# Risk-sensitive Consumption and Savings under Rational Inattention

*By* YULEI LUO AND ERIC R, YOUNG*

*This paper studies the consumption-savings behavior of households who have risk-sensitive preferences and suffer from limited information-processing capacity (rational inattention or RI). We first solve the model explicitly and show that RI increases precautionary savings by interacting with income uncertainty and risk-sensitivity. Given the closed-form solutions, we find that the RI model displays a wide range of observational equivalence properties, implying that consumption and savings data cannot distinguish between risk-sensitivity, robustness, or the discount factor, in any combination. We then show that the welfare costs from RI are larger for risk-sensitive households than any other observationally-equivalent settings.*

*JEL: C61, D81, E21.*
*Keywords: Risk-sensitive Control, Robust Control, Rational Inattention, Consumption, Precautionary Savings.*

Christopher A. Sims (1998) first argued that rational inattention (RI) 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. Sims (2003) and Yulei Luo (2008) use this model to explore anomalies in the consumption literature, particularly the well-known "excess sensitivity" and "excess smoothness" puzzles, employing a linear-quadratic version of the standard permanent income model (as in Robert E. Hall 1978). In that model RI is equivalent to confronting the household with a signal extraction problem regarding the value of permanent income (as in Milton Friedman 1957) but permitting the agents to choose the distribution of the noise terms, subject to their limited capacity. With normal innovations to income the optimal distribution of noise is also normal, leading to a simple and analytically tractable problem.

Of course, the linear-quadratic model has some undesirable features as well. One objection that is particularly relevant for this paper is that it satisfies the certainty equivalence property, ruling out any response of saving to uncertainty (that is, precautionary behavior). Given that

* Luo (Corresponding author): School of Economics and Finance, The University of Hong Kong, Pokfulam Road, Hong Kong, yluo@econ.hku.hk. Young: Department of Economics, 242 Monroe Hall, University of Virginia, Charlottesville, VA 22904, ey2d@virginia.edu. We thank Olivier Blanchard, Thomas Muench, Jonathan Parker, Ricardo Reis, Chris Sims, Mirko Wiederholt, Chi-Wa Yuen, and an anonymous referee, as well as seminar and conference participants at the 2008 North American Summer Meetings of the Econometric Society, the 2008 Meeting of the Society for Economic Dynamics, the 2008 Midwest Macroeconomic Meetings, SUNY-Stony Brook, University of Hawaii at Manoa, the University of Hong Kong, and the Chinese University of Hong Kong for helpful comments and suggestions. Luo thanks the General Research Fund (GRF) in Hong Kong for financial support. Young thanks the Bankard Fund for Political Economy at Virginia for financial support. All remaining errors are our responsibility.

the important component of RI is the introduction of endogenous uncertainty into the household problem, it is not particularly desirable to use a model in which this uncertainty cannot manifest itself in decision rules. Fully nonlinear versions of the RI problem are solved in Sims (2005, 2006), Kurt F. Lewis (2006), Altantsetseg Batchuluun, Yulei Luo, and Eric R. Young (2008), and Antonella Tutino (2008); these papers show that the precautionary aspect of RI is important when channel capacity is small (which it must be to produce any interesting results). But the models solved in those papers 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). It is important to find a class of models that can produce precautionary behavior while maintaining tractability in the RI setup, if the properties of finite channel capacity are going to be thoroughly explored.

One class of models that satisfies these desiderata is the risk-sensitive permanent income model from Lars Peter Hansen, Thomas J. Sargent, and Thomas D. Tallarini (1999) (henceforth, HST).[1] In this model agents effectively compute expectations through a distorted lens, increasing their effective risk aversion by overweighting negative outcomes.[2] 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, leading again to the optimality of Gaussian noise in the RI model. Thus, we preserve the tractability of the LQ PIH model without being forced to accept certainty equivalence; for future reference we denote the model with risk sensitivity and rational inattention as the RS-RI model.[3] Formally, one can view risk-sensitive agents as ones who have non-state-separable preferences, as in Larry G. Epstein and Stanley E. Zin (1989), but with a restricted value for the intertemporal elasticity of substitution (see Tallarini 2000).

A second class of models that produces precautionary savings but remains within the class of LQ-Gaussian models is the robust control model of Hansen and Sargent (1995). In robust control problems, agents are concerned about the possibility that their model is misspecified in a manner that is difficult to detect statistically; as a result, 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). Hansen and Sargent (2007) present an observational equivalence result for RS and Robust models for consumption and savings decisions: any consumption path that could be generated by a model featuring risk sensitivity can also be generated by a model with robustness. Thus, introducing RI into the Robust model – which we denote the RB-RI model – is again straightforward, since the model retains the optimality of Gaussian noise.

Our purpose in this paper is to study the properties of RS-RI and RB-RI models. In particular, we are interested in two main questions. The first question involves the question of observational equivalence – do RS-RI and RB-RI models possess combinations of parameters such that their

[1] Possibly the first contribution to this literature is Frederick van der Ploeg (1993). Hansen and Sargent (2004) demonstrate that the risk-sensitive model displays a form of certainty equivalence, but that it generates precautionary savings anyway.

[2] Note that risk-sensitive agents in HST still have rational expectations, but act as if they are pessimistic.

[3] Of course, quadratic utility carries with it other objectionable properties, such as increasing relative risk aversion.

implied consumption-savings rules are the same? And can those decision rules also be generated by a standard RI model? To preview the results, which we discuss in more detail below, the answer is yes – any member of the class of LQ-Gaussian models with information-processing capacity can produce the same consumption function with appropriate adjustments in some combination of the discount factor, the degree of risk sensitivity, or the preference for robustness. The second question is how the costs of finite information-processing capacity differ within the class of observationally-equivalent models; that is, do observationally-equivalent agents suffer more or less from RI depending on which model generates their behavior? As a preview, we find here that the costs are not equal – RS-RI models produce larger welfare losses from RI than do RB-RI and RI models with higher discount factors. Thus, we identify a potential channel through which risk sensitivity, robustness, and the discount factor can be disentangled – simply ask agents what they would pay to relax their Shannon channel constraint.[4]

Our first step in studying the various RI models is to perform a reduction of the state space. Multivariate versions of the RI model require a constraint we term 'no subsidization.' With more than one state variable, agents can allocate their attention differently across these variables and thus reduce their uncertainty at different rates; RI requires that the uncertainty regarding one variable cannot be increased in order to reduce uncertainty regarding another, as this would permit reductions that exceed the channel capacity. This constraint is nonlinear when the dimension is larger than one, even in the LQ model, so multivariate versions of the RI model are not any more tractable than the nonlinear models mentioned above. We show that a PIH model with a general labor income process can be reduced to a model with a unique state variable – permanent income – that has iid innovations, enabling us to study a model with an empirically-reasonable income process within the LQ setup.

We then present the basic risk-sensitive (RS) model and derive some results regarding the behavior of consumption and saving. We show explicitly that both precautionary savings and the marginal propensity to consume out of permanent income are increasing in the parameter that measures risk sensitivity. From here we proceed by introducing finite Shannon channel capacity; RI implies that behavior is as if the agent were observing a noisy signal of current period's true permanent income, so the agent uses the Kalman filter to update the perceived state. Our RS-RI model predicts that the precautionary saving increment is determined by the interaction of three factors: labor income uncertainty, enhanced risk aversion due to RS, and incomplete information due to finite channel capacity (RI). Specifically, we show that it increases with the degrees of income uncertainty, RS, and RI.[5] Since the evolution of the true state is affected by precautionary motives, the evolution of the perceived state is as well. Thus, RS and RI interact – while the household uses the same Kalman gain parameter independent of the degree of risk sensitivity, the mean value of the signal received increases with RS so the perceived state evolves differently.

Because we are able to provide closed-form solutions to the RS-RI model we can provide two general results. First, introducing RI into the RS model improves the predictions about the aggregate joint dynamics of consumption and income; this result is obtained because the basic RS model actually *reduces* the ability of the model to match the sensitivity of consumption to

[4] Of course, figuring out how to conduct this experiment may be considerably more difficult.

[5] In a continuous-time setting, Neng Wang (2004) shows that (exogenously) partial observability of individual components of labor income generates additional precautionary saving due to estimation risk.

income changes.[6] Second, we show that observational equivalence holds for quantities – we can find more than one combination of discount factors and risk-sensitivity that imply the same savings and consumption behavior, whether RI is present or not. The intuition is the same as that in HST (1999) – an agent with RS looks just like a standard but more patient agent.

We then turn to the RB-RI model.[7] Here again we derive closed-form solutions and generate an observational equivalence result, which now extends in multiple directions. That is, we show that the RB-RI model produces the same consumption-savings decisions as the RS-RI model and, therefore, the RE and RI models. The intuition remains the same as above – with a preference for robust decision-making the agent chooses to save more than before, appearing more patient when viewed through the lens of a standard RE or RI model. Econometrically, these observational equivalence properties imply that it is impossible to test for the presence of RI, RS, or robustness using this model and consumption-savings data alone.

Given the observational equivalence result, we ask whether anything fundamentally distinguishes the RS-RI model from the RI and RB-RI models. We find a key difference when we measure the welfare costs of RI – that is, how much would agents pay an agency to reveal the true state of the world, given that their own capacity is fixed exogenously? Equivalently, one can think of this question as asking what an agent would pay to increase their information-processing capacity to $\infty$. Luo (2008) studied the costs of RI in the standard model and concluded that they are trivial. Generalizing his result to the RS-RI and RB-RI models, we find that the costs may not be trivial; in particular, agents in the RS-RI model can suffer very large losses from RI. In fact, we can prove that the RS-RI model implies that the welfare losses from RI are necessarily larger than those in RB-RI, and that this gap is increasing as channel capacity converges to 0. One key difference is that agents in the RS-RI model change their consumption behavior as their capacity changes, whereas RB-RI agents do not. The second key aspect is related to the asset pricing results in HST (1999) – RS-RI agents are more sensitive to the fluctuations in consumption and would pay more to remove them (mirroring the fact that the RS-RI model produces better asset pricing).

The remainder of the paper is organized as follows. Section 2 presents three PIH models with general income processes and discusses some results regarding optimal consumption-savings decisions and the joint dynamics of aggregate consumption and income (these models are the standard PIH model, the RS model, and the RB model). Section 3 introduces RI into both the RS and RB models and the behavior of consumption and savings; we provide closed-form solutions for all models here. Section 4 presents our battery of observational equivalence results. Section 5 then discusses the welfare costs of RI. Section 6 concludes with a discussion of some potential directions that the literature could proceed.

## I. Three Rational Expectations Permanent Income Models

In this section, we present three benchmark models. In the first model, we show that within the LQG framework, the multivariate permanent income model with general income processes can

[6] This fact explains the use of habit formation and preference shocks in RS models (such as HST 1999).

[7] Hansen, Sargent, and Wang (2002) consider a robust LQ permanent income model with filtering, and examine equilibrium market prices of risk. They model imperfect information by assuming that the consumer cannot observe the components in the aggregate endowment.

be reduced to the univariate model with iid innovations to permanent income that can be solved in closed-form solution. We then present the second model in which risk-sensitivity is introduced into the univariate PIH model in Section A. We finally introduce the third model in which the preference for robustness is incorporated into the first standard RE PIH model.

### A. The Standard Permanent Income Model

In the standard permanent income model households solve the dynamic consumption-savings problem

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

subject to

$$w_{t+1} = R(w_t - c_t) + y_{t+1}, \tag{2}$$

where $u(c_t) = -\frac{1}{2}(c_t - \overline{c})^2$ is the period utility function, $\overline{c} > 0$ is the bliss point, equation (2) is the flow budget constraint, $c_t$ is consumption, $w_t$ is cash-on-hand (or market resources), $y_t$ is a general income process with Gaussian white noise innovations, $\beta$ is the discount factor, and $R$ is the constant gross interest rate at which the consumer can borrow and lend freely.[8] For the rest of the paper we will restrict attention to points where $c_t < \overline{c}$, so that utility is increasing and concave. This specification follows that in Hall (1978) and Marjorie Flavin (1981) and implies that optimal consumption is determined by permanent income:

$$c_t = \left(1 - \frac{1}{\beta R^2}\right) s_t - \frac{1}{R-1}\left(1 - \frac{1}{\beta R}\right)\overline{c}, \tag{3}$$

where

$$s_t = w_t + \sum_{j=1}^{\infty} R^{-j} E_t \left[ y_{t+j} \right] \tag{4}$$

is the expected present value of lifetime resources, consisting of financial wealth plus human wealth. If $s_t$ is defined as a new state variable, we can reformulate the above PIH model as

$$v(s_0) = \max_{\{c_t\}} E_0 \left[ \sum_{t=0}^{\infty} \beta^t u(c_t) \right] \tag{5}$$

[8] As in Hansen and Sargent (2007) or HST (1999), we can also interpret this problem as applying to a linear production technology that generates consumption and capital,

$$c_t + k_t = R_k k_{t-1} + y_t,$$

where $R_k$ is the net physical return on capital and investment is given by $i_t = k_t - k_{t-1}$. If the allocation is decentralized the return coincides with the risk-free interest rate $R$.

subject to

$$s_{t+1} = R(s_t - c_t) + \zeta_{t+1}; \tag{6}$$

a formal proof is contained in Luo (2008). The stochastic term $\zeta_{t+1}$ is equal to

$$\zeta_{t+1} = \sum_{j=t+1}^{\infty} R^{-j+(t+1)} \left(E_{t+1} - E_t\right) \left[y_j\right]; \tag{7}$$

that is, $\zeta_{t+1}$ the time $(t+1)$ innovation to permanent income. $v(s_0)$ is the consumer's value function under RE. By defining a new state variable $s_t$, the original multivariate optimization problem has been reduced to a univariate problem; under quadratic utility this model leads to the well-known random walk result of Hall (1978) when $\beta R = 1$:

$$\Delta c_{t+1} = \frac{R-1}{R} \zeta_{t+1}; \tag{8}$$

in this case, the change in consumption depends neither on the past history of labor income nor on anticipated changes in labor income. We also point out the well-known result that the standard PIH model with quadratic utility implies the certainty equivalence property holds; thus, uncertainty has no impact on optimal consumption, so that there is no precautionary saving. We assume for the remainder of this section that $\beta R = 1$, since this setting is the only one that implies stationary consumption.

For simplicity, following Jorn-Steffen Pischke (1995) we suppose that income $y_t$ can be expressed as the sum of *aggregate* permanent and *idiosyncratic* transitory components:

$$y_{t+1} = y^p_{t+1} + y^t_{t+1}, \tag{9}$$

where the superscripts $p$ and $t$ denote permanent and transitory, respectively. Each of these components follows its own stochastic process; $y^p_{t+1}$ follows a random walk

$$y^p_{t+1} = y^p_t + \varepsilon_{t+1}, \tag{10}$$

and $y^t_{t+1}$ follows an iid process

$$y^t_{t+1} = \overline{y} + \epsilon_{t+1}, \tag{11}$$

where $\varepsilon_{t+1}$ and $\epsilon_{t+1}$ are white noises with mean 0 and variance $\omega^2_\varepsilon$ and $\omega^2_\epsilon$ respectively. Substituting these income processes into (8) gives

$$\Delta c_{t+1} = \varepsilon_{t+1} + \frac{R-1}{R} \epsilon_{t+1} \tag{12}$$

because the innovation to permanent income $\zeta_{t+1} = \frac{R}{R-1}\varepsilon_{t+1} + \epsilon_{t+1}$. The different load fac-

tors on the two shocks reflects the standard consumption-smoothing motives: consume all of a permanent increase in income ($\varepsilon_{t+1}$) but only the annuity value of a transitory one ($\epsilon_{t+1}$).

