat{s}_t$.

Notice that Σ measures the effects of both model uncertainty and state uncertainty, which is bounded by 0 and 1.[14] As argued in Sims (2003), although the randomness in an individual's response to aggregate shocks will be idiosyncratic because it arises from the individual's own information-processing constraint, there is likely a significant common component. 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, $\overline{\xi}_t$, lies between 0 and the part of the idiosyncratic error, ξ_t, caused by the common shock to permanent income, ζ_t. Formally, assume that ξ_t consists of two independent noises: $\xi_t = \overline{\xi}_t + \xi_t^i$, where $\overline{\xi}_t = E^i[\xi_t]$ and ξ_t^i are the common and idiosyncratic components of the error generated by ζ_t, respectively. A single parameter,

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

can be used to measure the common source of coded information on the aggregate component (or the relative importance of $\overline{\xi}_t$ vs. ξ_t).[15]

Next, we briefly list the facts we focus on (Table 1). First, the correlation between the current account and net income is positive but small (and insignificant when detrended with the HP filter). Second, the relative volatility of the current account to net income is smaller in emerging countries than in developed economies, although the difference is not statistically significant when the series are detrended with the HP filter. Third, the persistence of the current account is smaller than that of net income, and less persistent in emerging economies. And fourth, the volatility of consumption growth relative to income growth is larger in emerging economies than in developed economies.

---

[14] See Luo, Nie, and Young (2011a) for the proof.

[15] It is worth noting that the special case that $\lambda = 1$ can be viewed as a representative-agent model in which we do not need to discuss the aggregation issue.

Finally, let's compare the model implications, as summarized in Table 2. First, we have seen that in this case ($\lambda = 1$ and $\theta = 50\%$) the interaction of RB and RI make the model fit the data quite well along dimensions (3) and (4), while also quantitatively improving the model's predictions along dimensions (1) and (2). Second, this improvement does not preclude the model from matching the first two dimensions as well (i.e., the contemporaneous correlation between the current account and net income and the volatility of the current account). For example, holding λ equal to 1 and further reducing θ can generate a smaller contemporaneous correlation between the current account and net income which is closer to the data. And holding $\theta = 50\%$ and reducing λ to 0.1 can make the relative volatility of the current account to net income very close to the data.

4.2 Resolving The International Consumption Puzzle

The same framework can be used to address an old puzzle in the international economics literature. That is, the cross-country consumption correlations are very low in the data (lower than the cross-country correlations of outputs) while standard models imply the opposite.[16]

To show the flexibility of the general framework summarized by (5) and (6), we slightly deviate from the assumption we used in the previous subsection (example 1) to introduce state uncertainty (SU). We assume that consumers in the model economy cannot observe the true state s_t perfectly and only observes the noisy signal

$$s_t^* = s_t + \xi_t, \tag{23}$$

when making decisions, where ξ_t is the iid Gaussian noise due to imperfect observations. The specification in (23) is standard in the signal extraction literature and captures the situation where agents happen or choose to have imperfect knowledge of the underlying shocks.[17] Since imperfect observations on the state lead to welfare losses, agents use the processed information to estimate the true state.[18] Specifically, we assume that households use the Kalman filter to update the perceived state $\widehat{s}_t = E_t[s_t]$ after observing new signals in the steady state:

$$\widehat{s}_{t+1} = (1-\theta)(R\widehat{s}_t - c_t) + \theta(s_{t+1} + \xi_{t+1}), \tag{24}$$

[16] For example, Backus, Kehoe, and Kydland (1992) solve a two-country real business cycles model and argue that the puzzle that empirical consumption correlations are actually lower than output correlations is the most striking discrepancy between theory and data.

[17] For example, Muth (1960), Lucas (1972), Morris and Shin (2002), and Angeletos and La'O (2009). It is worth noting that this assumption is also consistent with the rational inattention idea that ordinary people only devote finite information-processing capacity to processing financial information and thus cannot observe the states perfectly.

[18] See Luo (2008) for details about the welfare losses due to information imperfections within the partial equilibrium permanent income hypothesis framework.

where θ is the Kalman gain (i.e., the observation weight).[19]

In the signal extraction problem, the Kalman gain can be written as

$$\theta = \Upsilon \Lambda^{-1}, \tag{25}$$

where Υ is the steady state value of the conditional variance of s_{t+1}, $\mathrm{var}_{t+1}[s_{t+1}]$, and is the variance of the noise, $\Lambda = \mathrm{var}_t[\xi_{t+1}]$. Υ and Λ are linked by the following equation which updates the conditional variance in the steady state:

$$\Lambda^{-1} = \Upsilon^{-1} - \Psi^{-1}, \tag{26}$$

where Ψ is the steady state value of the *ex ante* conditional variance of s_{t+1}, $\Psi_t = \mathrm{var}_t[s_{t+1}]$.

