Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 163 (2016) 1–41

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Induced uncertainty, market price of risk, and the dynamics of consumption and wealth ☆

Yulei Luo [a,*], Eric R. Young [b]

[a] *Faculty of Business and Economics, The University of Hong Kong, Hong Kong*
[b] *Department of Economics, University of Virginia, Charlottesville, VA 22904, United States*

Received 16 March 2015; final version received 25 November 2015; accepted 9 January 2016

Available online 18 January 2016

Abstract

In this paper we examine the implications of model uncertainty or robustness (RB) for consumption and saving and the market price of uncertainty under limited information-processing capacity (rational inattention or RI). First, we show that RI by itself creates an additional demand for robustness that leads to higher "induced uncertainty" facing consumers. Second, if we allow capacity to be elastic, RB increases the optimal level of attention. Third, we explore how the induced uncertainty composed of (i) model uncertainty due to RB and (ii) state uncertainty due to RI, affects consumption and wealth dynamics, the market price of uncertainty, and the welfare losses due to incomplete information. We find that induced uncertainty can

☆ We are grateful to Ricardo Lagos (editor), an associate editor, and two anonymous referees for many constructive suggestions and comments, and to Tom Sargent for his invaluable guidance and discussions. We also would like to thank Anmol Bhandari, Jaime Casassus, Richard Dennis, Larry Epstein, Hanming Fang, Lars Hansen, Ken Kasa, Tasos Karantounias, Jae-Young Kim, Rody Manuelli, Jun Nie, Kevin Salyer, Martin Schneider, Chris Sims, Wing Suen, Laura Veldkamp, Mirko Wiederholt, Tack Yun, Shenghao Zhu, and Tao Zhu as well as seminar and conference participants at UC Davis, Hong Kong University of Science and Technology, City University of Hong Kong, University of Tokyo, Shanghai University of Finance and Economics, National University of Singapore, Seoul National University, the conference on "Putting Information Into (or Taking it out of) Macroeconomics" organized by LAEF of UCSB, the Summer Meeting of Econometric Society, the conference on "Rational Inattention and Related Theories" organized by CERGE-EI, Prague, the KEA annual meeting, the Fudan Conference on Economic Dynamics, and the Workshop on the Macroeconomics of Risk and Uncertainty at the Banco Central de Chile for helpful discussions and comments. Luo thanks the General Research Fund (GRF Nos. HKU749711 and HKU791913) in Hong Kong for financial support. Young thanks the Bankard Fund for Political Economy at the University of Virginia for financial support. All errors are the responsibility of the authors.

* Corresponding author.
E-mail addresses: yulei.luo@gmail.com (Y. Luo), ey2d@virginia.edu (E.R. Young).

http://dx.doi.org/10.1016/j.jet.2016.01.003

better explain the observed consumption-income volatility and market price of uncertainty – low attention increases the effect of model misspecification.

JEL classification: C61; D81; E21

Keywords: Robust control and filtering; Optimal inattention; Induced uncertainty; Market prices of uncertainty; Consumption and income volatility

1. Introduction

Hansen and Sargent (1995) first introduced robustness (RB, a concern for model misspecification) into linear-quadratic (LQ) economic models.[1] In robust control problems, agents do not know the true data-generating process and are concerned about the possibility that their model (denoted the approximating model) is misspecified; consequently, they choose optimal decisions as if the subjective distribution over shocks was chosen by an evil nature in order to minimize their expected utility.[2] Robustness (RB) models produce precautionary savings but remain within the class of LQ models, which leads to analytical simplicity. The effects of RB can be understood by viewing decisions through a related model, namely the risk-sensitive (RS) framework from Hansen and Sargent (1995) and Hansen et al. (1999) (henceforth HST). In the RS model agents effectively compute expectations through a distorted lens, increasing their effective risk aversion by overweighting negative outcomes. The resulting decision rules depend explicitly on the variance of the shocks, producing precautionary savings, but the value functions are still quadratic functions of the states. As shown in Hansen and Sargent (2007), risk-sensitive preferences can be used to interpret the desire for robustness as both models lead to the same consumption-saving decisions and similar asset pricing implications.[3]