We now want to derive some properties of aggregate consumption. Suppose that there are a continuum of consumers in the model economy. Given the expression of the change in individual consumption, (12), aggregating across all consumers yields the change in aggregate consumption as

$$\Delta C_{t+1} = \varepsilon_{t+1}, \tag{13}$$

so that aggregate consumption $C_{t+1}$ is unpredictable using past information and the smoothness ratio of aggregate consumption to income is 1.[9] In other words, the predictions of the standard PIH model for the joint behavior of aggregate consumption and income is not consistent with the empirical evidence; in the US data aggregate consumption growth is much smoother than income and is sensitive to past information.[10] These two anomalies have been termed the *excess smoothness* and *excess sensitivity* puzzles in the literature.[11] Furthermore, as documented in Ricardo Reis (2006), the impulse response of aggregate consumption to aggregate income takes a hump-shaped form, which means that aggregate consumption reacts to income shocks gradually. The standard RE PIH model cannot capture this feature in the US data.

### B. The Risk-sensitive Permanent Income Model

Risk-sensitivity (RS) was first introduced into the LQ-Gaussian framework by David H. Jacobson (1973) and extended by Peter 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. HST (1999) use this model to study a risk-sensitive version of the PIH model and show how RS alters the choices of consumption, investment, and asset prices; Tallarini (2000) addresses many of the same issues within the stochastic growth model.

We use an RE version of risk-sensitive control based on recursive preferences with an exponential certainty equivalence function. The household problem can be written recursively as

$$\upsilon(s_t) = \max_{c_t} \left\{ -\frac{1}{2}(c_t - \overline{c})^2 + \beta \mathcal{R}_t \left[ \upsilon(s_{t+1}) \right] \right\}, \tag{14}$$

subject to (6). The distorted expectation operator $\mathcal{R}_t$ is defined by

$$\mathcal{R}_t \left[ \upsilon(s_{t+1}) \right] = -\frac{1}{\alpha} \log E_t \left[ \exp\left(-\alpha \upsilon(s_{t+1})\right) \right], \tag{15}$$

where $\alpha > 0$ measures higher risk aversion *vis a vis* the von Neumann-Morgenstern specification. Solving this optimization problem yields the following proposition:

[9] The idiosyncratic shock, $\epsilon$, would be canceled out after aggregating over all consumers.

[10] In the US data, the smoothness ratio is around 0.5.

[11] See Angus Deaton (1992) for a recent review.

PROPOSITION 1: *The value function for the RS model is*

$$v(s_t) = -\frac{1}{2}\frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\zeta^2}s_t^2 + \frac{R\left(\beta R^2 - 1\right)}{\left(\beta R^2 - \alpha\omega_\zeta^2\right)(R-1)}\bar{c}s_t -$$

$$\left(\frac{1}{2}\frac{R^2\left(\beta R^2 - 1\right)}{(R-1)^2\left(\beta R^2 - \alpha\omega_\zeta^2\right)}\bar{c}^2 - \frac{\beta}{1-\beta}\frac{1}{2\alpha}\log\left(1 - \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\zeta^2}\alpha\omega_\zeta^2\right)\right), \tag{16}$$

*the consumption function is*

$$c_t = \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\zeta^2}s_t + \frac{\left(R - \beta R^2 + (1-R)\alpha\omega_\zeta^2\right)\bar{c}}{\left(\beta R^2 - \alpha\omega_\zeta^2\right)(R-1)}. \tag{17}$$

PROOF:

See Appendix A.

Equation (17) shows that the parameter controlling risk-sensitivity, $\alpha$, affects both the marginal propensity to consume (MPC) out of perceived permanent income, $\frac{\beta R^2-1}{\beta R^2-\alpha\omega_\zeta^2}$, and the precautionary savings increment, $-\frac{\left(R-\beta R^2+(1-R)\alpha\omega_\zeta^2\right)\bar{c}}{\left(\beta R^2-\alpha\omega_\zeta^2\right)(R-1)}$, where $\omega_\zeta^2 = \left(\frac{R}{R-1}\right)^2\omega_\varepsilon^2 + \omega_\epsilon^2$.[12] The larger the value of $\alpha$ the larger the MPC and the larger the precautionary saving increment (we impose the restriction $\beta R^2 > 1 > \alpha\omega_\zeta^2$ so that precautionary savings is positive), since

$$\frac{\partial\left(\frac{\beta R^2-1}{\beta R^2-\alpha\omega_\zeta^2}\right)}{\partial\alpha} > 0 \text{ and } \frac{\partial\left(-\frac{\left(R-\beta R^2+(1-R)\alpha\omega_\zeta^2\right)\bar{c}}{\left(\beta R^2-\alpha\omega_\zeta^2\right)(R-1)}\right)}{\partial\alpha} > 0.$$

As its name suggests, risk sensitivity makes agents more concerned about outcomes that are unfavorable; one way to minimize the probability of bad outcomes (here, outcomes where the marginal utility of consumption becomes very large) is to save a bit more, reducing consumption relative to the certainty equivalent setup where $\alpha = 0$. 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.[13] It is worth noting that this type of precautionary saving does not depend on the convexity of the marginal utility of consumption and occurs even when the period utility of consumers is quadratic as we have shown. In the long run agents will have more wealth as the extra precautionary savings generates more returns.

[12] Given the definition of $s_t$ in (6), the MPC out of permanent income equals the MPC out of financial wealth $w_t$ and human wealth $\sum_{j=1}^{\infty} R^{-j}E_t\left[y_{t+j}\right]$.

[13] Later we show that although RS and RB are observationally equivalent in the sense that they cannot be distinguished by consumption, savings and investment data alone under RE or RI, they have different implications for welfare and stabilization policies. Hence, we should be careful when using one to interpret the other.

Relative risk aversion over wealth gambles is given by rra $(s_t) = \frac{1}{\frac{R}{R-1}\frac{\bar{c}}{s_t}-1}$, which is an decreasing function of $\frac{\bar{c}}{s_t}$. Thus, $\alpha$ does not directly impact risk aversion in this model; what matters for risk aversion is only how close current permanent income is to the bliss point. $\alpha$ increases $s_t$, however, and thus indirectly increases the risk aversion of the agent; the increasing relative risk aversion property of quadratic utility functions remains in the RS setting.

We now present some basic results from the RS model, mainly to facilitate comparisons to the RI model of the next section. The change in consumption can be written as

$$\Delta c_{t+1} = -\frac{\beta R^2 - R - (1-R)\,\alpha\omega_\zeta^2}{\beta R^2 - \alpha\omega_\zeta^2}\,(c_t - \bar{c}) + \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\zeta^2}\zeta_{t+1}. \tag{18}$$

Equation (18) has two important implications for our purposes. Setting $\alpha > 0$ implies that consumption growth becomes more volatile relative to income and becomes more sensitive to unanticipated changes in permanent income, since $\frac{\beta R^2-1}{\beta R^2} < 1$. Furthermore, consumption growth is not mean zero unless $c_t = \bar{c}$ even when $\beta R = 1$; indeed, when $\beta R = 1$ consumption will drift upward over time. Counteracting this drift requires $\widehat{\beta} R < 1$ such that the tendency to save for precautionary reasons is exactly counterbalanced by the impatience effect of increasing consumption; this result is the basis of the observational equivalence result in HST (1999).

As in the last subsection, substituting the specified income processes, (9), (10), and (11), into (18) gives

$$\Delta c_{t+1} = \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\zeta^2}\left[\left(\frac{R}{R-1}\varepsilon_{t+1} + \epsilon_{t+1}\right) + (\rho_1 - 1)\,\frac{\frac{R}{R-1}\varepsilon_t + \epsilon_t}{1-\rho_1 L}\right], \tag{19}$$

where $\rho_1 = R\frac{1-\alpha\omega_\zeta^2}{\beta R^2 - \alpha\omega_\zeta^2} \in (0, 1)$ and $L$ is the lag operator.

Aggregating (19) across all consumers gives the change in aggregate consumption:

$$\Delta C_{t+1} = \frac{R}{R-1}\frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\zeta^2}\left[\varepsilon_{t+1} + (\rho_1 - 1)\,\frac{\varepsilon_t}{1-\rho_1 L}\right]. \tag{20}$$

Figure 1 illustrates the response of aggregate consumption growth to an aggregate income shock $\varepsilon_{t+1}$; RS raises the sensitivity of consumption growth to unanticipated changes in aggregate income.

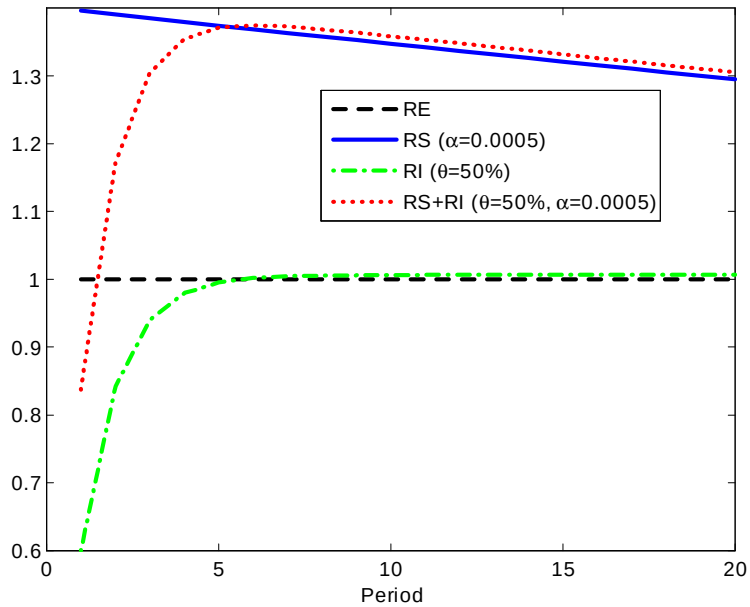


Figure 1: Responses of consumption to income shock, $\varepsilon$

Furthermore, the smoothness ratio of aggregate consumption to income is seen to be

$$\mu = \frac{\mathrm{sd}\left[\Delta C_t\right]}{\mathrm{sd}\left[\varepsilon_t\right]} = \frac{R}{R-1}\frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\zeta^2}\sqrt{\frac{2}{1+\rho_1}}, \tag{21}$$

where "sd" denotes standard deviation. It is straightforward to prove the following proposition, which is the key point from this section.

PROPOSITION 2: $\frac{\partial \mu}{\partial \alpha} > 0$.

PROOF:

Clearly the first term is increasing in $\alpha$. Then

$$\frac{\partial \rho_1}{\partial \alpha} = R\frac{\omega_\zeta^2 (1-R)}{\left(R - \alpha\omega_\zeta^2\right)^2} < 0.$$

Thus, $\mu$ is increasing in $\alpha$.

RS increases the response of aggregate consumption growth to a change in income, just as it does for individual income. Thus, RS by itself *worsens* the standard RE PIH model's prediction

for the joint behavior of aggregate consumption and income growth by exacerbating the excess smoothness puzzle, and therefore needs to be combined with other assumptions to resolve the anomalies.[14]

### C. The Robust Permanent Income Model

Robust control emerged in the engineering literature in the 1970s, and was introduced into economics and further developed by Hansen, Sargent, and others. A simple version of robust optimal control considers such a question: How to make decisions when the agent does not know the probability model that generates the data? The agent with the preference for robustness considers a range of models, and makes decisions that maximize utility given the worst possible model. Following Hansen and Sargent (2007), the simple robust permanent income problem can be written as

$$v(s_t) = \max_{c_t} \min_{\nu_t} \left\{ -\frac{1}{2}(\overline{c} - c_t)^2 + \beta \left[ \vartheta \nu_t^2 + E_t \left[ v(s_{t+1}) \right] \right] \right\} \tag{22}$$

subject to

$$s_{t+1} = R(s_t - c_t) + \zeta_{t+1} + \omega_\zeta \nu_t, \tag{23}$$

where $\nu_t$ distorts the mean of the innovation and $\vartheta > 0$ controls how bad the error can be.[15] Solving this robust control problem yields the following proposition:[16]

PROPOSITION 3: *The value function for the above RB model is*

$$\begin{aligned} v(s_t) \quad = \quad & -\frac{\beta R^2 - 1}{2\beta R^2 - \omega_\zeta^2/\vartheta} s_t^2 - \frac{R}{R-1} \frac{\beta R^2 - 1}{\beta R^2 - \omega_\zeta^2/(2\vartheta)} \overline{c} s_t \\ & - \left( \frac{R^2 \left(\beta R^2 - 1\right)}{2(R-1)^2 \left(\beta R^2 - \omega_\zeta^2/(2\vartheta)\right)} \overline{c}^2 + \frac{\beta}{1-\beta} \frac{\beta R^2 - 1}{2\beta R^2 - \omega_\zeta^2/\vartheta} \omega_\zeta^2 \right), \end{aligned} \tag{24}$$

*and the consumption function is*

$$c_t = \frac{\beta R^2 - 1}{\beta R^2 - \omega_\zeta^2/(2\vartheta)} s_t + \frac{(R-1)\left(1 - \omega_\zeta^2/(2\vartheta)\right) + 1 - \beta R^2}{(R-1)\left(\beta R^2 - \omega_\zeta^2/(2\vartheta)\right)}. \tag{25}$$

[14] In HST (1999), habit persistence is incorporated into the RS version of the PIH model. Hansen and Sargent (2007) use habit persistence and/or adjustment costs.

[15] Formally, this setup is a game between the decision-maker and a malevolent nature that chooses the distortion process $\nu_t$. $\vartheta \geq 0$ 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.

[16] The recursive robust control problem can be solved using the procedure adopted in David K. Backus, Bryan R. Routledge, and Stanley E. Zin (2004) and Hansen and Sargent (2007). The detailed derivation is available from the corresponding author by request.

The consumption function under RB, (25), shows that the preference for robustness, $\vartheta$, affects both the MPC out of permanent income, $\frac{\beta R^2-1}{\beta R^2-\omega_\zeta^2/(2\vartheta)}$, and the precautionary savings increment, $-\frac{(R-1)\left(1-\omega_\zeta^2/(2\vartheta)\right)+1-\beta R^2}{(R-1)\left(\beta R^2-\omega_\zeta^2/(2\vartheta)\right)}$. The less the value of $\vartheta$ the larger the MPC and the larger the precautionary saving increment (we impose the restriction $\beta R^2 > 1 > \omega_\zeta^2/(2\vartheta)$ so that precautionary savings is positive). Furthermore, comparing (25) with (17), it is clear that when $\alpha = 1/(2\vartheta)$, the RS and RB models are observationally equivalent (OE) in the sense that they generate the same consumption and savings decisions, as well as the same joint dynamics of consumption and income. For this reason, HST (1999) and Hansen and Sargent (2007) 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. However, (16) and (24) clearly show that the two models under OE have different welfare implications, as

$$-\frac{1}{\alpha}\log\left(1-\frac{\beta R^2-1}{\beta R^2-\alpha\omega_\zeta^2}\alpha\omega_\zeta^2\right) > \frac{\beta R^2-1}{\beta R^2-\omega_\zeta^2/(2\vartheta)}\omega_\zeta^2$$

for $\alpha > 0$. In other words, RS consumers are more risk averse than RB consumers in our models, and thus are willing to pay more to avoid the uncertainty due to labor income. Furthermore, since stabilization policies are closely related to welfare costs due to uncertainty, the two models would have different implications for stabilization policies, and thus they cannot be viewed observationally equivalent in this aspect.