Multiplying ω_ζ^2 on both sides of (26) and using the fact that $\Psi = R^2\Upsilon + \omega_\zeta^2$, we have

$$\omega_\zeta^2 \Lambda^{-1} = \omega_\zeta^2 \Upsilon^{-1} - \left[R^2\left(\omega_\zeta^2 \Upsilon^{-1}\right)^{-1} + 1\right]^{-1}, \tag{27}$$

where $\omega_\zeta^2 \Upsilon^{-1} = \left(\omega_\zeta^2 \Lambda^{-1}\right)\left(\Lambda \Upsilon^{-1}\right)$.

Define SNR as $\pi = \omega_\zeta^2 \Lambda^{-1}$. We obtain the following equality linking SNR (π) and the Kalman gain (θ):

$$\pi = \theta\left(\frac{1}{1-\theta} - R^2\right). \tag{28}$$

Solving for θ from the above equation yields

$$\theta = \frac{-(1+\pi) + \sqrt{(1+\pi)^2 + 4R^2(\pi + R^2)}}{2R^2}, \tag{29}$$

where we omit the negative values of θ because both Υ and Λ must be positive. Note that given π, we can pin down Λ using $\pi = \omega_\zeta^2 \Lambda^{-1}$ and Υ using (25) and (29).

Combining (4) with (24), we obtain the following equation governing the perceived state $\widehat{s}_t$:

$$\widehat{s}_{t+1} = R\widehat{s}_t - c_t + \eta_{t+1}, \tag{30}$$

where

$$\eta_{t+1} = \theta R(s_t - \widehat{s}_t) + \theta(\zeta_{t+1} + \xi_{t+1}) \tag{31}$$

is the innovation to the mean of the distribution of perceived permanent income,

$$s_t - \widehat{s}_t = \frac{(1-\theta)\zeta_t}{1-(1-\theta)R\cdot L} - \frac{\theta\xi_t}{1-(1-\theta)R\cdot L} \tag{32}$$

[19] Note that θ measures how much uncertainty about the state can be removed upon receiving the new signals about the state.

is the estimation error where L is the lag operator, and $E_t[\eta_{t+1}] = 0$.Note that η_{t+1} can be rewritten as

$$\eta_{t+1} = \theta \left[\left(\frac{\zeta_{t+1}}{1-(1-\theta)R \cdot L} \right) + \left(\xi_{t+1} - \frac{\theta R \xi_t}{1-(1-\theta)R \cdot L} \right) \right], \quad (33)$$

where $\omega_\xi^2 = \mathrm{var}[\xi_{t+1}] = \frac{1}{\theta}\frac{1}{1/(1-\theta)-R^2}\omega_\zeta^2$. Expression (33) clearly shows that the estimation error reacts to the fundamental shock positively, while it reacts to the noise shock negatively. In addition, the importance of the estimation error is decreasing with θ. More specifically, as θ increases, the first term in (33) becomes less important because $(1-\theta)\zeta_t$ in the numerator decreases, and the second term also becomes less important because the importance of ξ_t decreases as θ increases.[20]

Although the assumption we use to introduce state uncertainty is different, the general framework is still the same. More importantly, the solution strategy is also the same. Basically, we can explicitly derive the expressions for consumption and the worst-case shock and then substitute them into (30). Together with (32), it forms a state-space representation of the model.

Table 3 reports the implied consumption correlations (between the domestic country and ROW) between the RE, RB, and RB-SU models. There are two interesting observations in the table. First, given the degrees of RB and SU (θ), $\mathrm{corr}(c_t, c_t^*)$ decreases with the aggregation factor (λ). Second, when λ is positive (even if it is very small, e.g., 0.1 in the table), $\mathrm{corr}(c_t, c_t^*)$ is decreasing with the degree of inattention (i.e., increasing with θ). The intuition is that when there are common noises, the effect of the noises could dominate the effect of gradual consumption adjustments on cross-country consumption correlations.

As we can see from Table 3, for all the countries we consider here, introducing SU into the RB model can make the model better fit the data on consumption correlations at many combinations of the parameter values. For example, for Italy, when $\theta = 60\%$ (60% of the uncertainty is removed upon receiving a new signal about the innovation to permanent income) and $\lambda = 1$, the RB-SU model predicts that $\mathrm{corr}(c_t, c_t^*) = 0.27$, which is very close to the empirical counterpart, 0.25.[21] For France, when $\theta = 90\%$ and $\lambda = 0.5$, the RB-SU model predicts that $\mathrm{corr}(c_t, c_t^*) = 0.46$, which exactly matches the empirical counterpart. Note that a small value of θ can be rationalized by examining the welfare effects of finite channel capacity.[22]

---

[20] Note that when $\theta = 1$, $var[\xi_{t+1}] = 0$.