Sims (2003) first introduced rational inattention into economics and argued that it is a plausible method for introducing sluggishness, randomness, and delay into economic models. In his formulation agents have finite Shannon channel capacity, limiting their ability to process signals about the true state of the world. As a result, an impulse to the economy induces only gradual responses by individuals, as their limited capacity requires many periods to discover just how much the state has moved. Since RI introduces additional uncertainty, the endogenous noise due

[1] See Hansen and Sargent (2007) for a textbook treatment on robustness. For decision-theoretic foundations of the robustness preference, see Maccheroni et al. (2006) and Strzalecki (2011) for detailed discussions. It is worth noting that both the preference for "wanting robustness" proposed by Hansen and Sargent and "ambiguity aversion" proposed by Epstein and his coauthors (e.g., Epstein and Wang, 1994) can be used to capture the same idea of the multiple priors model of Gilboa and Schmeidler (1989). See Epstein and Schneider (2010) for a recent review on this topic. In this paper, we use Hansen and Sargent's "wanting robustness" specification to introduce model misspecification.

[2] The solution to a robust decision-maker's problem is the equilibrium of a max–min game between the decision-maker and nature.

[3] An alternative tractable setup is constant absolute risk aversion preferences (CARA). Although both RB (or RS) and CARA preferences (i.e., Caballero, 1990 and Wang, 2003) increase the precautionary savings premium via the intercept terms in the consumption function, they have distinct implications for the marginal propensity to consume out of permanent income (MPC). Specifically, CARA preferences do not alter the MPC relative to the LQ case, whereas RB or RS increases the MPC. That is, under RB, in response to a negative wealth shock the consumer would choose to reduce consumption more than that predicted in the CARA model (i.e., save more to protect themselves against the negative shock).

to finite capacity, into economic models, RI by itself creates an additional demand for robustness. In addition, agents with finite capacity need to use a filter to update their perceived state upon receiving noisy signals, which may lead to another demand for robustness, namely robustness against the process generating the filtering errors; in response, agents would use the robust Kalman filter.[4]

In this paper we construct a discrete-time robust permanent income model with inattentive consumers who have concerns about two types of model misspecification: (i) the disturbances to the perceived permanent income (the disturbances here include both the fundamental shock and the RI-induced noise shock) and (ii) the Kalman gain.[5] For ease of presentation, we will refer to the first type of model misspecification as Type I and the second as Type II.[6] In the standard RI problem, the decision-maker (DM) combines a pre-specified prior over the state with the new noisy state observations to construct the perceived value of the state, and is assumed to have only a single prior (i.e., no concerns about model misspecification). However, given the difficulty in estimating permanent income, the sensitivity of optimal decisions to finite capacity, and the substantial empirical evidence that agents are not neutral to ambiguity, it is important to consider inattentive consumers with multiple priors who are concerned about model misspecification and hence desire robust decision rules that work well for a set of possible models. The optimal consumption-saving problem under RI and RB can be formulated by making two additions to the standard full-information rational expectations (FI-RE) model: (i) imposing an additional constraint on the information-processing ability of the DM that gives rise to endogenous noises; and (ii) introducing an additional minimization over the set of probability models subject to the additional constraint. The additional constraint recognizes that the probability model of the perceived state is not unique. Furthermore, the additional minimization procedure reflects the preference for robustness of the DM who understands that he only has finite information-processing capacity.

We first examine how a desire for robustness affects optimal consumption and precautionary savings via interactions with finite capacity. We show that, given finite capacity, concerns about the two types of model misspecification have opposing impacts on the marginal propensity to consume out of perceived permanent income (MPC) and precautionary savings. In the case with only Type I model misspecification, 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.[7] As for Type II misspecification, an increase in the strength of the preference for robustness increases the Kalman gain, which leads to lower total uncertainty about the true level of permanent income and then lower precautionary savings. In addition, the strength of the

[4] The key assumption in Luo and Young (2010) is that agents with finite capacity distrust their budget constraint, but still use an ordinary Kalman filter to estimate the true state; in this case, a distortion to the mean of permanent income is introduced to represent possible model misspecification. However, this case ignores the effect of the RI-induced noise on the demand for robustness.