## II. Permanent Income Models with Rational Inattention

We now introduce rational inattention (RI) by assuming the consumer faces information processing constraints; as first suggested in Sims (1998), households have only finite Shannon channel capacity that they can apply to observations of the state of the world. As in Sims (2003), we use the concept of *entropy* from information theory to characterize the rate of information flow; the reduction in entropy is a natural measure of information flow, where entropy is defined as a measure of the uncertainty about a random variable.[17] With finite channel capacity, the consumer will choose a signal that reduces the uncertainty of the state variable subject to limitations on the extent that entropy can be reduced. Formally, this idea is described by the information processing constraint

$$\mathcal{H}(s_{t+1}|\mathcal{I}_t) - \mathcal{H}(s_{t+1}|\mathcal{I}_{t+1}) \leq \kappa, \tag{26}$$

where $\kappa$ is the consumer's channel capacity that imposes an upper bound on the amount of information that can be transmitted via the channel, $\mathcal{H}(s_{t+1}|\mathcal{I}_t)$ denotes the entropy of the state prior to observing the new signal at $t+1$, and $\mathcal{H}(s_{t+1}|\mathcal{I}_{t+1})$ denotes the entropy after observing the

[17] See Claude Shannon (1948) and Thomas M. Cover and Joy A. Thomas (1991) for details.

new signal.[18]

Given the LQG specification, Sims (2003, 2005) showed that $(s_t|\mathcal{I}_t)$ is distributed $N\left(\widehat{s}_t, \sigma_t^2\right)$, where $\widehat{s}_t = E_t\left[s_t\right]$ and $\sigma_t^2 = \text{var}_t\left[s_t\right]$ are the conditional mean and variance of the state variable $s_t$, respectively. In models with imperfect observations, the optimal decisions are determined by the *perceived* state, rather than the *actual* state. Computational difficulties arise in RI models because the perceived state is the distribution of the actual state variable conditional on the information set available at time $t$, $\mathcal{I}_t$, and as we have noted already this object typically has high dimension. Fortunately, given the LQG specification and the Gaussian distribution of $s_t$, the first two moments, $\widehat{s}_t$ and $\sigma_t^2$, are enough to characterize the perceived state. In addition, as we will show, the problem will simplify further because $\sigma_t^2$ converges to a constant in the stochastic steady state.

The constraint (26) can be rewritten as

$$\log\left|\psi_t^2\right| - \log\left|\sigma_{t+1}^2\right| \leq 2\kappa, \tag{27}$$

where $\sigma_{t+1}^2 = \text{var}_{t+1}\left[s_{t+1}\right]$ and $\psi_t^2 = \text{var}_t\left[s_{t+1}\right] = R^2\sigma_t^2 + \text{var}_t\left[\zeta_{t+1}\right]$ are the posterior and prior variance of the state variable, $s_{t+1}$, respectively.[19] As shown in Sims (2003), in any univariate case, this information constraint completes the characterization of the optimization problem with RI and the model can be solved explicitly. Furthermore, with a finite capacity $\kappa$ the optimizing consumer will choose a signal that reduces the conditional variance of $s_{t+1}$ by a limited amount. Hence, (27) must be binding for the optimizing consumer:

$$\log\left|\psi_t^2\right| - \log\left|\sigma_{t+1}^2\right| = 2\kappa. \tag{28}$$

It is straightforward to show that in the univariate case (28) has a steady state $\overline{\sigma}^2$. In that steady state, $\sigma_t^2 = \overline{\sigma}^2 = \frac{\omega_\zeta^2}{\exp(2\kappa)-R^2}$, where $\omega_\zeta^2 = \text{var}_t\left[\zeta_{t+1}\right] = \left(\frac{R}{R-1}\right)^2\omega_\varepsilon^2 + \omega_\epsilon^2$, and the consumer behaves as if he is observing a noisy signal, $s_{t+1}^* = s_{t+1} + \xi_{t+1}$, where $\xi_{t+1}$ is the iid endogenous noise and its variance, $\varrho_t^2 = \text{var}_t\left[\xi_{t+1}\right]$, is determined by the usual updating formula of the variance of a Gaussian distribution:

$$\sigma_{t+1}^2 = \psi_t^2 - \psi_t^2\left(\psi_t^2 + \varrho_t^2\right)^{-1}\psi_t^2. \tag{29}$$

[18] We regard $\kappa$ as a technological parameter. If the base for logarithms is 2, the unit used to measure information flow is a 'bit', if we use the natural logarithm $e$, the unit is a 'nat', and if we use a base 10 the unit is a 'dit' (also called a 'ban' or a 'hartley'). 1 nat is equal to $\log_2 e = 1.433$ bits. We will use the nat in this paper. Lewis (2006) studies a model where households can increase their channel capacity by reducing their leisure; we have not investigated whether that model can be mapped into the LQ-Gaussian framework.

[19] Note that here we use the fact that the entropy of a Gaussian random variable is equal to half of its logarithm variance plus some constant term.

Thus, in the steady state we have

$$\varrho_t^2 = \overline{\varrho}^2 = \mathrm{var}_t\left[\xi_{t+1}\right] = \frac{\left(\omega_\zeta^2 + R^2\overline{\sigma}^2\right)\overline{\sigma}^2}{\omega_\zeta^2 + (R^2-1)\overline{\sigma}^2}.$$

For simplicity, we assume that initially the model economy is at the steady state, implying $s_0 | \mathcal{I}_0 \sim N\left(\widehat{s}_0, \overline{\sigma}^2\right)$.

To draw attention to the value of reducing the multivariate PIH model to a univariate one we note that RI with multiple state variables requires a second constraint:

$$\Psi - \Sigma \succeq 0,$$

where $\Psi$ and $\Sigma$ are the prior and posterior variance of the state variable. This constraint is to be read as 'the matrix defined as the difference between the posterior and the prior element-by-element must be positive semidefinite.' We term this constraint a 'no subsidization' condition, as it implies that an agent cannot reduce one variable's entropy by more than $\kappa$ by increasing the entropy of another. This constraint takes the form of restrictions on eigenvalues and thus is not linear except in the univariate case where it is unnecessary.[20] As a result, the multivariate version of the RI PIH model is no longer tractable; we therefore exploit the state reduction presented above and use permanent income as the unique state variable.

Sims (2003) and Luo (2008) introduce RI into Hall's permanent income models, and show that RI can provide an endogenous propagation mechanism that disentangles the short-run and long-run responses of consumption to income shocks (see the dotted-dashed line in Figure 1). In particular, Luo (2008) introduces information-processing constraints, (26), into the standard PI model proposed in Section A and solves for the following expression for consumption growth:

$$\Delta c_t = \theta\frac{R-1}{R}\left[\left(\frac{\zeta_t}{1-(1-\theta)RL}\right) + \left(\xi_t - \frac{\theta R\xi_{t-1}}{1-(1-\theta)RL}\right)\right]; \tag{30}$$

this MA($\infty$) expression clearly shows that with finite capacity, consumption reacts to the income shock gradually and with delay. Comparing with the predictions of the RS-PIH model (see the solid line in Figure 1), it is clear that RS (or RB) and RI have distinct effects on the responses of consumption to income shocks: RS (or RB) just increases the sensitivity of consumption to the unanticipated income shock, whereas RI dampens the consumption reaction and thus generates a hump-shaped response to the shock.[21] Kenneth Kasa (2006) shows in a continuous-time filtering model that RI and RB are observationally equivalent, in the sense that a higher filter gain can either be interpreted as a stronger preference for robustness or an increased information-processing capacity; however, as shown above, RI and RB (or RS) are totally different, in the sense that they generate different dynamics of consumption and savings.

Luo (2008) contains a complete discussion of the basic linear-quadratic-Gaussian RI model,

[20] It is also likely that some variables would generate corner solutions where their entropy is not reduced at all, further exacerbating the nonlinearity of the constraint.

[21] Note that RS and RB generate the same consumption-savings rules in the permanent income model.

so we proceed by introducing two models that generalize the RI model to permit precautionary effects. The first, denoted RS-RI, introduces the risk sensitivity motive that we discussed in the previous section. The second, denoted RB-RI, assumes that agents have a demand for robust decision rules (those that are good against a wide range of misspecifications in the model they use for forecasting income movements). We lay out each model and solve it in this section.

### A. RS-RI Model

The household problem is characterized by the Bellman equation

$$\widehat{v}\left(\widehat{s}_t\right) = \max_{c_t}\left\{-\frac{1}{2}\left(c_t-\overline{c}\right)^2+\beta\mathcal{R}_t\left[\widehat{v}\left(\widehat{s}_{t+1}\right)\right]\right\} \tag{31}$$

subject to the budget constraint, (6), and the Kalman filter equation

$$\widehat{s}_{t+1} = (1-\theta)\,R\left(\widehat{s}_t - c_t\right) + \theta\left(s_{t+1}+\xi_{t+1}\right). \tag{32}$$

The distorted expectation operator is now given by

$$\mathcal{R}_t\left[\widehat{v}\left(\widehat{s}_{t+1}\right)\right] = -\frac{1}{\alpha}\log E_t\left[\exp\left(-\alpha\widehat{v}\left(\widehat{s}_{t+1}\right)\right)\right], \tag{33}$$

$s_0|\,\mathcal{I}_0 \sim N\left(\widehat{s}_0,\overline{\sigma}^2\right)$, $\widehat{s}_t = E_t\left[s_t\right]$ is the perceived state variable, $\theta$ is the optimal weight on the new observation of the state, and $\xi_{t+1}$ is the endogenous noise. The optimal choice of the weight $\theta$ is given by $\theta(\kappa) = \frac{\exp(2\kappa)-1}{\exp(2\kappa)} \in [0,1]$; therefore, we choose to parameterize RI using $\theta$ directly. Furthermore, $\lim_{\kappa\to\infty}\theta(\kappa) = 1$, so that we nest the standard RS specification as a special case of our model. To make economic sense in subsequent results we impose the condition $\theta > \frac{R^2-1}{R^2} > \frac{R-1}{R}$, which implies that agents must be able to process some minimum amount of information. This condition is similar to one imposed in Luo and Young (2009) that guarantees an agent with higher capacity optimally chooses a noise process with lower variance; we regard this requirement to be intuitively reasonable.[22]

The following proposition summarizes the solution to the RI-RS model.

PROPOSITION 4: *Given finite channel capacity $\kappa$ and the degree of risk-sensitivity $\alpha$, the value function of a risk-sensitive consumer under RI is*

$$\begin{aligned}\widehat{v}\left(\widehat{s}_t\right) = &-\frac{\beta R^2-1}{2\beta R^2-2\alpha\omega_\eta^2}\widehat{s}_t^2+\frac{R\left(1-\beta R^2\right)}{\left(\beta R^2-\alpha\omega_\eta^2\right)(R-1)}\overline{c}\widehat{s}_t- \\ &\left[\frac{1}{2}\frac{R^2\left(\beta R^2-1\right)}{(R-1)^2\left(\beta R^2-\alpha\omega_\eta^2\right)}\overline{c}^2-\frac{\beta}{1-\beta}\frac{1}{2\alpha}\log\left(1-2\alpha\frac{\beta R^2-1}{2\beta R^2-2\alpha\omega_\zeta^2}\omega_\eta^2\right)\right],\end{aligned} \tag{34}$$

[22] This assumption is not needed in fully nonlinear versions of the RI model, suggesting an important breakdown of the linear-quadratic setup as channel capacity gets very small.

*and the consumption function is*

$$c_t = \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2}\widehat{s}_t + \frac{\left(1 - \alpha\omega_\eta^2\right)(R-1) + 1 - \beta R^2}{\left(\beta R^2 - \alpha\omega_\eta^2\right)(R-1)}\overline{c}, \tag{35}$$

*where*

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

*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 B.

To interpret these expressions, we remind the reader that $\xi_t$ is the noise shock induced by RI and $\zeta_t$ is the innovation to permanent income. $\eta_{t+1}$ will turn out to be the stochastic component in the growth rate of consumption (see below); unlike the basic RS model, $\eta_{t+1}$ contains a forecastable stochastic component due to the presence of the MA($\infty$) term. There are two special cases of interest that are nested in our specification: if $\alpha = 0$ we obtain the same expressions as in Luo (2008) (the basic RI model), and if $\theta = 1$ we obtain $\eta_{t+1} = \zeta_{t+1}$ (the basic RS model).

Note that the permanent income shock $\zeta_t$ and the noise shock $\xi_t$ have the same effect on $\widehat{s}_{t+1}$ (they both increase perceived permanent income by $\theta \leq 1$) but only $\zeta_t$ affects $s_{t+1}$ (it increases by 1); because the variance of the noise shock will be decreasing in $\theta$, agents with high processing capacity will end up with $\widehat{s}_t$ tracking $s_t$ quite closely. The consumption function shows that $\theta$ affects both the MPC out of the perceived state variable $\frac{\beta R^2-1}{\beta R^2-\alpha\omega_\eta^2}$ and the precautionary savings increment $\frac{\left(1-\alpha\omega_\eta^2\right)(R-1)+1-\beta R^2}{\left(\beta R^2-\alpha\omega_\eta^2\right)(R-1)}$ only through the variance of the noise distribution. With finite capacity $\omega_\eta^2 > \omega_\zeta^2$ and it then follows that $\omega_\eta^2$ is decreasing in $\theta$.

PROPOSITION 5: $\frac{\partial\omega_\eta^2}{\partial\theta} < 0$.

PROOF:

By simple calculation we obtain

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

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

Thus, both the MPC out of the perceived state variable (the responsiveness of $c_t$ to $\widehat{s}_t$) and the precautionary savings increment (the intercept of the consumption profile) are decreasing in $\theta$ for fixed $\alpha$, meaning that agents with lower capacity save a larger fraction of any perceived income increase. Figure 2 illustrates how the combination of $(\alpha, \theta)$ affects the MPC out of financial and human wealth (which are equal here) when $R = 1.01$. The relative risk aversion expression is

now rra $(\widehat{s_t}) = \frac{1}{\frac{R}{R-1}\frac{\bar{c}}{\widehat{s_t}}-1}$; the key difference is that risk aversion depends not on how far actual permanent income is from the bliss point but rather how far perceived permanent income is from the bliss point. An RS-RI agent is more risk averse than a standard RS agent whenever $\widehat{s_t} > s_t$, which occurs when the agent receives negative shocks to $\zeta$ or positive shocks to $\xi$.