[21] For example, Adam (2005) found $\theta = 40\%$ based on the response of aggregate output to monetary policy shocks. Luo (2008) found that if $\theta = 50\%$, the otherwise standard permanent income model can generate realistic relative volatility of consumption to labor income.

[22] See Luo and Young (2010) for details about the welfare losses due to imperfect observations in the RB model; they are uniformly small.

4.3 Other Possible Applications

This linear-quadratic framework which incorporates model uncertainty (due to RB) and state uncertainty (either due to RI or imperfect information) can be applied to study other topics as well. We will briefly discuss several more in this subsection. We will not write down the model equations again as we have shown in Section 2 and 3 that these models can be written in a similar framework.

First, as shown in the previous section, model uncertainty due to RB is particularly promising and interesting for studying emerging and developed small-open economies because it has the potential to generate the *different* joint behaviors of consumption and current accounts observed *across the two groups of economies*. This novel theoretical contribution can also be used to address the observed U.S. Great Moderation in which the volatility of output changed after 1984. Specifically, this feature can be used to address different macroeconomic dynamics (e.g., consumption volatility) given that output volatility changed *before and after* the Great Moderation.

Second, inventories in the standard production smoothing model can be viewed as a stabilizing factor. Cost-mininizing firms facing sales fluctuations smooth production by adjusting their inventories. As a result, production is less volatile than sales. However, in the data, real GDP is more volatile than final sales measured by real GDP minus inventory investment. The existing studies find supportive evidence that real GNP is more volatile than final sales in industry-level data. The key question is that if cost-minimizing firms use inventories to smooth their production, why is production more volatile than sales? In the future research, we can examine whether introducing RB can help improve the prediction of an otherwise standard production smoothing model with inventories on the joint dynamics of inventories, production, and sales.

Third, as shown in Luo, Nie, and Young (2011c), the standard tax-smoothing model proposed by Barro (1979) cannot explain the observed volatility of the tax rates and the joint behavior of the government spending and deficits. As shown in Example 2 of Section (2), the tax-smoothing model used in the literature falls well into the linear-quadratic framework we described. It's easy to show that the same mechanisms presented in Section 4.1 and 4.2 will work in the tax-smoothing model which incorporates model uncertainty and state uncertainty. Specifically, Luo, Nie, and Young (2011c) shows that it can help the standard model to better explain the relative volatility of the changes in tax rates to government spending and the co-movement between government deficits and spending in the data.

Fourth, this framework can also be extended to study optimal monetary policy under model uncertainty and imperfect state observation. A central bank sets nominal interest rate to minimize prices fluctuations and the output gap (i.e., the deviation of the output from the potential maximum output level). Following the literature, the standard objective function of a central bank can be described by a quadratic function which is a weighted average of the deviation of the inflation from its target and the output gap.[23] Therefore, the framework presented in this paper can be used to study optimal monetary policy when a central bank has concerns that the model is misspecified and it faces noisy data when making decisions.[24]

4.4 Quantifying Model Uncertainty

One remaining question from previous sections is how to quantify the incorporated degree of model uncertainty.[25] In this section, we will show how to use the state-space representation of s_t and $\widehat{s_t}$ to simulate the model and calibrate the key parameters. For convenience and consistence, we continue to use the small-open economy model described in Example 1 as the illustration example.

Let model A denote the approximating model and model B be the distorted model. Define p_A as

$$p_A = \text{Prob}\left(\log\left(\frac{L_A}{L_B}\right) < 0 \middle| A \right), \tag{34}$$

where $\log\left(\frac{L_A}{L_B}\right)$ is the log-likelihood ratio. When model A generates the data, p_A measures the probability that a likelihood ratio test selects model B. In this case, we call p_A the probability of the model detection error. Similarly, when model B generates the data, we can define p_B as

$$p_B = \text{Prob}\left(\log\left(\frac{L_A}{L_B}\right) > 0 \middle| B \right). \tag{35}$$

Following Hansen, Sargent, and Wang (2002) and Hansen and Sargent (2007b), the detection error probability, p, is defined as the average of p_A and p_B:

$$p(\vartheta) = \frac{1}{2}(p_A + p_B), \tag{36}$$

where ϑ is the robustness parameter used to generate model B. Given this definition, we can see that $1 - p$ measures the probability that econometricians can distinguish the approximating model from the distorted model.

Now we show how to compute the model detection error probability due to model uncertainty and state uncertainty.

---

[23] For example, see Svensson (2000), Gali and Monacelli (2005), Walsh (2005), Leitemo and Soderstrom (2008a,b).