[5] Anderson et al. (2003) provided a general framework to study and quantify robustness in continuous-time. See Cagetti et al. (2002) and Maenhout (2004) for the applications of robustness in pricing, growth, and portfolio choice in continuous-time.

[6] When modeling Type II misspecification, we assume that the agent faces the commitment on the part of the minimizing agent to previous distortions.

[7] Luo et al. (2012) applied Type I RB in a small-open economy real business cycles (RBC) model and showed that this type of RB can help generate realistic relative volatility of consumption to income and the current account dynamics observed in emerging and developed small-open economies.

precautionary effect is positively related to the amount of this uncertainty that always increases as finite capacity gets smaller.

We also show that increasing RB increases the robust Kalman filter gain and thus leads to lower relative volatility of consumption to income (a smoother consumption process) when we only consider Type II misspecification. In contrast, RB increases the relative volatility of consumption by increasing the MPC out of changes in permanent income when we only consider Type I misspecification. After inspecting the consumption and saving decisions, we find that Type I misspecification dominates the Type II misspecification in the robust control and filtering problem under RI. In addition, we show that the ex post Gaussianity and additive iid Gaussian noise that obtained in the RI-LQG model are still optimal in the presence of RB. Specifically, although introducing RB can significantly change the RI model's dynamics and welfare implications, it does not change the key properties of the ex post distribution of the state and the RI-induced noise.

Furthermore, we show that if we assume that the marginal cost of information-processing is fixed, capacity or attention will be elastic with respect to a change in fundamental uncertainty or a change in policy in the RB-RI model. Specifically, optimal attention is increasing with the degree of RB because agents with strong preference for RB are more sensitive to the risk they face and thus choose to devote more capacity to monitoring the state. We then compare the implications of RS and RB for consumption and savings when considering both control and filtering decisions of inattentive consumers. In the risk-sensitive permanent income model with imperfect-state-observation due to RI, the classical Kalman filter that extremizes the expected value of a certain quadratic objective function is still optimal. After solving the RB and RS models with filtering, we establish the observational equivalence (OE) conditions between RB and RS. We find that the simple and linear OE between RB and RS established in Hansen and Sargent (2007) and Luo and Young (2010) no longer holds, we instead have a complicated and nonlinear OE between RB and RS under RI.

We next explore how the interaction of RB and RI affects the consumption-income inequality. Using the estimated individual income process documented in the literature, we find that the relative dispersion of consumption to income obtained in the full-information RE-PIH model is well below its empirical counterpart, and RI by itself still cannot generate sufficiently high consumption-income dispersion. In contrast, we show that the interaction of RB and RI can generate the realistic consumption and income inequality for plausibly calibrated RB parameter values. In addition, we also find that allowing for optimal attention helps the model explain the evolution of the consumption and income inequality.

Finally, we investigate the asset pricing implications of RB and RI.[8] Following Hansen (1987) and HST (1999), we interpret the consumption-saving decisions in terms of a social planning problem and these decisions are equilibrium allocations for a competitive equilibrium. We can then deduce asset prices as in the consumption-based asset pricing literature by finding the shadow prices that clear security markets. Since these asset prices include information about the agent's intertemporal preferences, they measure the risk and uncertainty aversion of the agent. Given the explicit solutions for consumption and saving decisions, we can explicitly solve for the

[8] See Epstein and Wang (1994), Chen and Epstein (2002), and Ju and Miao (2012) for ambiguity, risk aversion, and asset returns. See Peng (2004), Luo (2008), Mondria (2010), and Van Nieuwerburgh and Veldkamp (2009, 2010) for applications of rational inattention in consumption, portfolio selections, and asset pricing.

market prices of induced uncertainty under RB and RI.[9] We find that the interaction of RB and RI significantly increases the market price of uncertainty, and thus makes the model better explain the market price of risk estimated from the data. The mechanism is straightforward to describe. Under RB, the market price of uncertainty is related to the norm of the worst-case shock (that is, the size of the pessimistic distortion to the underlying stochastic process for income); adding rational inattention increases the size of these distortions and therefore amplifies the effect on asset prices. We find that our model, under plausible calibrations of the fear of model misspecification based on detection error probabilities (as in Hansen and Sargent, 2007), produces stochastic discount factors that satisfy the Hansen–Jagannathan bounds.