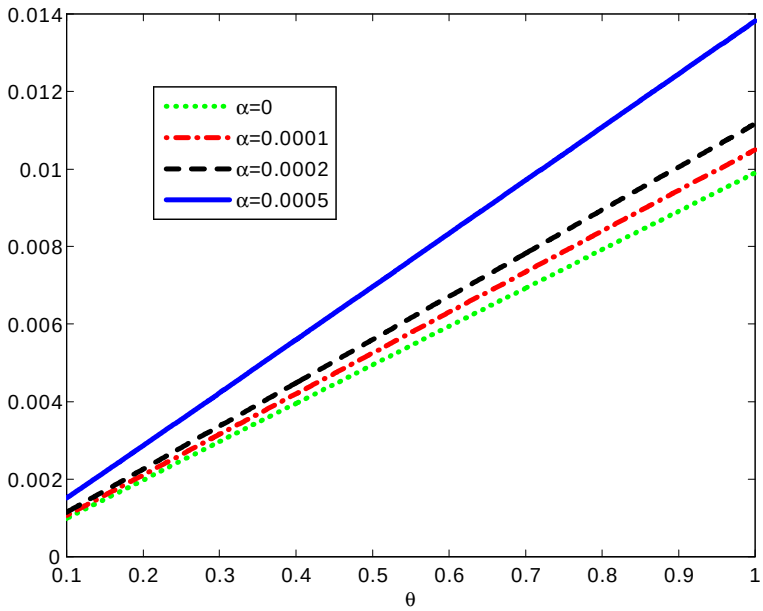


Figure 2: MPC out of financial wealth $\omega$ and human wealth $\frac{y}{R-1}$

Since agents with low capacity are very concerned about the confluence of low permanent income and high consumption (meaning they believe their permanent income is high so they consume a lot and then their new signal indicates that in fact their permanent income was low), they take actions which reduce the probability of this bad event – they save more.[23] The strength of the precautionary effect is positively related to the amount of uncertainty regarding the true level of permanent income, and this uncertainty increases as $\theta$ gets smaller. Figure 3 illustrates the effect of $(\alpha, \theta)$ on the precautionary savings increment. In the absence of risk sensitivity ($\alpha = 0$) the precautionary savings increment is zero independent of $\theta$; if $\alpha > 0$, however, precautionary savings is decreasing in $\theta$, so that agents with lower capacity will consume less on average.[24] Therefore, the precautionary savings increment in the RS-RI model is determined by the interaction of three factors: labor income uncertainty, enhanced risk aversion relative to

[23] A more complete discussion of precautionary savings in an RI model can be found in Batchuluun, Luo, and Young (2008).

[24] Note that agents with RI display precautionary behavior with respect to second-order increases in risk without requiring convex marginal utility, whereas RE households do not (see Louis Eeckhoudt and Harris Schlesinger 2008).

the expected-utility setting (RS), and finite information-processing capacity (RI). Many empirical studies have estimated the importance of precautionary saving when people are exposed to future income uncertainty. For example, Christopher Carroll and Andrew Samwick (1997) estimate the precautionary savings for typical U.S. households to be in the range of $20-50\%$ of total savings, while Luigi Guiso, Tullio Jappelli, and Daniele Terlizzese (1992) use a survey of Italian households to estimate that the precautionary component explains only a small fraction of total saving. Hence, the result in Guiso, Jappelli, and Terlizzese (1992) casts doubt on the empirical relevance of precautionary saving generated by labor income uncertainty, but does not contradict the importance of the precautionary motive *per se*. Besides exogenous income uncertainty, the endogenous uncertainty due to RI generated from our RS-RI model may be a key determinant of precautionary savings.

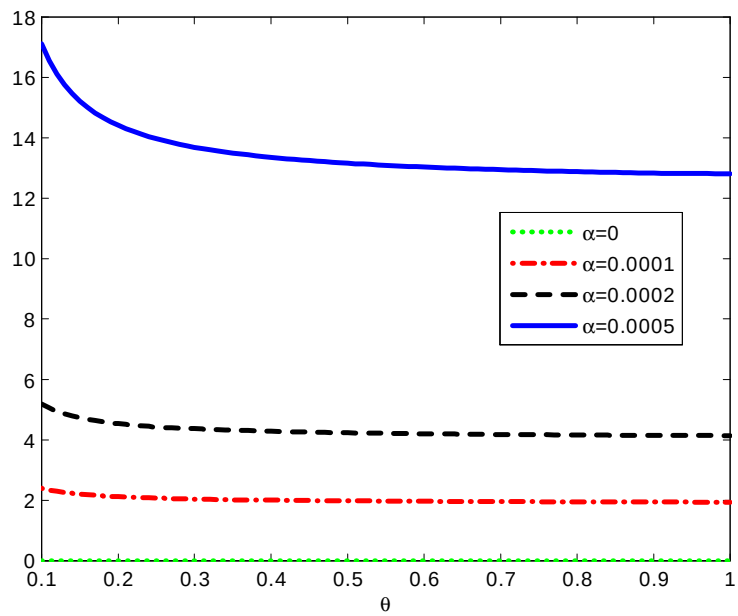


Figure 3: Effect of $(\alpha, \theta)$ on precautionary savings

INDIVIDUAL DYNAMICS. — The change in individual consumption in the RI-RS economy can be written as

$$\Delta c_{t+1} = \frac{R - \beta R^2 + (1-R)\,\alpha\omega_\eta^2}{\beta R^2 - \alpha\omega_\eta^2}\,(c_t - \overline{c}) + \theta R \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2}\left(\frac{(1-\theta)\,\zeta_t - \theta\xi_t}{1-(1-\theta)\,RL}\right) + \theta\frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2}\left(\zeta_{t+1} + \xi_{t+1}\right)$$

so that

(36)
$$\Delta c_t = \theta \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2} \left\{ \sum_{j=0}^{\infty} \Gamma_j \zeta_{t-j} + \sum_{j=0}^{\infty} \Gamma_j \xi_{t-j} - R \sum_{j=0}^{\infty} \Gamma_j \xi_{t-1-j} \right\} + \frac{R - \beta R^2 + (1-R)\,\alpha\omega_\eta^2}{\beta R^2 - \alpha\omega_\eta^2} \overline{c}$$

where $\rho_1 = \frac{R - \alpha\omega_\eta^2 R}{\beta R^2 - \alpha\omega_\eta^2} \in (0,1)$, $\rho_2 = (1-\theta)\,R \in (0,1)$, and $\Gamma_j = \sum_{k=0}^{j} \left(\rho_1^{j-k} \rho_2^k\right) - \sum_{k=0}^{j-1} \left(\rho_1^{j-1-k} \rho_2^k\right)$. Thus, consumption growth is seen to be a weighted average of all past permanent income and noise shocks.

Figure 1 also clearly shows that RI has a similar qualitative effect on consumption in the standard RE model and the RS model – consumption reacts gradually to income shocks, with monotone adjustments to the corresponding asymptote – while quantitatively RI has larger impacts on consumption in the RS model because the deviation of the asymptote from that for the standard RE case is larger in this case. With a stronger preference for risk-sensitivity, the precautionary savings increment is larger and thus an income shock that is initially undetected would have larger impacts on consumption during the adjustment process. In addition, we note that the immediate response of consumption to a shock to permanent income is given by

(37)
$$\mathrm{SR}\,(\varepsilon_t) = \theta \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2}.$$

This expression is increasing in $\alpha$, implying that the more risk-sensitive the household the more consumption responds initially to changes in permanent income. It is also increasing in $\theta$, implying that agents with lower capacity have smaller initial responses, so that RI and RS have opposing effects. In Figure 4 we plot the short-run responses of consumption growth as a function of $\theta$ for various values of $\alpha$.

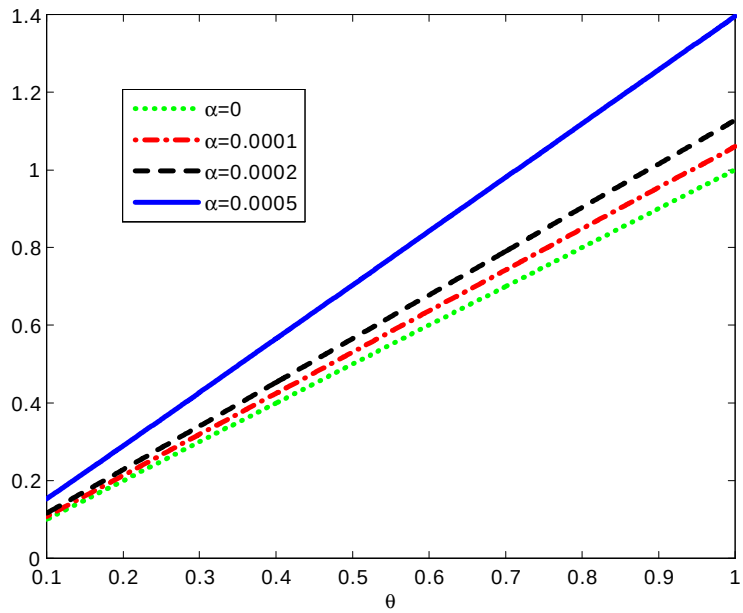


Figure 4: Short-run responses of consumption

With RI agents respond slowly, similar to their response if they faced convex adjustment costs for their wealth; one resulting implication is that the short-run and long-run responses of consumption are different. The long-run response is given by

$$\text{LR}(\varepsilon_t) = \frac{\partial\left(\sum_{s=0}^{T}\Delta c_{t+s}\right)}{\partial\varepsilon_t} = \theta\frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2}\sum_{k=0}^{T}\left(\rho_1^{T-k}\rho_2^k\right). \tag{38}$$

Note that when $\alpha = 0$ and $\rho_1 = 1$ and taking the limit as $T \to \infty$ this expression reduces to the result reported in Luo (2008):

$$\text{LR}(\varepsilon_t) = \frac{\theta}{1-\rho_2}. \tag{39}$$

It is clear from the above two expressions that RS has two opposite impacts on the long-run responses of consumption to income shocks under RI. First, $\alpha$ increases the long-run response by a factor $\theta\frac{R}{R-\alpha\omega_\eta^2}$; second, it reduces the long-run response by a factor $\lim_{T\to\infty}\sum_{k=0}^{T}\left(\rho_1^{T-k}\rho_2^k\right) =$

$0 < \frac{1}{1-\rho_2}$ because $|\rho_1| < 1$. Furthermore, the volatility of individual consumption growth is

(40)
$$\operatorname{var}[\Delta c_t] = \left(\theta \frac{R-1}{R-\alpha\omega_\eta^2}\right)^2 \left\{ \sum_{j=0}^{\infty} \Gamma_j^2 \omega_\zeta^2 + \sum_{j=0}^{\infty} \left[ \Gamma_j - (1+R) \sum_{k=0}^{j-1} \left(\rho_1^{j-1-k} \rho_2^k\right) + \sum_{k=0}^{j-2} \left(\rho_1^{j-1-k} \rho_2^k\right) \right]^2 \omega_\xi^2 \right\}.$$

Intuitively, this expression should be decreasing in $\theta$ and increasing in $\alpha$. With lower $\theta$, agents have less channel capacity and therefore choose more volatile noise shocks, leading to changes in consumption that are more sensitive to changes in perceived permanent income.[25]

AGGREGATE DYNAMICS. — Since the expression for the change in individual consumption (36) permits exact aggregation, we can obtain the change in aggregate consumption as

$$\Delta C_t = \theta \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2} \left\{ \sum_{j=0}^{\infty} \Gamma_j \varepsilon_{t-j} + \sum_{j=0}^{\infty} \Gamma_j E^i \left[\xi_{t-j}\right] - R \sum_{j=0}^{\infty} \Gamma_j E^i \left[\xi_{t-1-j}\right] \right\}, \tag{41}$$

where $i$ denotes a particular individual and $E^i[\cdot]$ is the population average. This expression shows that even if every consumer only faces the common shock $\varepsilon$, the RI economy still has heterogeneity since each consumer faces the idiosyncratic noise induced by finite channel capacity. 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 information-processing constraint, there is likely a significant common component; provided that agents face similar needs for coding macroeconomic information, they will rely on common sources and generate aggregate noise shocks. Existing theory does not provide a way to determine the common component of the noise term; we can only state that the common term $E^i\left[\xi_t^i\right]$ is between 0 and the part of the idiosyncratic error $\xi_t$ generated by the aggregate income shock $\varepsilon_t$. Formally, assume that $\xi_t$ consists of two independent noises: $\xi_t = \overline{\xi}_t + \xi_t^i$, where $\overline{\xi}_t = E^i\left[\xi_t\right]$ and $\xi_t^i$ are the common and idiosyncratic components of the error generated by $\zeta_t$, respectively. A parameter, $\lambda = \frac{\operatorname{var}\left[\overline{\xi}_t\right]}{\operatorname{var}\left[\xi_t\right]} \in [0,1]$, can be used to measure the common source of coded information on the aggregate component (or the relative importance of $\overline{\xi}_t$ to $\xi_t$). In order to simplify expressions we first consider the case where all noises are idiosyncratic, that is, $\lambda = 0$ (so that individuals live on isolated islands and do not interact with each other directly or indirectly via conversation, imitation, newspapers, or other media); in this special case the change in aggregate consumption can be written as

$$\Delta C_t = \theta \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2} \left\{ \sum_{j=0}^{\infty} \Gamma_j \varepsilon_{t-j} \right\}. \tag{42}$$

[25] The fact that the noise shocks are necessarily more volatile with lower channel capacity is a consequence of normality. Otherwise, the distribution of noise shocks would merely have more entropy, which might not necessarily translate into higher variance.

Therefore, the volatility of aggregate consumption relative to income can be written as

$$\mu = \frac{\mathrm{sd}\,[\Delta C_t]}{\mathrm{sd}\,[\varepsilon_t]} = \left(\theta \frac{\beta R^2 - 1}{\beta R^2 - \alpha \omega_\eta^2}\right) \sqrt{\sum_{j=0}^{\infty} \Gamma_j^2}, \tag{43}$$

which means that RI reduces the volatility of aggregate consumption growth in the risk-sensitive PIH model. Figure 5 illustrates how the combination of RI and RS affects the smoothness ratio of aggregate consumption. It is clear that incorporating RI into the risk-sensitive PIH model can reduce the volatility of aggregate consumption and thus bring the model and the data closer along this dimension. For example, when $\alpha = 5 \times 10^{-5}$, $\omega_\varepsilon = 5.6$, $R = 1.01$, and $\theta = 30\%$, the excess smoothness ratio is about 0.44, very close to its empirical counterpart in the US data.