[24] For the examples of the model equations describing the inflation and output dynamics in a closed economy, see Leitemo and Soderstrom (2008a).

[25] This includes the two versions of the model presented in previous sections which incorporates the model uncertainty due to RB: one uses the regular Kalman filter; the other one assumes that the agent does not trust the Kalman filter either (robust filtering).

In the model with both the RB preference and RI, the approximating model can be written as

$$s_{t+1} = Rs_t - c_t + \zeta_{t+1}, \tag{37}$$
$$\widehat{s}_{t+1} = (1-\theta)(R\widehat{s}_t - c_t) + \theta(s_{t+1} + \lambda\xi_{t+1}), \tag{38}$$

and the distorted model is

$$s_{t+1} = Rs_t - c_t + \zeta_{t+1} + \omega_\zeta \nu_t, \tag{39}$$
$$\widehat{s}_{t+1} = (1-\theta)\left(R\widehat{s}_t - c_t + \omega_\zeta \nu_t\right) + \theta(s_{t+1} + \lambda\xi_{t+1}), \tag{40}$$

where we remind the reader that $\lambda = \frac{\mathrm{var}[\overline{\xi}_t]}{\mathrm{var}[\xi_t]} \in [0,1]$ is the parameter measuring the relative importance of $\overline{\xi}_t$ vs. ξ_t.

After substituting the consumption function and the worst-case shock expression into (38) and (40) we can put the equations in the following matrix form:

$$\begin{bmatrix} s_{t+1} \\ \widehat{s}_{t+1} \end{bmatrix} = \begin{bmatrix} R & -\frac{R-1}{1-\Sigma} \\ \theta R & \frac{1-R+R(1-\theta)(1-\Sigma)}{1-\Sigma} \end{bmatrix} \begin{bmatrix} s_t \\ \widehat{s}_t \end{bmatrix} + \begin{bmatrix} \zeta_{t+1} \\ \theta(\zeta_{t+1} + \lambda\xi_{t+1}) \end{bmatrix} + \begin{bmatrix} \frac{\Sigma}{1-\Sigma}\overline{c} \\ \frac{\Sigma}{1-\Sigma}\overline{c} \end{bmatrix} \tag{41}$$

and

$$\begin{bmatrix} s_{t+1} \\ \widehat{s}_{t+1} \end{bmatrix} = \begin{bmatrix} R & -(R-1) \\ \theta R & 1-\theta R \end{bmatrix} \begin{bmatrix} s_t \\ \widehat{s}_t \end{bmatrix} + \begin{bmatrix} \zeta_{t+1} \\ \theta(\zeta_{t+1} + \lambda\xi_{t+1}) \end{bmatrix}. \tag{42}$$

Given the RB parameter, ϑ, and RI parameter, θ, we can compute p_A and p_B and thus the detection error probability as follows.

1. Simulate $\{s_t\}_{t=0}^T$ using (41) and (42) a large number of times. The number of periods used in the simulation, T, is set to be the actual length of the data for each individual country.
2. Count the number of times that $\log\left(\frac{L_A}{L_B}\right) < 0 \Big| A$ and $\log\left(\frac{L_A}{L_B}\right) > 0 \Big| B$ are each satisfied.
3. Determine p_A and p_B as the fractions of realizations for which $\log\left(\frac{L_A}{L_B}\right) < 0 \Big| A$ and $\log\left(\frac{L_A}{L_B}\right) > 0 \Big| B$, respectively.

4.5 Discussions: Risk-sensitivity and Robustness under Rational Inattention

Risk-sensitivity (RS) was first introduced into the LQG 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 overweighting 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.[26] 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.

Following Luo and Young (2010), we formulate an RI version of risk-sensitive control based on recursive preferences with an exponential certainty equivalence function as follows:

$$\widehat{v}(\widehat{s}_t) = \max_{c_t} \left\{ -\frac{1}{2}(c_t - \overline{c})^2 + \beta \mathscr{R}_t [\widehat{v}(\widehat{s}_{t+1})] \right\} \tag{43}$$

subject to the budget constraint (6) and the Kalman filter equation 14. The distorted expectation operator is now given by

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

where $s_0 | \mathscr{I}_0 \sim N(\widehat{s}_0, \overline{\sigma}^2)$, $\widehat{s}_t = E_t[s_t]$ is the perceived state variable, θ is the optimal weight on the new observation of the state, and ξ_{t+1} is the endogenous noise. The optimal choice of the weight θ is given by $\theta(\kappa) = 1 - 1/\exp(2\kappa) \in [0,1]$. The following proposition summarizes the solution to the RI-RS model when $\beta R = 1$:

Proposition 1. *Given finite channel capacity κ and the degree of risk-sensitivity α, the consumption function of a risk-sensitive consumer under RI*

$$c_t = \frac{R-1}{1-\Pi}\widehat{s}_t - \frac{\Pi \overline{c}}{1-\Pi}, \tag{44}$$

where

$$\Pi = R\alpha\omega_\eta^2 \in (0,1), \tag{45}$$

$$\omega_\eta^2 = \operatorname{var}[\eta_{t+1}] = \frac{\theta}{1-(1-\theta)R^2}\omega_\zeta^2, \tag{46}$$

η_{t+1} is defined in (16), and $\theta(\kappa) = 1 - 1/\exp(2\kappa)$.