Literature review This paper contributes to the literature on consumption-saving dynamics and asset pricing with incomplete information. This paper is closely related to HST (1999), Hansen et al. (2002, henceforth HSW), Luo (2008), and Luo and Young (2010). HST (1999) explored how model uncertainty due to robustness affects consumption-saving decisions and asset prices within the LQG setting, and found that the interaction of RB with habit formation and adjustment costs can help generate sufficiently high market price of risk. HSW (2002) extended HST (1999) and considered a robust control and filtering problem when part of the state vector is unobservable. Luo (2008) studied how RI affects consumption dynamics and helps resolve two well-known consumption puzzles. Luo and Young (2010) discussed the key differences between risk-sensitivity, robustness, and the discount factor in determining consumption-saving decisions when consumers are inattentive. Unlike HST (1999), HSW (2002), and Luo and Young (2010), the present paper focuses on the rich interaction of model uncertainty due to robustness and state uncertainty due to inattention and shows that RI by itself creates an additional demand for model uncertainty. We then use the model to explore the dynamics of consumption and income as well as asset prices.[10]

The remainder of the paper is organized as follows. Section 2 presents a rational inattention version of the permanent income model. Section 3 discusses how to model robust control and filtering under rational inattention, and examines how the preference for robustness affects individual consumption and saving decisions. Section 4 explores how the interaction of the two informational frictions affects individual consumption and saving dynamics and its welfare implications. Section 5 computes how induced uncertainty due to RB and RI affects the market prices of risk. Section 6 concludes.

2. A rational inattention version of the standard permanent income model

In this section we consider a rational inattention (RI) version of the standard permanent income model. In the standard permanent income model (Hall, 1978; Flavin, 1981), households solve the dynamic consumption-savings problem

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

subject to

[9] To explore how induced uncertainty due to RB and RI affects market prices of uncertainty, we follow the procedure adopted in Epstein and Wang (1994) and Hansen et al. (1999).

[10] Luo (2016) considered a robustly strategic consumption-portfolio choice problem of inattentive investors who face labor income risk and have constant-absolute-risk-averse (CARA) utility in a continuous-time setting.

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

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

$$s_t = b_t + \frac{1}{R}\sum_{j=0}^{\infty} R^{-j} E_t\left[y_{t+j}\right] \tag{2}$$

is permanent income, i.e., the expected present value of lifetime resources, consisting of financial wealth (b_t) plus human wealth (i.e., the discounted expected present value of current and future labor income: $\sum_{j=0}^{\infty} R^{-j} E_t\left[y_{t+j}\right]/R$),

$$\zeta_{t+1} \equiv \frac{1}{R}\sum_{j=t+1}^{\infty}\left(\frac{1}{R}\right)^{j-(t+1)}(E_{t+1} - E_t)\left[y_j\right], \tag{3}$$

is the time $(t+1)$ innovation to permanent income, b_t is financial wealth (or cash-on-hand), y_t is a labor income process with Gaussian white noise innovations, β is the discount factor, and $R > 1$ is the constant gross interest rate at which the consumer can borrow and lend freely.[11] In this paper, we assume that income y_t takes the following AR(1) process with the persistence coefficient $\phi \in [0, 1]$,

$$y_{t+1} = \phi y_t + (1-\phi)\overline{y} + \varepsilon_{t+1}, \tag{4}$$

where $\overline{y}$ is the mean of income, and ε_{t+1} is iid with mean 0 and variance ω^2. Given this income specification, we have $s_t \equiv b_t + y_t/(R-\phi) + (1-\phi)\overline{y}/[(R-1)(R-\phi)]$ and $\zeta_{t+1} = \varepsilon_{t+1}/(R-\phi)$, where and $\omega_\zeta^2 \equiv \text{var}(\zeta_{t+1}) = \omega^2/(R-\phi)^2$.[12] Finally, financial wealth (b) follows the process

$$b_{t+1} = Rb_t + y_t - c_t. \tag{5}$$

This specification follows that in Hall (1978) and Flavin (1981) and implies that optimal consumption is determined by permanent income:

$$c_t = \left(R - \frac{1}{\beta R}\right)s_t - \frac{1}{R-1}\left(1 - \frac{1}{\beta R}\right)\overline{c}. \tag{6}$$

We assume for the remainder of this section that $\beta R = 1$, since this setting is the only one that implies zero drift in consumption under rational expectations. Under this assumption the model leads to the well-known random walk result of Hall (1978):

$$\Delta c_{t+1} = (R-1)\zeta_{t+1}; \tag{7}$$

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: uncertainty has no effect on consumption, so that there is no precautionary saving.