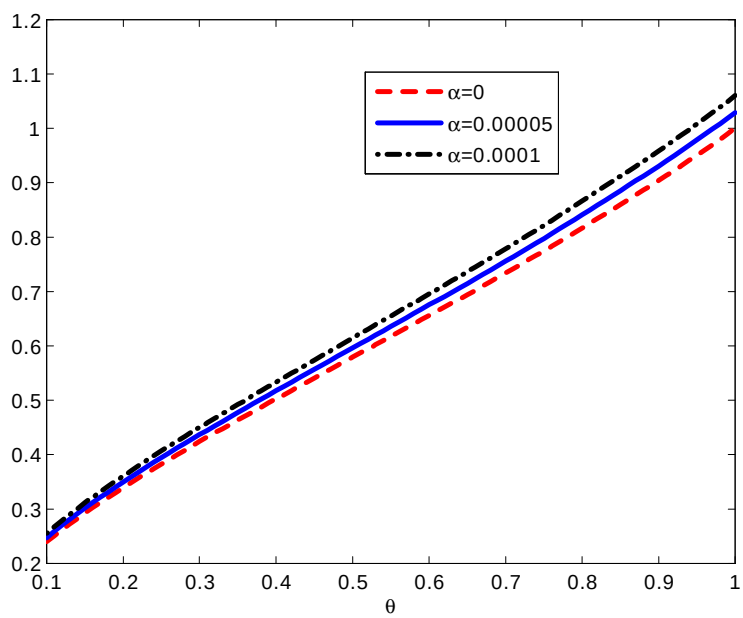


Figure 5: Effect of $(\alpha, \theta)$ on smoothness of aggregate consumption

However, $\lambda = 0$ is a strong assumption as consumers in most countries do not live in isolated islands and they interact with each other by conversations, reading the same newspapers, watching the same TV programs, and so on. We therefore need to consider the more general case in which $\lambda \in (0, 1]$. In this case, the volatility of aggregate consumption relative to income can be

written as

$$\mu = \left(\theta \frac{\beta R^2 - 1}{\beta R^2 - \alpha \omega_\eta^2}\right) \sqrt{\sum_{j=0}^{\infty} \Gamma_j^2 + \lambda^2 \frac{1-\theta}{\theta\left[1-(1-\theta)R^2\right]} \sum_{j=0}^{\infty} \left(\Gamma_j - R\Gamma_{j-1}\right)^2}, \tag{44}$$

where we use the fact that $\omega_\xi^2 = \text{var}\left[\xi_t\right] = \frac{1-\theta}{\theta\left[1-(1-\theta)R^2\right]}\omega_\varepsilon^2$. Figure 6 illustrates how the combination of $\theta$ and $\lambda$ affects the smoothness ratio of aggregate consumption when $\alpha = 5 \times 10^{-5}$, $\omega_\varepsilon = 5.6$, and $R = 1.01$. It is clear that the aggregation factor, $\lambda$, increases the volatility of aggregate consumption because the presence of the common noise, $\overline{\xi}_t$, offsets the smoothness of consumption caused by the gradual responses to fundamental shocks. Some combinations of $(\theta, \lambda)$ can still explain the observed smoothness of aggregate consumption. For example, when $\theta = 0.3$ and $\lambda = 0.3$, the excess smoothness ratio is about 0.48, also very close to its empirical counterpart.

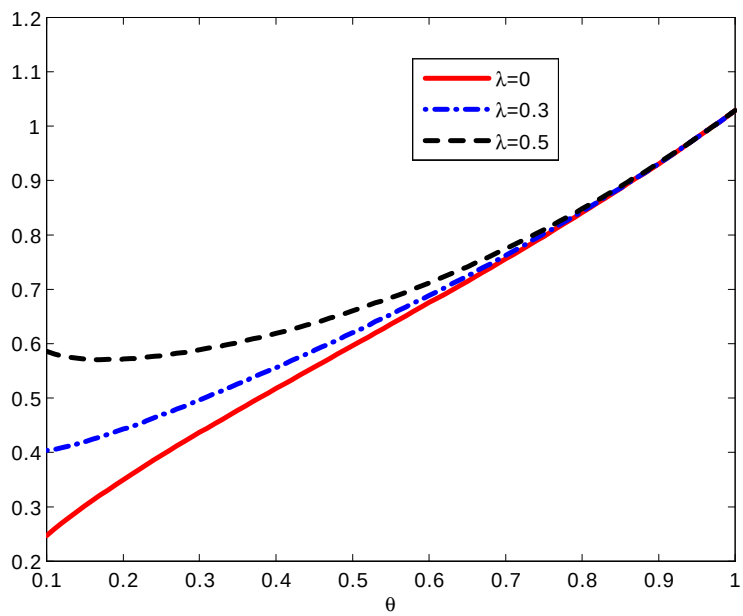


Figure 6: Effect of $(\theta, \lambda)$ on smoothness of aggregate consumption

### B. Robust-RI Model

We now consider the connections between risk sensitivity and robustness. Hansen and Sargent (2007) demonstrate a deep connection between the choices made by an agent that is risk-sensitive and one that has a preference for robustness. An agent with a preference for robust decision rules

considers a range of models and makes decisions that maximize utility given the worst possible model. In this section, we assume that agents with finite capacity distrust their model generating the data, and use an ordinary Kalman filter to estimate the true state.[26] The robust PIH problem with inattentive consumers can thus be written as

$$\widehat{v}(\widehat{s}_t) = \max_{c_t} \min_{\nu_t} \left\{ -\frac{1}{2}(c_t - \overline{c})^2 + \beta E_t \left[ \vartheta \nu_t^2 + \widehat{v}(\widehat{s}_{t+1}) \right] \right\} \tag{45}$$

subject to the distorted flow budget constraint

$$s_{t+1} = R(s_t - c_t) + \zeta_{t+1} + \omega_\zeta \nu_t \tag{46}$$

and the Kalman filter equation

$$\widehat{s}_{t+1} = (1-\theta)\left[R(\widehat{s}_t - c_t) + \omega_\zeta \nu_t\right] + \theta\left(s_{t+1} + \xi_{t+1}\right), \tag{47}$$

where $\nu_t$ distorts the mean of the innovation and $\vartheta > 0$ controls how bad the error can be.[27] After substituting (46) into (47), the Kalman filter equation can be rewritten as

$$\widehat{s}_{t+1} = R(\widehat{s}_t - c_t) + \omega_\zeta \nu_t + \eta_{t+1}. \tag{48}$$

where $\eta_{t+1} = \theta R(s_t - \widehat{s}_t) + \theta\left(\zeta_{t+1} + \xi_{t+1}\right)$.

The following proposition summarizes the solution to the RB-RI model.

PROPOSITION 6: *Given $\vartheta$ the value function of an agent with preference for robustness under RI is given by*

$$\begin{aligned} \widehat{v}(\widehat{s}_t) = &-\frac{\beta R^2 - 1}{2\beta R^2 - \omega_\zeta^2/\vartheta}\widehat{s}_t^2 - \frac{R}{R-1}\frac{\beta R^2 - 1}{\beta R^2 - \omega_\zeta^2/(2\vartheta)}\overline{c}\widehat{s}_t - \\ &\left(\frac{R^2\left(\beta R^2 - 1\right)}{2\left(\beta R^2 - \omega_\zeta^2/(2\vartheta)\right)(R-1)^2}\overline{c}^2 + \frac{\beta}{1-\beta}\frac{\beta R^2 - 1}{2\beta R^2 - \omega_\zeta^2/\vartheta}\omega_\eta^2\right), \end{aligned} \tag{49}$$

*the consumption function is*

$$c_t = \frac{\beta R^2 - 1}{\beta R^2 - \omega_\zeta^2/(2\vartheta)}\widehat{s}_t + \frac{\left(1 - \omega_\zeta^2/(2\vartheta)\right)(R-1) + 1 - \beta R^2}{\left(\beta R^2 - \omega_\zeta^2/(2\vartheta)\right)(R-1)}\overline{c}, \tag{50}$$

[26] In particular, a distortion to the mean of permanent income is introduced to represent possible model misspecification.

[27] Formally, this setup is a game between the decision-maker and a malevolent nature that chooses the distortion process $\nu_t$. $\vartheta \geq 0$ 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.

*the optimal worst-case distribution is given by*

$$\nu_t = \frac{\omega_\zeta^2}{2\vartheta} \frac{\beta R^2 - 1}{\beta R \left(\beta R^2 - \omega_\zeta^2 / (2\vartheta)\right)} \widehat{s}_t + \frac{\omega_\zeta^2}{2\vartheta} \left(\frac{R}{R-1}\right) \frac{\left(\beta R^2 - 1\right)}{\beta R \left(\beta R^2 - \omega_\zeta^2 / (2\vartheta)\right)} \overline{c}. \tag{51}$$

PROOF:

See Appendix C.

Equation (50) displays an important property – it is independent of $\theta$. Thus, in the RB-RI model limited channel capacity does not affect consumption and precautionary savings except through changes in the level of perceived permanent income. In this model, the preference for robustness inspires precautionary savings because agents wants to protect themselves against mistakes in specifying conditional means of the innovation to permanent income. Just like in the RS model, this kind of precautionary savings does not depend on the convexity of the marginal utility of consumption. The same property holds for (51); the worst-case distribution does not directly depend on the channel capacity. The value function therefore only directly depends on $\theta$ through the constant term. Individual and aggregate dynamics of consumption under robustness are very similar to those under RS, so we omit an explicit discussion.

## III. Observational Equivalence

In this section we detail the observational equivalence results that connect the RI, RS-RI, and RB-RI models. All three models are capable of producing the same consumption-savings decisions for some combination of parameters $(\alpha, \beta, \vartheta, \theta)$, rendering them fundamentally unidentified by this data alone. Our purpose is to provide intuition about why the observational equivalence result obtains in three settings – between RI and RS-RI, between RI and RB-RI, and between RS-RI and RB-RI.

### *A. Observational Equivalence Between RI and RS-RI*

HST (1999) show that as far as the *quantity* observations on consumption and investment are concerned, the risk-sensitive version $(\alpha > 0, \widetilde{\beta})$ of the PIH model is observationally equivalent to the standard version $(\alpha = 0, \beta)$ 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] However, under RI, both the marginal propensity to consume (MPC) and precautionary savings are determined by the interactions of RI and RS (or RB). Since there is no prior knowledge about the mechanism of the impacts of interactions of RI and RS (or RB) on MPC and precautionary savings, the OE result proposed by HST (1999) might not hold under RI. In this section we

[28] HST (1999) derive the observational equivalence result by fixing all parameters, including $R$, except for the pair $(\alpha, \beta)$.

[29] As shown in HST (1999), the two models have different implications for asset prices because continuation valuations would alter as one alters $(\alpha, \beta)$ within the observational equivalence set.

show that a generalized observational equivalence result holds under RI. Specifically, holding all parameters constant except the pair $(\alpha, \beta)$, the RI version of the PIH model with risk-sensitive consumers ($\alpha > 0$ and $\beta R < 1$) is observationally equivalent to the standard RI version of the model ($\alpha = 0$ and $\beta R = 1$), extending the observational equivalence result to a broader class of models. To do so, we fix $R$ and assume that $\beta$ differs across the two settings; thus, $\beta^{RS,RI}$ is the discount factor needed in the RS-RI model to achieve OE with the RI model where $\beta R = 1$.

PROPOSITION 7: *Let*

$$\beta^{RS,RI} = \frac{R - (R-1)\,\alpha\omega_\eta^2}{R^2} < \frac{1}{R}. \tag{52}$$

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

PROOF:

See Appendix D.

This proposition is straightforward to prove by simply setting the coefficients in the consumption function equal and solving for the discount factor. We show in the appendix that the same outcome is obtained by setting the slope and curvature coefficients in the value functions equal. (52) implies that the required discount factor is decreasing as a function of $\alpha$ and increasing as a function of $\theta$:

$$\frac{\partial}{\partial\alpha}\left(\beta^{RS,RI}\right) = -\frac{(R-1)\,\omega_\eta^2}{R^2} < 0$$
$$\frac{\partial}{\partial\theta}\left(\beta^{RS,RI}\right) = -\frac{(R-1)\,\alpha}{R^2}\frac{\partial\omega_\eta^2}{\partial\theta} > 0.$$

Holding fixed $\theta$, an agent who is more sensitive to risk (higher $\alpha$) will save more; thus, to match an agent who is not sensitive to risk that agent must be less patient. Similarly, holding fixed $\alpha$ an agent with a higher channel capacity (higher $\theta$) will save less, requiring them to be more patient to match the saving of an agent who is risk sensitive but has lower channel capacity. Note that $\frac{\partial}{\partial\theta}\left(\beta^{RS,RI}\right) = 0$ if $\alpha = 0$, so that the discount factor would not need to be adjusted if agents are not risk sensitive; this result simply restates the result in Luo (2008) and arises because RI does not alter savings plan when $\alpha = 0$. Figure 7 plots the locus of OE between $(\alpha, \beta)$ for various different values of $\theta$ (we fix $R = 1.01$).

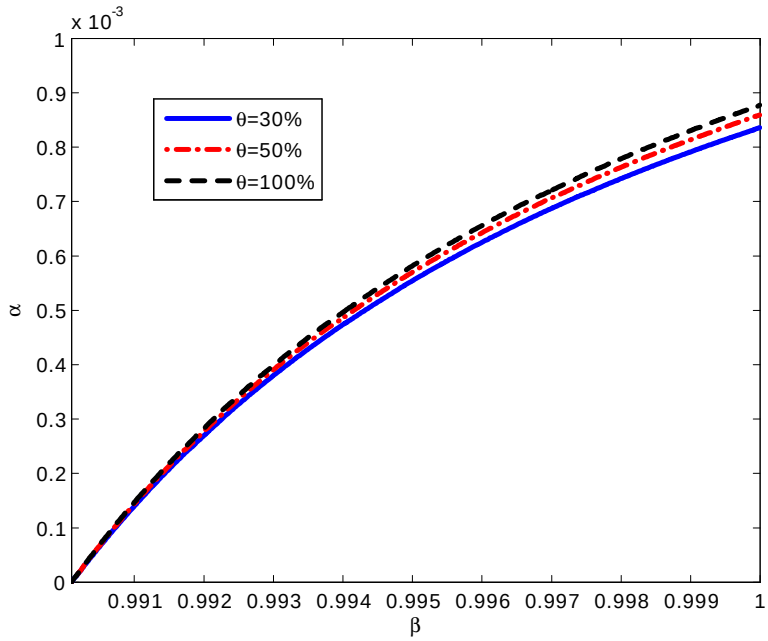


Figure 7: Observational equivalence between $\alpha$ and $\beta$

### B. Observational Equivalence Between RI and RB-RI

We now turn to a comparison of RI and RB-RI. As above, we can derive an expression that links the discount factor in the RB-RI model to the robustness parameter $\vartheta$ in such a way that the agent would make the same choices in the RI and RB-RI models.

PROPOSITION 8: *Let*

$$\beta^{RB,RI} = \frac{R - (R-1)\,\omega_\zeta^2/(2\vartheta)}{R^2} < \frac{1}{R}. \tag{53}$$

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

PROOF:

See Appendix D.

Here, the discount factor needed in the RB-RI model does not depend on the channel capacity parameter $\theta$; this independence is a consequence of the previously-noted independence of consumption in this model. It is immediate that the discount factor is an increasing function of $\vartheta$:

$$\frac{\partial}{\partial\vartheta}\left(\beta^{RB,RI}\right) = \frac{R-1}{R^2}\frac{\omega_\zeta^2}{2\vartheta^2} > 0.$$

An agent with higher $\vartheta$ is less concerned about model misspecification and therefore tends to save less (equivalently, they consider only models that are very close to the trusted one, leading to less distortion and therefore less additional saving); in order to match the behavior of an agent who does not require robust decisions the robust agent must be made more patient. If $\vartheta = \infty$ then $\frac{\partial}{\partial \vartheta}\left(\beta^{RB,RI}\right) = 0$; again, this result matches that found in Luo (2008).