Comparing (18) and (44), it is straightforward to show that it is impossible to distinguish between RB and RS under RI using only consumption-savings decisions.

Proposition 2. *Let the following expression hold:*

$$\alpha = \frac{1}{2\vartheta}. \tag{47}$$

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

Then consumption and savings are identical in the RS-RI and RB-RI models.

Note that (47) is *exactly the same as* the observational equivalence condition obtained in the full-information RE model (see Backus, Routledge, and Zin 2004). That is, under the assumption that the agent distrusts the Kalman filter equation, the OE result obtained under full-information RE still holds under RI.[27]

HST (1999) show that as far as the *quantity* observations on consumption and savings are concerned, the robustness version ($\vartheta > 0$ or $\alpha > 0, \widetilde{\beta}$) of the PIH model is observationally equivalent to the standard version ($\vartheta = \infty$ or $\alpha = 0, \beta = 1/R$) of the PIH model for a unique pair of discount factors.[28] The intuition is that introducing a preference for risk-sensitivity (RS) or a concern about robustness (RB) increases savings in the same way as increasing the discount factor, so that the discount factor can be changed to offset the effect of a change in RS or RB on consumption and investment.[29] Alternatively, holding all parameters constant except the pair (α, β), the RI version of the PIH model with RB consumers ($\vartheta > 0$ and $\beta R = 1$) is observationally equivalent to the standard RI version of the model ($\vartheta = \infty$ and $\widetilde{\beta} > 1/R$).

Proposition 3. *Let*

$$\widetilde{\beta} = \frac{1}{R}\frac{1 - R\omega_\eta^2/(2\vartheta)}{1 - R^2\omega_\eta^2/(2\vartheta)} = \frac{1}{R}\frac{1 - R\alpha\omega_\eta^2}{1 - R^2\alpha\omega_\eta^2} > \frac{1}{R}.$$

Then consumption and savings are identical in the RI, RB-RI, and RS-RI models.

5 Conclusions

In this paper we show that a state-space representation can be explicitly derived from a classic macroeconomic framework which has incorporated model uncertainty due to concerns for model misspecification (robustness or RB) and state uncertainty due to limited information constraints (rational inattention or RI). We show the state-space representation can also be used to quantify the key model parameters. Several applications are also provided to show how this general framework can be used to address a range of interesting economic questions.

---

[27] Note that the OE becomes

$$\frac{\alpha\theta}{1-(1-\theta)R^2} = \frac{1}{2\vartheta},$$

if we assume that the agents distrust the income process hitting the budget constraint, but trust the RI-induced noise hitting the Kalman filtering equation.

[28] HST (1999) derive the observational equivalence result by fixing all parameters, including R, except for the pair (α, β).

[29] As shown in HST (1999), the two models have different implications for asset prices because continuation valuations would alter as one alters (α, β) within the observationally-equivalent set of parameters.

Acknowledgements This paper is prepared for the book volume "State Space Models — Application in Economics and Finance" in a new Springer Series. Luo thanks the Hong Kong GRF under grant No. 748209 and 749510 and HKU seed funding program for basic research for financial support. All errors are the responsibility of the authors. The views expressed here are the opinions of the authors only and do not necessarily represent those of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

Appendix

6 Solving the Current Account Model Explicitly under Model Uncertainty

To solve the Bellman equation (7), we conjecture that

$$v(s_t) = -As_t^2 - Bs_t - C,$$

where A, B, and C are undetermined coefficients. Substituting this guessed value function into the Bellman equation gives

$$-As_t^2 - Bs_t - C = \max_{c_t}\min_{\nu_t}\left\{-\frac{1}{2}(\overline{c}-c_t)^2 + \beta E_t\left[\vartheta\nu_t^2 - As_{t+1}^2 - Bs_{t+1} - C\right]\right\}. \tag{48}$$

We can do the min and max operations in any order, so we choose to do the minimization first. The first-order condition for ν_t is

$$2\vartheta\nu_t - 2AE_t\left[\omega_\zeta\nu_t + Rs_t - c_t\right]\omega_\zeta - B\omega_\zeta = 0,$$

which means that

$$\nu_t = \frac{B + 2A(Rs_t - c_t)}{2\left(\vartheta - A\omega_\zeta^2\right)}\omega_\zeta. \tag{49}$$