---

11 We only require that y_t and R are such that permanent income is finite.

12 For the rest of the paper we will restrict attention to points where $c_t < \overline{c}$, so that utility is increasing and concave.

2.1. A detour on consumption, the PIH, and robust control

To motivate what follows, we now remind readers why (7) is inadequate as an empirical representation of consumption. In the U.S. data aggregate consumption exhibits both "excessive smoothness" to unanticipated changes in income and "excessive sensitivity" to anticipated changes in income. An alternative but equivalent representation of these puzzles is to say that consumption changes too little in response to permanent changes in income and too much in response to temporary ones. Campbell and Deaton (1989) provided a detailed discussion on these consumption puzzles and how they are related to each other. Unfortunately for the benchmark PIH model, (7) implies that consumption should be orthogonal to anticipated income changes, and changes in consumption are too volatile relative to the change in income if income is difference-stationary (which cannot be rejected in U.S. data).[13] HST (1999) improved upon the basic model by introducing robustness, but need habit formation to avoid exacerbating the excess sensitivity problem (robust agents respond more strongly to changes in permanent income). Luo and Young (2010) explicitly showed that robustness by itself worsens the standard FI-RE 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. In Section 4, we will examine how robustness interacts with rational inattention and affects the relative volatility of consumption growth to income growth at the individual level.

With respect to asset prices, we can price financial assets by treating the consumption process, (7), as though it were an endowment process (we will be more specific on this point in Section 5). Asset prices are therefore just the shadow prices that leave the consumer content with that endowment process; for the benchmark model the equity premium puzzle is in full force. As shown in HST (1999), a robust permanent income model can generate sufficiently high market price of risk, but as noted already they require strong habit formation to get reasonable asset prices.

In contrast, we will use rational inattention instead of habit formation. Luo (2008) and Luo et al. (2015) (henceforth LNWY) discussed the key differences between habit and rational inattention in partial equilibrium and general equilibrium models, respectively; while both lead to slow adjustment in consumption, they have different implications for consumption volatility and equilibrium interest rates.

2.2. Information-processing constraints

To this end we follow Sims (2003, 2010) and incorporate rational inattention (RI) due to finite information-processing capacity into the model. Under RI, consumers have only finite Shannon channel capacity to observe the state of the world. Specifically, we use the concept of entropy from information theory to characterize the uncertainty about a random variable; the reduction in entropy is thus a natural measure of information flow.[14] With finite capacity $\kappa \in (0, \infty)$, a variable s following a continuous distribution cannot be observed without error and thus the information set at time $t+1$, $\mathcal{I}_{t+1}$, is generated by the entire history of noisy signals $\left\{s_j^*\right\}_{j=0}^{t+1}$.

[13] Luo (2008) discussed how rational inattention can help resolve the two consumption puzzles.

[14] Formally, entropy is defined as the expectation of the negative of the (natural) log of the density function, $-E\left[\ln(f(s))\right]$. The entropy of a discrete distribution with equal weight on two points is simply $E\left[\ln(f(s))\right] = -0.5\ln(0.5) - 0.5\ln(0.5) = 0.69$, and the unit of information contained in this distribution is 0.69 "nats". In this case, an agent can remove all uncertainty about s if the capacity devoted to monitoring s is $\kappa = 0.69$ nats.