### C. Observational Equivalence Between RS-RI and RB-RI

We now turn the final comparisons, where we provide parameter constellations for $(\alpha, \beta, \theta, \vartheta)$ that jointly produce observational equivalence. These results imply that there exists a fundamental lack of identification between the class of models we are considering here (linear-quadratic-Gaussian models); it would be impossible to distinguish between any of them using only consumption-savings decisions.[30]

PROPOSITION 9: *Let the following expression hold:*

$$\frac{\alpha\theta}{1-(1-\theta)\,R^2} = \frac{1}{2\vartheta}. \tag{54}$$

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

PROOF:

See Appendix D.

To interpret these expressions we fix $\beta$ at some common value. An agent who is risk-sensitive ($\alpha > 0$) and suffers from finite Shannon channel capacity ($\theta < 1$) will make the same choices as an agent who is concerned with robustness ($\vartheta < \infty$). The "required" $\vartheta$ is given by

$$\vartheta = \frac{1-(1-\theta)\,R^2}{2\alpha\theta}, \tag{55}$$

where $\frac{\partial\vartheta}{\partial\theta} = \frac{R^2-1}{2\alpha\theta^2} > 0$ and $\frac{\partial\vartheta}{\partial\alpha} = -\frac{1-(1-\theta)R^2}{2\alpha^2\theta} < 0$. That is, if we compare two agents with the same $\alpha$, the one with higher channel capacity will behave the same as an agent who is less concerned with model misspecification. If we compare two agents with the same $\theta$, the one with higher risk sensitivity will behave the same as an agent who is more concerned with model misspecification.

Another observational equivalence result arises if we fix $(\alpha, \vartheta)$ for two agents who have the same $\theta$. It is straightforward to show that there exists a pair of discount factors, $\beta^{RI,RS}$ and $\beta^{RI,RB}$, such that the two agents choose the same consumption plans.

PROPOSITION 10: *Let the two discount factors satisfy the relation*

$$\beta^{RI,RB} = \frac{\left(1-R^2\beta^{RI,RS}\right)\omega_\zeta^2 + 2\vartheta\left(R^2\beta^{RI,RS} - \alpha\omega_\eta^2\right)}{2\vartheta\,R^2\left(1-\alpha\omega_\eta^2\right)}.$$

[30] By extension, if we interpret our model as a model of capital accumulation with a linear technology, it is impossible to distinguish the models using only consumption and investment data.

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

## IV. Welfare Costs From RI

Since information-processing constraints cannot help in individuals' optimization – agents under RI cannot observe the state perfectly when making optimal decisions – the average welfare difference between the RI and RE economies is greater than 0. We present here the welfare cost of RI – how much utility does an agent lose if the actual consumption path he chooses under RI deviates from the first-best RE consumption path? Alternatively, we ask what an agent would pay to increase channel capacity $\kappa$ to $\infty$ (so that the optimal choice is $\theta = 1$). Luo (2008) shows that the welfare costs of RI are fairly small in the certainty equivalent environment; our interest here is assessing how these costs differ in the RS-RI, RB-RI, and RI models.[31] To make the comparisons meaningful, we restrict our attention to combinations of preferences that imply observational equivalence – that is, do agents who look the same (but for different reasons) suffer the same welfare losses from information-processing capacity limitations?

### A. *Welfare Losses in RS-RI and RB-RI*

We assume for this section that $(\alpha, \vartheta)$ are such that OE obtains between the two models (note that OE requires that $\alpha$ depend on $\theta$). Within this class we can derive the lifetime utility in the RS-RI model,

$$\widehat{v}^{RS}(\widehat{s}_t) = -\frac{R-1}{2R}\widehat{s}_t^2 + \widehat{\overline{c}s}_t - \frac{1}{2}\left(\frac{R}{R-1}\overline{c}^2 - \frac{\beta}{1-\beta}\frac{1}{\alpha}\log\left(1 - \frac{R-1}{R}\alpha\omega_\eta^2\right)\right), \tag{56}$$

and in the RB-RI model,

$$\widehat{v}^{RB}(\widehat{s}_t) = -\frac{R-1}{2R}\widehat{s}_t^2 + \widehat{\overline{c}s}_t - \frac{1}{2}\left(\frac{R}{R-1}\overline{c}^2 + \frac{\beta}{1-\beta}\frac{R-1}{R}\omega_\eta^2\right), \tag{57}$$

where $\beta = \beta^{RS,RI} = \beta^{RB,RI}$. The only difference between the two functions is the constant term; this equivalence arises because the slope and curvature parameters are pinned down by the equalization of the consumption decisions. We note that

$$\lim_{\alpha\to 0}\left\{-\frac{1}{\alpha}\log\left(1 - \frac{R-1}{R}\alpha\omega_\eta^2\right)\right\} = \frac{R-1}{R}\omega_\eta^2, \tag{58}$$

so that the two expressions are equal if $\alpha = 0$. We also note that

$$\frac{\partial}{\partial\alpha}\left(-\frac{1}{\alpha}\log\left(1 - \frac{R-1}{R}\alpha\omega_\eta^2\right)\right) > 0.$$

[31] Our calculation is related to Tallarini (2000), who shows how risk-sensitivity increases the welfare cost of business cycles and the equity premium.

Thus, we have

$$-\frac{1}{\alpha}\log\left(1-\frac{R-1}{R}\alpha\omega_\eta^2\right) > \frac{R-1}{R}\omega_\eta^2$$

for $\alpha > 0$. This result implies that RS households have lower utility than RB agents do, conditional on being observationally equivalent. We summarize the previous discussion as a formal proposition.

PROPOSITION 11: *Given the same level of the perceived state $\widehat{s}_t$ and assuming (54) is satisfied by $(\alpha, \vartheta)$ for fixed $(\theta, \beta)$, it follows that $\widehat{v}^{RS}(\widehat{s}_t) < \widehat{v}^{RB}(\widehat{s}_t)$.*

A similar proposition holds in the absence of RI simply by replacing $\omega_\eta^2$ with $\omega_\zeta^2$, so that it holds when $\theta = 1$. This result implies that the costs of uncertainty are larger for RS agents than RB agents whenever the parameters are such that their observable behavior is identical. Both types of agents respond by saving more than a standard RI agent would; under the OE restriction for parameters, they therefore generate the same consumption path. But RS agents are more averse to the remaining fluctuations in consumption than RB agents are, resulting in higher costs of uncertainty and lower lifetime utility.

PROPOSITION 12: *Define the difference between the welfare losses of RS and RB consumers as*

$$\Delta = -\frac{1}{\alpha}\log\left(1-\frac{R-1}{R}\alpha\omega_\eta^2\right) - \frac{R-1}{R}\omega_\eta^2. \tag{59}$$

*$\Delta$ is decreasing in $\theta$.*

PROOF:

By straightforward differentiation we obtain

$$\frac{\partial\Delta}{\partial\theta} = \frac{\partial\Delta}{\partial\omega_\eta^2}\frac{\partial\omega_\eta^2}{\partial\theta} < 0,$$

because

$$\frac{\partial\Delta}{\partial\omega_\eta^2} = \frac{R-1}{R-(R-1)\alpha\omega_\eta^2} - \frac{R-1}{R} > 0 \text{ for } \alpha > 0 \text{ and } \frac{\partial\omega_\eta^2}{\partial\theta} < 0.$$

RI increases the difference between the welfare losses of RS and RB consumers who have the same consumption-savings decisions; as $\theta$ decreases, the cost of RI for the RS agents increases faster than the cost for the RB agents.

Furthermore, we can use (56) and (57) to compute the *marginal* welfare losses due to RI at different channel capacities $(\theta)$.[32] Specifically, following Robert Barro (2007), the marginal

[32] We thank the anonymous referee for suggesting that we examine the marginal welfare losses due to RI under OE.

welfare costs (mwc) due to RI in the RS and RB models can be written as

$$\text{mwc}^{RS} = \frac{\partial \widehat{v}^{RS}/\partial\theta}{\left(\partial \widehat{v}^{RS}/\partial \widehat{s}_0\right)\widehat{s}_0} = \frac{-\frac{1}{2}\frac{\beta}{1-\beta}\frac{R-1}{R-(R-1)\alpha\omega_\eta^2}\frac{\partial\omega_\eta^2}{\partial\theta}}{-\frac{R-1}{R}\widehat{s}_0^2 + \overline{c}\widehat{s}_0} \tag{60}$$

$$\text{mwc}^{RB} = \frac{\partial \widehat{v}^{RB}/\partial\theta}{\left(\partial \widehat{v}^{RB}/\partial \widehat{s}_0\right)\widehat{s}_0} = \frac{-\frac{1}{2}\frac{\beta}{1-\beta}\frac{R-1}{R}\frac{\partial\omega_\eta^2}{\partial\theta}}{-\frac{R-1}{R}\widehat{s}_0^2 + \overline{c}\widehat{s}_0}, \tag{61}$$

where $\frac{\partial \widehat{v}^{RS}}{\partial\theta}$ and $\frac{\partial \widehat{v}^{RB}}{\partial\theta}$ are evaluated for given $\widehat{s}_0$, and the welfare costs due to RI are compared with that from a small proportional change in the initial level of the perceived state $\widehat{s}_0$. Therefore, the following ratio, $\pi$, measures the *relative* marginal welfare losses due to RI in the two economies at various capacities ($\theta$):

$$\pi = \frac{\text{mwc}^{RS}}{\text{mwc}^{RB}} = \frac{R}{R-(R-1)\,\alpha\omega_\eta^2}. \tag{62}$$

Expression (62) clearly shows that this ratio is different at various capacities and is determined by the interaction of channel capacity and labor income uncertainty given the interest rate $R$: $\omega_\eta^2 = \frac{\theta}{1-(1-\theta)R^2}\omega_\zeta^2$. It is worth emphasizing that $\pi$ can be rewritten as $\frac{1}{\beta R}$ as the OE holds, where $\beta = \beta^{RS,RI} = \beta^{RB,RI} = \frac{R-(R-1)\alpha\omega_\eta^2}{R^2} < \frac{1}{R}$. Therefore, $\beta$, as determined by the interaction of $R, \alpha, \theta$, and $\omega_\varepsilon^2$, measures the relative marginal welfare losses of RS and RB agents at the channel capacity $\theta$. For example, using the parameters specified in Section A, $R = 1.01$, $\alpha = 5 \times 10^{-5}$, and $\omega_\varepsilon = 5.6$, we have $\pi = 1.24$ when $\theta = 10\%$ and $\pi = 1.2$ when $\theta = 30\%$. The above simple calculations simply show that although RS agents suffer more from lower channel capacity than RB agents do, the relative marginal welfare losses of RS and RB agents may not be very significant. In the next subsection, we will calculate the welfare losses due to deviating from the first-best full information path and will then provide more quantitative results.

### B. *Quantitative Results*

To compute the welfare losses of RS and RB consumers from RI due to deviating from the first-best RE path, we need the costs of deviating from the RE benchmark. Specifically, we first assume that at the beginning both RS and RB consumers have unlimited capacity and choose the same consumption-savings decisions, that is, the OE under RE holds ($\alpha = 1/(2\vartheta)$); we then increase the degree of inattention and compute the welfare costs of RI due to deviating from the RE path. After deriving the value functions in all cases, it is straightforward to show that under

the OE, $\alpha = 1/(2\vartheta)$, the expressions for the expected welfare losses from RI are

(63)

$$E[\Delta v(\alpha)] = \frac{\beta R^2 - 1}{2}\left(\frac{1}{\beta R^2 - \alpha\omega_\eta^2} E\left[\widehat{s}_t^2\right] - \frac{1}{\beta R^2 - \alpha\omega_\zeta^2} E\left[s_t^2\right]\right) + \frac{\left(\beta R^2 - 1\right) R\overline{c}}{R-1}\left(\frac{1}{R^2\beta - \alpha\omega_\zeta^2} E[s_t] - \frac{1}{R^2\beta - \alpha\omega_\eta^2} E[\widehat{s}_t]\right) + \frac{R^2\left(R^2\beta - 1\right)\overline{c}^2}{2(R-1)^2}\left(\frac{1}{R^2\beta - \alpha\omega_\eta^2} - \frac{1}{R^2\beta - \alpha\omega_\zeta^2}\right) + \frac{\beta}{1-\beta}\frac{1}{2\alpha}\log\left(\frac{1 - \alpha\omega_\zeta^2 \frac{R^2\beta - 1}{\beta R^2 - \alpha\omega_\zeta^2}}{1 - \alpha\omega_\eta^2 \frac{R^2\beta - 1}{\beta R^2 - \alpha\omega_\eta^2}}\right),$$

where $\beta = \frac{R-(R-1)\alpha\omega_\zeta^2}{R^2}$ for the RS-RI model and

(64)
$$E[\Delta v(\vartheta)] = \frac{R-1}{2R}\left(\operatorname{var}(\widehat{s}_t) - \operatorname{var}(s_t)\right) + \frac{\beta}{1-\beta}\frac{R-1}{2R}\left(\omega_\eta^2 - \omega_\zeta^2\right),$$

where $\beta = \frac{R-(R-1)\omega_\zeta^2/(2\vartheta)}{R^2}$ for the RB-RI model.

To do quantitative welfare analysis in the models we need to know the level of $E[s_t]$ (note that it is equal to $E[\widehat{s}_t]$). First, denote by $\gamma$ the local coefficient of relative risk aversion, which equals

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

for the utility function $u(\cdot)$ evaluated at mean income $E\left[y\right]$. To calculate the welfare losses due to RI, we set the parameters according to those estimated from post-World War U.S. time series by HST (1999) for a PIH model with two-factor endowment process. Specifically, we follow the procedure used in Hansen and Sargent (2004) and use the estimated one-factor endowment process as follows

(65)
$$y_{t+1} = 0.9992 y_t + \varepsilon_{t+1},$$

and $\varepsilon_{t+1}$ follows an iid process distributed as $N(0, 0.2559)$. Since the estimated persistence coefficient is difficult to distinguish from 1, for simplicity we assume that income follows a random walk. Following HST (1999) we also set the mean income level $E\left[y\right] = 16$ and then find the value of the bliss point $\overline{c}$ that generates reasonable relative risk aversion $\gamma$. For example, if $\gamma$ is equal to 1, $\overline{c} = 2E\left[y\right] = 32$. Furthermore, assume that the ratio of mean financial wealth

to mean labor income is 5, that is, $E[w]/E[y] = 5$.[33] Since $s_t = w_t + \frac{1}{R-1} y_t$, we have

$$E[s_t] = \left(5 + \frac{1}{R-1}\right) E[y_t].$$

We will use this specification when we make quantitative statements.