Substituting (49) back into (48) gives

$$-As_t^2 - Bs_t - C = \max_{c_t}\left\{-\frac{1}{2}(\overline{c}-c_t)^2 + \beta E_t\left[\vartheta\left[\frac{B+2A(Rs_t-c_t)}{2\left(\vartheta - A\omega_\zeta^2\right)}\omega_\zeta\right]^2 - As_{t+1}^2 - Bs_{t+1} - C\right]\right\},$$

where

$$s_{t+1} = Rs_t - c_t + \zeta_{t+1} + \omega_\zeta\nu_t.$$

The first-order condition for c_t is

$$(\overline{c}-c_t) - 2\beta\vartheta\frac{A\omega_\zeta}{\vartheta - A\omega_\zeta^2}\nu_t + 2\beta A\left(1 + \frac{A\omega_\zeta^2}{\vartheta - A\omega_\zeta^2}\right)\left(Rs_t - c_t + \omega_\zeta\nu_t\right) + \beta B\left(1 + \frac{A\omega_\zeta^2}{\vartheta - A\omega_\zeta^2}\right) = 0.$$

Using the solution for v_t the solution for consumption is

$$c_t = \frac{2A\beta R}{1 - A\omega_\zeta^2/\vartheta + 2\beta A} s_t + \frac{\bar{c}\left(1 - A\omega_\zeta^2/\vartheta\right) + \beta B}{1 - A\omega_\zeta^2/\vartheta + 2\beta A}. \tag{50}$$

Substituting the above expressions into the Bellman equation gives

$$\begin{aligned}
&-As_t^2 - Bs_t - C \\
&= -\frac{1}{2}\left(\frac{2A\beta R}{1 - A\omega_\zeta^2/\vartheta + 2\beta A} s_t + \frac{-2\beta A\bar{c} + \beta B}{1 - A\omega_\zeta^2/\vartheta + 2\beta A}\right)^2 \\
&+ \frac{\beta\vartheta\omega_\zeta^2}{\left(2\left(\vartheta - A\omega_\zeta^2\right)\right)^2}\left(\frac{2AR\left(1 - A\omega_\zeta^2/\vartheta\right)}{1 - A\omega_\zeta^2/\vartheta + 2\beta A} s_t + B - \frac{2\bar{c}\left(1 - A\omega_\zeta^2/\vartheta\right)A + 2\beta AB}{1 - A\omega_\zeta^2/\vartheta + 2\beta A}\right)^2 \\
&- \beta A\left(\left(\frac{R}{1 - A\omega_\zeta^2/\vartheta + 2\beta A} s_t - \frac{-B\omega_\zeta^2/\vartheta + 2c + 2B\beta}{2\left(1 - A\omega_\zeta^2/\vartheta + 2\beta A\right)}\right)^2 + \omega_\zeta^2\right) \\
&- \beta B\left(\frac{R}{1 - A\omega_\zeta^2/\vartheta + 2\beta A} s_t - \frac{-B\omega_\zeta^2/\vartheta + 2c + 2B\beta}{2\left(1 - A\omega_\zeta^2/\vartheta + 2\beta A\right)}\right) - \beta C.
\end{aligned}$$

Given $\beta R = 1$, collecting and matching terms, the constant coefficients turn out to be

$$A = \frac{R(R-1)}{2 - R\omega_\zeta^2/\vartheta}, \tag{51}$$

$$B = -\frac{R\bar{c}}{1 - R\omega_\zeta^2/(2\vartheta)}, \tag{52}$$

$$C = \frac{R}{2\left(1 - R\omega_\zeta^2/2\vartheta\right)}\omega_\zeta^2 + \frac{R}{2\left(1 - R\omega_\zeta^2/2\vartheta\right)(R-1)}\bar{c}^2. \tag{53}$$

Substituting (51) and (52) into (50) yields the consumption function. Substituting (53) into the current account identity and using the expression for s_t yields the expression for the current account.