Agents with finite capacity will choose a new signal $s_{t+1}^* \in \mathcal{I}_{t+1} = \left\{s_1^*, s_2^*, \cdots, s_{t+1}^*\right\}$ that reduces their uncertainty about the state variable s_{t+1} as much as possible. Formally, this idea can be described by the information constraint

$$\mathcal{H}\left(s_{t+1}|\mathcal{I}_t\right) - \mathcal{H}\left(s_{t+1}|\mathcal{I}_{t+1}\right) \leq \kappa, \tag{8}$$

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

Under the linear-quadratic-Gaussian (LQG) setting, as has been shown in Sims (2003, 2010), ex post Gaussian distribution, $s_t|\mathcal{I}_t \sim N\left(E\left[s_t|\mathcal{I}_t\right], \Sigma_t\right)$, where $\Sigma_t = E_t\left[\left(s_t - \widehat{s}_t\right)^2\right]$, is optimal.[15] In addition, Maćkowiak and Bartosz (2009) also show that when the variables being tracked are stationary Gaussian process, signals which take the form of "true state plus white noise error" (i.e., $s_{t+1}^* = s_{t+1} + \xi_{t+1}$, where ξ_{t+1} is the iid endogenous noise due to RI) are optimal.[16] In the next section, after taking RB into account, we will show that the ex post Gaussianity and additive iid Gaussian noise are still optimal under RI-RB. As will been shown in the next subsection, although introducing RB can significantly change the RI model's dynamics and welfare implications, it does not change the key properties of the ex post distribution of the state and the RI-induced noise. The logic for modeling RI and RB jointly this way is that we first conjecture that RB does not change the optimality of ex post Gaussianity, and then verify that the conjecture is correct after incorporating RB into the RI model. (Note that when we introduce RB, we treat the RI model as the approximating model.)

Specifically, within this robust LQG setting, the information-processing constraint, (8), can be reduced to

$$\log\left(R^2\Sigma_t + \omega_\zeta^2\right) - \log\left(\Sigma_{t+1}\right) \leq 2\kappa. \tag{9}$$

Since this constraint is always binding, we can compute the value of the steady state conditional variance Σ: $\Sigma = \omega_\zeta^2/\left(\exp\left(2\kappa\right) - R^2\right)$. Given this Σ, we can use the usual formula for updating the conditional variance of a Gaussian distribution Σ to recover the variance of the endogenous noise (Λ):

$$\Lambda = \left(\Sigma^{-1} - \Psi^{-1}\right)^{-1}, \tag{10}$$

where $\Psi = R^2\Sigma + \omega_\zeta^2$ is the posterior variance of the state. Finally, $\widehat{s}_t$ is governed by the following Kalman filtering equation:

$$\widehat{s}_{t+1} = (1-\theta)\left(R\widehat{s}_t - c_t\right) + \theta\left(s_{t+1} + \xi_{t+1}\right), \tag{11}$$

given $s_0 \sim N\left(\widehat{s}_0, \Sigma\right)$, where $\theta = \Sigma\Lambda^{-1} = 1 - \exp\left(-2\kappa\right)$ is the Kalman gain. In the next section, after introducing RB into the RI model, we will show that κ and θ can be endogenously determined in the RI-RB model by assuming that the marginal cost of information processing

15 Shafieepoorfard and Raginsky (2013) derived the result formally, as opposed to the heuristic approach from Sims (2003).

16 This result is often assumed as a matter of convenience in signal extraction models with exogenous noises, and RI can rationalize this assumption.

(i.e., the shadow price of information-processing ability) is constant. We will also show that in the robust LQG case these two RI modeling strategies are observationally equivalent in the sense that they lead to the same conditional variance and the Kalman gain. Under the observational equivalence, we can construct a mapping between fixed information-processing cost and fixed channel capacity.

Note that after substituting (1) into (11), we have an alternative expression of the regular Kalman filter:

$$\widehat{s}_{t+1} = R\widehat{s}_t - c_t + \eta_{t+1}, \tag{12}$$

where

$$\eta_{t+1} = \theta R\left(s_t - \widehat{s}_t\right) + \theta\left(\zeta_{t+1} + \xi_{t+1}\right) \tag{13}$$