Following John H. Cochrane (1989), Pischke (1995), and others, we use a money metric to measure the welfare cost of deviating from the standard RS and RB solutions. Specifically, dividing the expected welfare losses $E[\Delta v] = E[v(s_t; \theta = 1) - \widehat{v}(\widehat{s_t}; \theta < 1)]$ by the marginal utility of a dollar at time $t$ and converting it to dollars per quarter yields

$$\text{(66)} \qquad \text{\$ Loss/quarter:} \qquad \Delta^{RS} = \frac{R-1}{R} \frac{E[\Delta v(\alpha)]}{u'(\overline{y})};$$

$$\text{(67)} \qquad \text{\$ Loss/quarter:} \qquad \Delta^{RB} = \frac{R-1}{R} \frac{E[\Delta v(\vartheta)]}{u'(\overline{y})},$$

where $\Delta v(\alpha)$ and $\Delta v(\vartheta)$ are defined in (63) and (64), respectively. Table 1 reports welfare costs for several values of $\alpha$ and the optimal weight on observations $\theta$. (The parameters are set as follows: $\overline{c} = 32$, $E[y] = 16$, $E[w]/E[y] = 5$, $R = 1.01$, $\gamma = 1$, and $\omega_\varepsilon = 0.2559$.) It is clear from Table 1 that the welfare losses due to RI are trivial. For example, for $\alpha = 5 \times 10^{-4}$ and $\theta = 10\%$, (that is, 10% of the uncertainty is removed upon the receipt of a new signal), the loss only amounts to 76 cents per quarter. This result is similar to the findings by Pischke (1995) and Luo (2008).[34] Table 1 also shows that for plausible assumptions about the coefficient for local relative risk aversion ($\gamma$) and the ratio of mean financial wealth to mean income $\left(E[w]/E[y]\right)$, the welfare losses due to RI are not significant even if agents are risk-sensitive.[35] This thus provides some evidence that it is reasonable for agents to devote low channel capacity to observing and processing information because the welfare improvement from increasing capacity is trivial. In other words, although consumers can devote much more capacity to processing economic information and then improve their optimal consumption decisions, it is rational for them not to do so because the welfare improvement is tiny.

Our main purpose in this section is to demonstrate that the welfare losses from RI can in fact be considerably larger for RS agents. To this end, Table 2 reports the ratio of the welfare losses of RS to RB, $\Delta^{RS}/\Delta^{RB}$.[36] We set $\theta = 1$ and $\alpha = 1/(2\vartheta)$ to reflect the OE between robustness and risk-sensitivity in the RE context and then consider lower values of $\theta$. The costs are uniformly higher for the RS model, and can in fact be much larger (the rightmost column has ratios that approach 8). Similarly, Table 3 reports the ratio of the RI-induced welfare losses in the RS-RI and basic RI models for different values of $\theta$ given observationally-equivalent consumption-

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

[34] Pischke (1995) calculates utility losses in the no-aggregate-information model and finds that in most cases the utility losses due to no information about aggregate shocks are less than \$1 per quarter.

[35] We find that changing the value of $E[w]/E[y]$ from 5 to 10 only has minor effects on quantitative results and does not affect the main results of the RS-RI model.

[36] In Tables 2 and 3, the parameters are set to be the same as that used in Table 1.

TABLE 1—WELFARE LOSSES UNDER RS-RI, $\Delta^{RS}$ (\$/QUARTER)

| | $\alpha = 10^{-4}$ | $\alpha = 5 \cdot 10^{-4}$ | $\alpha = 10^{-3}$ |
|---|---|---|---|
| $\theta = 10\%$ | 0.1389 | 0.7566 | 3.1455 |
| $\theta = 40\%$ | 0.0193 | 0.0984 | 0.3263 |
| $\theta = 70\%$ | 0.0054 | 0.0272 | 0.0885 |
| $\theta = 90\%$ | 0.0014 | 0.0070 | 0.0226 |

savings decisions. Note that when calculating the welfare losses in the RS model, the value of $\beta$ is determined by (52) given the values of $\alpha$ and $\theta$, while the value of $\beta$ used in the basic RI model is just $1/R$. Risk-sensitive consumers also suffer much more from finite information-processing constraints than do the consumers who are simply more patient. It is worth noting that as (54) shows, changing channel capacity while holding RS and RB fixed would break the OE; consequently, the welfare consequences reported above are partly affected by the different consumption-savings rules.

TABLE 2—COMPARISON OF WELFARE LOSSES IN THE RS-RI AND RB-RI MODELS $\left(\Delta^{RS}/\Delta^{RB}\right)$

| | $(\alpha, \vartheta) = \left(10^{-5}, 5 \cdot 10^{4}\right)$ | $\left(5 \cdot 10^{-5}, 10^{4}\right)$ | $\left(10^{-4}, 5 \cdot 10^{3}\right)$ | $\left(5 \cdot 10^{-4}, 10^{3}\right)$ |
|---|---|---|---|---|
| $\theta = 10\%$ | 2.5040 | 4.7313 | 7.9944 | 69.8826 |
| $\theta = 40\%$ | 2.5013 | 4.7050 | 7.9028 | 64.5429 |
| $\theta = 70\%$ | 2.5012 | 4.7023 | 7.8928 | 63.9736 |
| $\theta = 90\%$ | 2.5009 | 4.7010 | 7.8891 | 63.7887 |

TABLE 3—COMPARISON OF WELFARE LOSSES IN THE RS-RI AND BASIC RI MODELS $\left(\Delta^{RS}/\Delta^{\beta}\right)$

| | $(\alpha, \beta) = \left(10^{-5}, 0.9901\right)$ | $\left(5 \cdot 10^{-5}, 0.9901\right)$ | $\left(10^{-4}, 0.9901\right)$ | $\left(5 \cdot 10^{-4}, 10^{3}\right)$ |
|---|---|---|---|---|
| $\theta = 10\%$ | 2.4757 | 4.4690 | 7.1318 | 38.8461 |
| $\theta = 40\%$ | 2.4730 | 4.4442 | 7.0501 | 35.8781 |
| $\theta = 70\%$ | 2.4728 | 4.4416 | 7.0411 | 35.5598 |
| $\theta = 90\%$ | 2.4726 | 4.4404 | 7.0379 | 35.4621 |

What is the significance of these calculations? The fact that RS agents suffer much more from limited information processing capacity suggests that a careful experiment might be able to distinguish the three types of agents. If a researcher were to ask individuals 'What would you pay to receive more/better information about the state of the economy?' RS-RI, RB-RI, and RI agents would provide significantly different answers. Given the answers to this question,

along with consumption-savings data, we may be able to determine whether agents are risk-sensitive, concerned about model misspecification, or more patient, provided they in fact have limited information-processing capacity.

### C. *Implications for Countercyclical Policies*

Finally, it is worth noting that we can also distinguish RS, RB, and the discount factor by examining different welfare gains of stabilization policy aimed at reducing aggregate fluctuations for RS-RI consumers, RB-RI consumers, and RI consumers with higher discount factors. The traditional way to calculate welfare costs of business cycles proposed by Robert E. Lucas (1987) is to offer risk averse consumers two possible consumption streams, one of which is constant and the other has the same mean but fluctuates. Consequently, risk averse consumers would always prefer the constant consumption stream and thus require some consumption compensation to accept the fluctuating path. As shown above, RI introduces additional uncertainty into RS and RB models and thus amplifies the impacts of RS and RB on total uncertainty faced by the consumers as well as precautionary savings. Since counter-cyclical policies (stabilization policies) which reduce aggregate fluctuations affect our model economy by reducing $\omega_\zeta^2$, RS-RI consumers gain more welfare from countercyclical policies than do RB-RI consumers and RI consumers with higher discount factors.[37]

## V. Conclusion

This paper has provided a characterization of the consumption-savings behavior of a single agent who is both risk-sensitive and limited in their capacity to process signals. The key component to our model is the presence of precautionary savings, permitting the limited processing capacity to affect the average level of consumption. We showed that there still exists an observational equivalence between models with and without risk-sensitivity (and/or robustness) even when agents face information-processing constraints, extending the results of HST (1999) to a broader class of models. However, within the observationally-equivalent class of models the welfare costs of rational inattention are not constant – models with higher discount factors and correspondingly lower risk-sensitivity generate larger welfare costs of RI than observationally-equivalent ones with lower discount factors and higher risk-sensitivity. We also showed a connection between robustness, risk-sensitivity, and the discount factor, in the sense that any combination of the three can plausibly generate the same consumption-savings decisions in the RI context. However, as in the case with RS and RI, the welfare costs of RI are not equivalent across these observably-equivalent specifications.

While our model does permit precautionary savings, it is not as general as we would prefer. As we have noted above, solving fully-nonlinear versions of the RI model is extremely difficult given the shape of the optimal distribution of consumption, meaning that only models with short

[37] Note that $\omega_\zeta^2$ includes both idiosyncratic and aggregate components. Countercyclical policies can reduce aggregate fluctuations by either reducing individual income risk directly or reducing the correlation across individuals in their income risk.

horizons or very small state spaces can be computed.[38] We suspect that methods to deal with this problem can be developed, particularly for simple one-agent problems like the one studied here. A more difficult problem with RI models is noted in Sims (2005) – a general equilibrium setting will involve nonstandard features including inventories, retailing, and probably search. Implementation of RI solutions is also problematic – under general settings consumption is not a deterministic function, meaning that the theory does not make clear predictions about what any individual's consumption will be, only what will prevail in a dataset of sufficiently large size. These difficulties will need to be overcome in order to make rational inattention a useful component of economic models.

Because our model is not necessarily general, we think it advisable to comment on whether the observational equivalence results are likely to hold in more general environments. The precautionary savings induced by a concern about robustness (or enhanced risk aversion) differs from the usual precautionary savings motive that emerges when labor income uncertainty interacts with the convexity of the marginal utility of consumption. This type of precautionary savings emerges because consumers need to save more to protect themselves against model misspecification and occurs even in models with quadratic utilities. Most existing robustness models assume that the objective functions are quadratic and the state transition equation is linear, consequently, worst case distributions are Gaussian. However, if the objective functions are not quadratic or the transition equations are not linear, worst case distributions are generally *non-Gaussian*. As discussed in Hansen and Sargent (2007), the most difficult part in solving such non-LQ models is computational: representing the worst-case distribution parsimoniously enough that the model is tractable.

Some progress in establishing a connection between risk-sensitivity and robustness has been made, though. In a continuous-time non-LQ setting, Pascal Maenhout (2004) shows that robustness is observationally equivalent to Duffie-Epstein-Zin recursive preferences (which contain risk-sensitive preferences as a strict subset): an investor with a preference for robustness measured by $\alpha$ and CRRA utility function $\frac{C_t^{1-\gamma}-1}{1-\gamma}$ is observationally equivalent to a Duffie-Epstein-Zin investor with EIS $\frac{1}{\gamma}$ and coefficient of relative risk aversion $\gamma + \alpha$, in the sense that they generate the same consumption and portfolio choice rules.[39] Anderson, Hansen, and Sargent (1998) extend the analysis of HST (1999) to non-LQ continuous-time economies that solve optimal resource allocation problems and also demonstrate that concerns about robustness can imitate risk sensitivity in the sense that they lead to the same resource allocation. However, to the best of our knowledge, it seems impossible to theoretically establish the OE in the discrete-time RE non-LQ setting, as most of them have to be solved numerically.

Under RI, it is even more difficult to establish the OE between RS and RB outside the LQ framework. RI means that imperfect information emerges endogenously and greatly complicates RB or RS models. The main reason is that the effective state in such models is not the traditional state variable, but is the so-called information state that is defined as the distribution of the true state conditional on processed information. In other words, it expands the state space to the

[38] Sims (2006), Lewis (2006), Batchuluun, Luo, and Young (2008), and Tutino (2008) present some nonlinear solutions to RI models. These models are generally very small, either in terms of the number of states or the number of periods.

[39] Backus, Routledge, and Zin (2004) also comment that the connection between the two models would hold outside the confines of the LQ-Gaussian model, but do not provide any formal proof.

space of distributions on the true state (in our paper, $s_t$), creating a severe "curse of dimensionality" problem. Within the LQ-Gaussian framework the conditional distributions are Gaussian, so that the first two moments, the conditional mean $E_t[s_t]$ and the conditional covariance matrix $\text{var}_t(s_t)$, are sufficient to characterize the information state. However, it is extremely difficult to solve non-LQ versions of the RI model, as the state of the world is a distribution and this distribution is not well-behaved except in rare circumstances (Sims 2005 finds one case where the posterior is distributed as a $F$ random variate, but this result only holds for a uniform prior and thus would not extend to horizons longer than two periods); consequently, only models with short horizons or very small state spaces can be computed, and those only to limited accuracy. It seems unlikely that observational equivalence can be proven. Some current work is focused on establishing an approximate result – how the introduction of RI alters our estimates of preference parameters.[40]

We conclude with some comments on the consequences of our results. Risk-sensitivity and concern for robustness are aspects of preferences; government policy should not be used to "counteract" their effects on consumption. But rational inattention is something different – it is an additional constraint on household actions that reduces welfare. Thus, the effects of rational inattention should be attenuated via policy whenever such actions are cost-effective and feasible. Our results suggest that it is important to dig deeper into how rational inattentive agents differ from other agents, for the optimal design of policies – indeed, even whether government policy is desirable or not – will depend on why agents choose the way they do.

Finally, there may be substantive differences across the various models that only manifest themselves in asset markets. In this paper we study a set of permanent income models in which the interest rate is constant. In a more realistic setting in which there are multiple assets and agents have to make portfolio choices, we can examine the distinct implications of the risk-sensitivity preference (a special case of the recursive utility theory) and rational inattention for long-run consumption risk – measured by the covariance of the asset return and consumption growth over the period of the return and many following periods. Jonathan A. Parker and Christian Julliard (2005) and Lars Peter Hansen, John C. Heaton, and Nan Li (2008) find strong empirical support for long-run consumption risk. Hansen, Heaton, and Li (2008) also show that recursive preferences provide a preference-based role for long-run consumption risk, as temporal dependence in consumption growth alters risk premia. In contrast, RI due to finite information-processing capacity affects the intertemporal composition of risk, as the RI model predicts that consumption reacts to the innovation to the equity return gradually and with delay (see Luo 2009). In the future it would be promising to examine how the two hypotheses lead to different implications for optimal portfolio choice and the equilibrium risk premium.

[40] Batchuluun (2009) and Batchuluun, Luo, and Young (2008) contain preliminary results on the connections between risk aversion, intertemporal substitution in consumption and labor supply, and channel capacity.