References

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

2. Backus, David K., Patrick J. Kehoe, and Finn E. Kydland (1992), "International Real Business Cycles," *Journal of Political Economy* 101(4), 745-775.
3. Gali, J., Monacelli, T., (2005), "Monetary policy and exchange rate volatility in a small open economy," *Review of Economic Studies* 72(3), 702-734.
4. Barro, Robert J. (1979), "On the Determination of the Public Debt," *Journal of Political Economy* 87, 940-971.
5. Ghosh, Atish R. (1995), "Intertemporal Tax-smoothing and the Government Budget Surplus: Canada and the United States," *Journal of Money, Credit and Banking* 27, 1033–1045.
6. Hall, Robert E. (1978), "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy* 86(6), 971-87.
7. Hansen, Lars Peter and Thomas J. Sargent (1995), "Discounted Linear Exponential Quadratic Gaussian Control," *IEEE Transactions on Automatic Control* 40, pp. 968-971.
8. Hansen, Lars Peter and Thomas J. Sargent (2007a), *Robustness*, Princeton University Press.
9. Hansen, Lars Peter and Thomas J. Sargent (2007b), "Robust Estimation and Control without Commitment," *Journal of Economic Theory* 136(1), 1-27.
10. Hansen, Lars Peter, Thomas J. Sargent, and Thomas D. Tallarini, Jr. (1999), "Robust Permanent Income and Pricing," *Review of Economic Studies* 66(4), 873-907.
11. Hansen, Lars Peter, Thomas J. Sargent, and Neng Wang (2002), "Robust Permanent Income and Pricing with Filtering," *Macroeconomic Dynamics* 6(1), 40-84.
12. Huang, Chao-Hsi, Lin, Kenneth S. (1993), "Deficits, Government Expenditures and Tax Smoothing in the United States: 1929–1988," *Journal of Monetary Economics* 31, 317–339.
13. Leitemo, Kai, Ulf Soderstrom (2008), "Robust Monetary Policy in the New Keynesian Framework," *Macroeconomic Dynamics* 12 (supplement 1), 126-135.
14. Leitemo, Kai, Ulf Soderstrom (2008), "Robust Monetary Policy in a Small Open Economy," *Journal of Economic Dynamics and Control* 32, 3218-3252.
15. Luo, Yulei (2008), "Consumption Dynamics under Information Processing Constraints," *Review of Economic Dynamics* 11(2), 366-385.
16. Luo, Yulei and Eric R. Young (2010), "Risk-sensitive Consumption and Savings under Rational Inattention," *American Economic Journal: Macroeconomics* 2(4), 281-325.
17. Luo, Yulei and Eric R. Young (2011), "Robustifying the Permanent Income Model with Rational Inattention," manuscript.
18. Luo, Yulei, Jun Nie, and Eric R. Young (2011a), "Robustness, Information-processing Constraints, and the Current Account in Small Open Economies," forthcoming, *Journal of International Economics*.
19. Luo, Yulei, Jun Nie, and Eric R. Young (2011b), "Robust Control, Informational Frictions, and International Consumption Correlations," Manuscript.
20. Luo, Yulei, Jun Nie, and Eric R. Young (2011c), "Model Uncertainty and Intertemporal Tax Smoothing," Manuscript.
21. Obstfeld, Maurice and Kenneth Rogoff (1995), "The Intertemporal Approach to the Current Account," in Grossman, Gene M. and Kenneth Rogoff, editors, *Handbook of International Economics*, volume 3.
22. Obstfeld, Maurice and Kenneth Rogoff (1996), *Foundations of International Macroeconomics*, MIT Press.
23. Reis, Ricardo (2006), "Inattentive Consumers," *Journal of Monetary Economics* 53(8), 1761-1800.
24. Sims, Christopher A. (2003), "Implications of Rational Inattention," *Journal of Monetary Economics* 50(3), 665-690.
25. Sims, Christopher A. (2006), "Rational Inattention: Beyond the Linear-Quadratic Case," *American Economic Review* 96(2), 158-163.
26. Svensson, L.E.O. (2000), "Open-economy inflation targeting," *Journal of International Economics* 50(1), 155-184.
27. Walsh, Carl E. (2004), "Robust optimal instrument rules and robust control: An equivalent result," *Journal of Money, Credit, and Banking* 36(6), 1105-1113.

Table 1 Emerging vs. Developed Countries (Averages)