## VI. Appendix

### A. Solving the RS model

To solve the Bellman equation (14), we conjecture that

$$v(s_t) = -As_t^2 - Bs_t - C, \tag{68}$$

where $A$, $B$, and $C$ are undetermined coefficients. We can then evaluate $E_t\left[\exp\left(-\alpha v\left(s_{t+1}\right)\right)\right]$ to obtain

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

where

$$a = \alpha A R^2 (s_t - c_t)^2 + \alpha B R (s_t - c_t) + \alpha C, \tag{69}$$

$$b = [2\alpha A R (s_t - c_t) + \alpha B]\,\omega_\zeta, \tag{70}$$

$$c = \alpha A \omega_\zeta^2. \tag{71}$$

Thus, the distorted expectations operator can be written as

$$\begin{aligned}
\mathcal{R}_t\left[v\left(s_{t+1}\right)\right] &= -\frac{1}{\alpha}\left\{-\frac{1}{2}\log(1-2c) + a + \frac{b^2}{2(1-2c)}\right\} \\
&= \frac{1}{2\alpha}\log\left(1 - 2\alpha A \omega_\zeta^2\right) - \frac{AR^2}{1-2\alpha A\omega_\zeta^2}(s_t - c_t)^2 - \frac{BR}{1-2\alpha A\omega_\zeta^2}(s_t - c_t) - \left[C + \frac{\alpha B^2 \omega_\zeta^2}{2\left(1-2\alpha A\omega_\zeta^2\right)}\right].
\end{aligned}$$

Maximizing the RHS of (14) with respect to $c_t$ yields the first-order condition

$$-(c_t - \overline{c}) + \frac{2\beta A R^2}{1-2\alpha A\omega_\zeta^2}(s_t - c_t) + \frac{B\beta R}{1-2\alpha A\omega_\zeta^2} = 0, \tag{72}$$

which means that

$$c_t = \frac{2A\beta R^2}{1-2\alpha A\omega_\zeta^2 + 2A\beta R^2} s_t + \frac{\overline{c}\left(1-2\alpha A\omega_\zeta^2\right) + B\beta R}{1-2\alpha A\omega_\zeta^2 + 2A\beta R^2}. \tag{73}$$

Substituting (73) and (68) into (14) gives

$$-As_t^2 - Bs_t - C = -\frac{1}{2}\left(\frac{2A\beta R^2}{1-2\alpha A\omega_\zeta^2 + 2A\beta R^2}s_t + \frac{B\beta R - 2A\beta R^2\overline{c}}{1-2\alpha A\omega_\zeta^2 + 2A\beta R^2}\right)^2 +$$
$$\beta\left[\begin{array}{l} \frac{1}{2\alpha}\log\left(1-2\alpha A\omega_\zeta^2\right) - \frac{AR^2}{1-2\alpha A\omega_\zeta^2}\left(\frac{1-2\alpha A\omega_\zeta^2}{1-2\alpha A\omega_\zeta^2+2A\beta R^2}s_t - \frac{\overline{c}\left(1-2\alpha A\omega_\zeta^2\right)+B\beta R}{1-2\alpha A\omega_\zeta^2+2A\beta R^2}\right)^2 \\ -\frac{BR}{1-2\alpha A\omega_\zeta^2}\left(\frac{1-2\alpha A\omega_\zeta^2}{1-2\alpha A\omega_\zeta^2+2A\beta R^2}s_t - \frac{\overline{c}\left(1-2\alpha A\omega_\zeta^2\right)+B\beta R}{1-2\alpha A\omega_\zeta^2+2A\beta R^2}\right) - \left(C + \frac{\alpha B^2\omega_\zeta^2}{2\left(1-2\alpha A\omega_\zeta^2\right)}\right) \end{array}\right].$$

Collecting and matching terms, the constant coefficients turn out to be

$$A = \frac{\beta R^2 - 1}{2\beta R^2 - 2\alpha\omega_\zeta^2}, \tag{74}$$

$$B = \frac{R\left(1-\beta R^2\right)}{\left(\beta R^2 - \alpha\omega_\zeta^2\right)(R-1)}\overline{c}, \tag{75}$$

$$C = \frac{1}{2}\frac{R^2\left(\beta R^2 - 1\right)}{(R-1)^2\left(\beta R^2 - \alpha\omega_\zeta^2\right)}\overline{c}^2 + \frac{\beta}{1-\beta}\frac{1}{2\alpha}\log\left(1-2\alpha A\omega_\zeta^2\right), \tag{76}$$

Substituting (74) and (75) into (73) yields the consumption function (17) in the text.

### B. *Solving the RI-RS PIH model*

To obtain the solution to the RI-RS model we simply need to replace $\omega_\zeta^2$ with $\omega_\eta^2$ (we omit the steps):

$$A = \frac{\beta R^2 - 1}{2\beta R^2 - 2\alpha\omega_\eta^2}$$
$$B = \frac{R\left(1-\beta R^2\right)}{\left(\beta R^2 - \alpha\omega_\eta^2\right)(R-1)}\overline{c}$$
$$C = \frac{1}{2}\frac{R^2\left(\beta R^2 - 1\right)}{(R-1)^2\left(\beta R^2 - \alpha\omega_\eta^2\right)}\overline{c}^2 - \frac{\beta}{1-\beta}\frac{1}{2\alpha}\log\left(1-2\alpha A\omega_\eta^2\right)$$

and the consumption function is

$$c_t = \frac{\beta R^2 - 1}{\beta R^2 - \alpha\omega_\eta^2}\widehat{s_t} + \frac{\left(1-\alpha\omega_\eta^2\right)(R-1) + 1 - \beta R^2}{\left(\beta R^2 - \alpha\omega_\eta^2\right)(R-1)}\overline{c} \tag{77}$$

### C. Solving the Robust RI Model

As before, we conjecture that $\widehat{v}(\widehat{s}_t) = -A\widehat{s}_t^2 - B\widehat{s}_t - C$, where $A$, $B$, and $C$ are undetermined coefficients. Substituting this guessed value function, (46), and (47) into the Bellman equation gives

$$-A\widehat{s}_t^2 - B\widehat{s}_t - C = \max_{c_t} \min_{\nu_t} \left\{ -\frac{1}{2}(\overline{c} - c_t)^2 + \beta E_t \left[ \vartheta \nu_t^2 - A\widehat{s}_{t+1}^2 - B\widehat{s}_{t+1} - C \right] \right\}. \tag{78}$$

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

$$2\vartheta \nu_t - 2AE_t \left[ \omega_\zeta \nu_t + R(\widehat{s}_t - c_t) \right] \omega_\zeta - B\omega_\zeta = 0,$$

which means that

$$\nu_t = \frac{B + 2AR(\widehat{s}_t - c_t)}{2\left(\vartheta - A\omega_\zeta^2\right)} \omega_\zeta. \tag{79}$$

Substituting (79) back the Bellman equation (45) gives

$$-A\widehat{s}_t^2 - B\widehat{s}_t - C = \max_{c_t} \left\{ -\frac{1}{2}(\overline{c} - c_t)^2 + \beta E_t \left[ \vartheta \left[ \frac{B + 2AR(\widehat{s}_t - c_t)}{2\left(\vartheta - A\omega_\zeta^2\right)} \omega_\zeta \right]^2 - A\widehat{s}_{t+1}^2 - B\widehat{s}_{t+1} - C \right] \right\}, \tag{80}$$

where

$$\begin{aligned} \widehat{s}_{t+1} &= (1-\theta)\left[ R(\widehat{s}_t - c_t) + \omega_\zeta \nu_t \right] + \theta \left( R(s_t - c_t) + \omega_\zeta \nu_t + \zeta_{t+1} + \xi_{t+1} \right) \\ &= R(\widehat{s}_t - c_t) + \omega_\zeta \nu_t + \eta_{t+1} \end{aligned}$$

The first-order condition for $c_t$ is

$$(\overline{c} - c_t) + 2\beta A \left( R + \frac{AR\omega_\zeta^2}{\vartheta - A\omega_\zeta^2} \right) \left[ R(\widehat{s}_t - c_t) + \omega_\zeta \nu_t \right] + \beta B \left( R + \frac{AR\omega_\zeta^2}{\vartheta - A\omega_\zeta^2} \right) - 2\beta\vartheta \frac{AR\omega_\zeta}{\vartheta - A\omega_\zeta^2} \nu_t = 0.$$

Using the solution for $\nu_t$ the solution for consumption is

$$c_t = \frac{2A\beta R^2}{1 - A\omega_\zeta^2/\vartheta + 2A\beta R^2} \widehat{s}_t + \frac{\overline{c}\left(1 - A\omega_\zeta^2/\vartheta\right) + B\beta R}{1 - A\omega_\zeta^2/\vartheta + 2A\beta R^2}. \tag{81}$$

Substituting the above expressions into the Bellman equation gives

$$
\begin{aligned}
&-A\widehat{s}_t^2 - B\widehat{s}_t - C \\
&= \left\{ \begin{array}{c} -\frac{1}{2}\left(\frac{2A\beta R^2}{1-A\omega_\zeta^2/\vartheta+2A\beta R^2}\widehat{s}_t + \frac{-2A\beta R^2\overline{c}+B\beta R}{1-A\omega_\zeta^2/\vartheta+2A\beta R^2}\right)^2 + \beta\vartheta\left[\frac{1}{2}\frac{B-2ARc}{2A\beta\vartheta R^2-A\omega_\zeta^2+\vartheta} + \frac{AR}{\vartheta-A\omega_\zeta^2+2A\beta R^2\vartheta}\widehat{s}_t\right]^2\omega_\zeta^2 \\ +\beta E_t\left[\begin{array}{c} -A\left\{\frac{R\vartheta}{2A\beta\vartheta R^2-A\omega_\zeta^2+\vartheta}\widehat{s}_t - \frac{1}{2}\frac{2B\beta\vartheta R^2+2c\vartheta R-B\omega_\zeta^2}{2A\beta\vartheta R^2-A\omega_\zeta^2+\vartheta} + \eta_{t+1}\right\}^2 \\ -B\left\{\frac{R\vartheta}{2A\beta\vartheta R^2-A\omega_\zeta^2+\vartheta}\widehat{s}_t - \frac{1}{2}\frac{2B\beta\vartheta R^2+2c\vartheta R-B\omega_\zeta^2}{2A\beta\vartheta R^2-A\omega_\zeta^2+\vartheta} + \eta_{t+1}\right\} - C \end{array}\right] \end{array} \right\}
\end{aligned}
$$

Collecting and matching terms, the constant coefficients turn out to be

$$A = \frac{\beta R^2 - 1}{2\beta R^2 - \omega_\zeta^2/\vartheta}, \tag{82}$$

$$B = \frac{R}{R-1}\frac{1-\beta R^2}{\beta R^2 - \omega_\zeta^2/(2\vartheta)}\overline{c}, \tag{83}$$

$$C = \frac{1}{2}\frac{R^2\left(\beta R^2 - 1\right)}{\left(\beta R^2 - \omega_\zeta^2/(2\vartheta)\right)(R-1)^2}\overline{c}^2 + \frac{\beta}{1-\beta}\frac{\beta R^2 - 1}{2\beta R^2 - \omega_\zeta^2/\vartheta}\omega_\eta^2. \tag{84}$$

Substituting (82) and (83) into (81) yields the consumption function (50) in the text.

### D. *Observational Equivalence Between RI-RS and RI-RB*

Comparing (35) with (50), we can derive the OE between RS and RB under RI:

$$\alpha\omega_\eta^2 = \omega_\zeta^2/(2\vartheta), \tag{85}$$

which implies that

$$\alpha\frac{\theta}{1-(1-\theta)R^2} = \frac{1}{2\vartheta}. \tag{86}$$

Note that when $\beta R = 1$, the standard RI PIH model implies that $c_t = \frac{R-1}{R}\widehat{s}_t$. Hence, under RS-RI ($\alpha > 0$), we can construct the following OE between $\alpha$ and $\beta$

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

which implies that

$$\beta^{RI,RS} = \frac{R-(R-1)\alpha\omega_\eta^2}{R^2} < \frac{1}{R}; \tag{87}$$

Similarly, under RS-RI ($\vartheta > 0$), we can construct the following OE between $\beta$ and $\vartheta$

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

which implies that

$$\beta^{RI,RB} = \frac{R - (R-1)\,\omega_\zeta^2 / (2\vartheta)}{R^2} < \frac{1}{R}.$$

## REFERENCES


Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent. 1998. "Risk and Robustness in General Equilibrium." Unpublished.

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

Barro, Robert. 2007. "On the Welfare Costs of Consumption Uncertainty." http://www.economics.harvard.edu/faculty/barro/workpapers_barro.

Batchuluun, Altantsetseg. 2009. "Labor Supply with Information-Processing Constraints." Unpublished.

Batchuluun, Altantsetseg, Yulei Luo, and Eric R. Young. 2008. "Portfolio Choice with Information-Processing Limits." http://people.virginia.edu/~ey2d/sv1.pdf.

Carroll, Christopher, and Andrew Samwick. 1997. "The Nature of Precautionary Wealth." *Journal of Monetary Economics,* 40(1): 41–71.

Cochrane, John H. 1989. "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives." *American Economic Review,* 79(3): 319-37.

Cover, Thomas M. and Joy A. Thomas. 1991. *Elements of Information Theory*. John Wiley and Sons.

Deaton, Angus. 1992. *Understanding Consumption*. Oxford University Press, Oxford, U.K.

Duffie, Darrell and Larry G. Epstein. 1992. "Stochastic Differential Utility." *Econometrica,* 60(2): 353–94.

Eeckhoudt, Louis and Harris Schlesinger. 2008. "Changes in Risk and the Demand for Saving." *Journal of Monetary Economics,* 55(7): 1329-36.

Epstein, Larry G. and Stanley E. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica,* 57(4): 937-69.

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

Friedman, Milton. 1957. *A Theory of the Consumption Function*. Cambridge, MA: National Bureau of Economic Research.

Guiso, Luigi, Tullio Jappelli and Daniele Terlizzese. 1992. "Earning Uncertainty and Precautionary Saving." *Journal of Monetary Economics,* 30(2): 307-38.

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

Hansen, Lars Peter, John C. Heaton, and Nan Li. 2008. "Consumption Strikes Back? Measuring Long Run Risk." *Journal of Political Economy,* 116(2): 260-302.

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

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

Hansen, Lars Peter and Thomas J. Sargent. 2007. *Robustness*. Princeton, NJ: Princeton University Press.

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.

Hansen, Lars Peter, Thomas J. Sargent, and Neng Wang. 2002. "Robust Permanent Income and Pricing with Filtering." *Macroeconomic Dynamics,* 6(1): 40-84.

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.

Kasa, Kenneth. 2006. "Robustness and Information Processing." *Review of Economic Dynamics,* 9(1): 1-33.

Lewis, Kurt F. 2006. "The Life-Cycle Effects of Information-Processing Constraints." Unpublished.

Lucas, Robert E. 1987. *Models of Business Cycles*. New York: Basil Blackwell.

Luo, Yulei. 2008. "Consumption Dynamics Under Information Processing Constraints." *Review of Economic Dynamics,* 11(2): 366-85.

Luo, Yulei. 2009. "Rational Inattention, Long-run Consumption Risk, and Portfolio Choice." http://yulei.luo.googlepages.com.

Luo, Yulei and Eric R. Young. 2009. "Rational Inattention and Aggregate Fluctuations." *The B.E. Journal of Macroeconomics: Vol. 9: Iss. 1* (*Contributions*), Article 14.

Maenhout, Pascal. 2004. "Robust Portfolio Rules and Asset Pricing." *Review of Financial Studies,* 17(4): 951-83.

Parker, Jonathan A. and Christian Julliard. 2005. "Consumption Risk and the Cross Section of Expected Returns." *Journal of Political Economy,* 113(1): 185-222.

Pischke, Jorn-Steffen. 1995. "Individual Income, Incomplete Information, and Aggregate Consumption." *Econometrica,* 63(4): 805-40.

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

Shannon, Claude. 1948. "A Mathematical Theory of Communication." *The Bell System Technical Journal,* 27: 623-56.

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

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics,* 50(3): 665-90.

Sims, Christopher A. 2005. "Rational Inattention: A Research Agenda." http://sims.princeton.edu/yftp/RIplus/RatInattPlus.pdf.

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

Tallarini, Thomas D., Jr. 2000. "Risk-Sensitive Real Business Cycles." *Journal of Monetary Economics,* 45(3): 507-32.

Tutino, Antonella. 2008. "The Rigidity of Choice: Life Cycle Savings with Information-Processing Limits." http://www.federalreserve.gov/pubs/feds/2008/200862/200862pap.pdf.

van der Ploeg, Frederick. 1993. "A Closed-Form Solution for a Model of Precautionary Saving." *Review of Economic Studies,* 60(2): 385-95.

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

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

Whittle, Peter. 1990. *Risk-Sensitive Optimal Control*. John Wiley and Sons.