A: Emerging vs. Developed Countries (HP Filter)		
$\sigma(y)/\mu(y)$	4.09(0.23)	1.98(0.09)
$\sigma(\Delta y)/\mu(y)$	4.28(0.23)	1.89(0.07)
$\rho(y_t, y_{t-1})$	0.53(0.03)	0.66(0.02)
$\rho(\Delta y_t, \Delta y_{t-1})$	0.28(0.05)	0.46(0.03)
$\sigma(c)/\sigma(y)$	0.74(0.02)	0.59(0.02)
$\sigma(\Delta c)/\sigma(\Delta y)$	0.71(0.02)	0.59(0.02)
$\sigma(ca)/\sigma(y)$	0.79(0.03)	0.85(0.04)
$\rho(c, y)$	0.85(0.02)	0.78(0.02)
$\rho(ca_t, ca_{t-1})$	0.30(0.05)	0.41(0.03)
$\rho(ca, y)$	−0.59(0.05)	−0.35(0.04)
$\rho\left(\frac{ca}{y}, y\right)$	−0.54(0.04)	−0.36(0.04)
B: Emerging vs. Developed Countries (Linear Filter)		
$\sigma(y)/\mu(y)$	7.97(0.40)	4.79(0.22)
$\sigma(\Delta y)/\mu(y)$	4.28(0.23)	1.89(0.07)
$\rho(y_t, y_{t-1})$	0.79(0.02)	0.89(0.01)
$\rho(\Delta y_t, \Delta y_{t-1})$	0.28(0.05)	0.46(0.03)
$\sigma(c)/\sigma(y)$	0.72(0.02)	0.58(0.02)
$\sigma(\Delta c)/\sigma(\Delta y)$	0.71(0.02)	0.59(0.02)
$\sigma(ca)/\sigma(y)$	0.54(0.03)	0.65(0.04)
$\rho(c, y)$	0.88(0.02)	0.85(0.02)
$\rho(ca_t, ca_{t-1})$	0.53(0.04)	0.71(0.02)
$\rho(ca, y)$	−0.17(0.06)	−0.08(0.05)
$\rho\left(\frac{ca}{y}, y\right)$	−0.32(0.05)	−0.20(0.04)

Table 2 Implications of Different Models (Emerging Countries)

	Data	RE	RB	RB+RI ($\theta = 0.9$)	RB+RI ($\theta = 0.8$)	RB+RI ($\theta = 0.7$)	RB+RI ($\theta = 0.5$)
				($\lambda = 1$)			
$\rho(ca, y)$	0.13	1.00	0.62	0.57	0.56	0.56	0.58
$\rho(ca_t, ca_{t-1})$	0.53	0.80	0.74	0.57	0.50	0.45	0.36
$\sigma(ca)/\sigma(y)$	0.80	0.71	0.49	0.52	0.55	0.59	0.79
$\sigma(\Delta c)/\sigma(\Delta y)$	1.35	0.28	0.90	0.89	0.89	0.91	1.36
				($\lambda = 0.5$)			
$\rho(ca, y)$	0.13	1.00	0.62	0.59	0.58	0.59	0.64
$\rho(ca_t, ca_{t-1})$	0.53	0.80	0.74	0.63	0.59	0.55	0.46
$\sigma(ca)/\sigma(y)$	0.80	0.71	0.49	0.50	0.52	0.53	0.64
$\sigma(\Delta c)/\sigma(\Delta y)$	1.35	0.28	0.90	0.85	0.81	0.79	0.99
				($\lambda = 0.1$)			
$\rho(ca, y)$	0.13	1.00	0.62	0.61	0.60	0.61	0.67
$\rho(ca_t, ca_{t-1})$	0.53	0.80	0.74	0.67	0.64	0.62	0.56
$\sigma(ca)/\sigma(y)$	0.80	0.71	0.49	0.49	0.50	0.51	0.57
$\sigma(\Delta c)/\sigma(\Delta y)$	1.35	0.28	0.90	0.84	0.79	0.75	0.82

Table 3 Theoretical $\mathrm{corr}(c, c^*)$ from Different Models

	Data	RE	RB	RB+SU ($\theta = 0.9$)	RB+SU ($\theta = 0.6$)	RB+SU ($\theta = 0.3$)
Canada						
($\lambda = 1$)	0.38	0.41	0.33	0.27	0.17	0.12
($\lambda = 0.5$)	0.38	0.41	0.33	0.31	0.26	0.23
($\lambda = 0.1$)	0.38	0.41	0.33	0.32	0.32	0.32
Italy						
($\lambda = 1$)	0.25	0.54	0.50	0.42	0.27	0.19
($\lambda = 0.5$)	0.25	0.54	0.50	0.48	0.41	0.36
($\lambda = 0.1$)	0.25	0.54	0.50	0.50	0.50	0.49
UK						
($\lambda = 1$)	0.21	0.69	0.45	0.38	0.25	0.17
($\lambda = 0.5$)	0.21	0.69	0.45	0.44	0.38	0.32
($\lambda = 0.1$)	0.21	0.69	0.45	0.46	0.46	0.45
France						
($\lambda = 1$)	0.46	0.51	0.49	0.40	0.26	0.18
($\lambda = 0.5$)	0.46	0.51	0.49	0.46	0.40	0.34
($\lambda = 0.1$)	0.46	0.51	0.49	0.49	0.48	0.48
Germany						
($\lambda = 1$)	0.04	0.45	0.40	0.33	0.22	0.15
($\lambda = 0.5$)	0.04	0.45	0.40	0.38	0.33	0.29
($\lambda = 0.1$)	0.04	0.45	0.40	0.40	0.40	0.40