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

$$s_t - \widehat{s}_t = \frac{(1-\theta)\,\zeta_t}{1-(1-\theta)R\cdot L} - \frac{\theta\xi_t}{1-(1-\theta)R\cdot L}, \tag{14}$$

and $E_t\left[\eta_{t+1}\right] = 0$ because the expectation is conditional on the perceived signals and inattentive agents cannot perceive the lagged shocks perfectly.[17] The variance of the innovation to the perceived state is:

$$\omega_\eta^2 = \mathrm{var}\left(\eta_{t+1}\right) = \frac{\theta}{1-(1-\theta)\,R^2}\omega_\zeta^2, \tag{15}$$

which means that ω_η^2 reflects two sources of uncertainty facing the consumer: (i) fundamental uncertainty, ω_ζ^2 and (ii) induced uncertainty, i.e., state uncertainty due to RI, $\left[\frac{\theta}{1-(1-\theta)R^2} - 1\right]\omega_\zeta^2$. Therefore, as κ decreases, the relative importance of induced uncertainty to fundamental uncertainty increases.

In the next section, we will discuss alternative ways to robustify this RI-PIH model and their different implications for consumption, precautionary savings, and the welfare costs of uncertainty. The RB-RI model proposed here encompasses the hidden state model discussed in HSW (2002) and Hansen and Sargent (2005); the main difference is that agents in the RB-RI model cannot observe the entire state vector perfectly, whereas agents in the RB-hidden state model can observe some part of the state vector (in particular, the part they control).

3. Robust control and filtering under rational inattention

3.1. *Concerns about the fundamental shock and the noise shock*

As shown in Hansen and Sargent (2007), we can robustify the permanent income model by assuming agents with finite capacity distrust their model of the data-generating process (i.e., their income process), but still use an ordinary Kalman filter to estimate the true state. Note that without the concern for model misspecification, the consumer has no doubts about the probability model used to form the conditional expectation of permanent income (s). It is clear that the

[17] In order that the variance of η be finite we need $\kappa > \ln(R) \approx R - 1$. For short time periods this requirement is obviously not very restrictive. Since $R > 1$, some minimum level of capacity is needed to control the conditional mean of permanent income and enforce the transversality condition.

Kalman filter under RI, (12), is not only affected by the fundamental shock (ζ_{t+1}), but also affected by the endogenous noise (ξ_{t+1}) induced by finite capacity; these noise shocks could be another source of the demand for robustness. We therefore need to consider this demand for robustness in the RB-RI model. By adding the additional concern for robustness developed here, we are able to strengthen the effects of robustness on decisions.[18] Specifically, we assume that the agent thinks that (12) is the approximating model.

A simple version of robust optimal control considers the question of how to make decisions when the agent does not know the probability model that generates the data. Specifically, an agent with a preference for robustness considers a range of models surrounding the given approximating model, (12):

$$\widehat{s}_{t+1} = R\widehat{s}_t - c_t + \omega_\eta w_t + \eta_{t+1}. \tag{16}$$

where w_t distorts the mean of the innovation, and makes decisions that maximize lifetime expected utility given this worst possible model (i.e., the distorted model).[19] To make that model (12) is a good approximation when (16) generates the data, we constrain the approximation errors by an upper bound ψ_0:

$$E_0\left[\sum_{t=0}^{\infty}\beta^{t+1}w_t^2\right] \leq \psi_0, \tag{17}$$

where $E_0[\cdot]$ denotes conditional expectations evaluated with model, and the left side of this inequality is a statistical measure of the discrepancy between the distorted and approximating models. Note that the standard full-information RE case corresponds to $\psi_0 = 0$. In the general case in which $\psi_0 > 0$, the evil agent is given an intertemporal entropy budget $\psi_0 > 0$ which defines the set of models that the agent is considering. Following Hansen and Sargent (2007), we compute robust decision rules by solving the following two-player zero-sum game: a minimizing decision maker chooses the optimal consumption process $\{c_t\}$ and a maximizing evil agent chooses the model distortion process $\{w_t\}$.

Following Hansen and Sargent (2007), a simple robustness version of the PIH model proposed above can be written as

$$v(\widehat{s}_t) = \max_{c_t}\min_{w_t}\left\{-\frac{1}{2}(\overline{c} - c_t)^2 + \beta\left(\frac{1}{2}\vartheta w_t^2 + E_t\left[v(\widehat{s}_{t+1})\right]\right)\right\} \tag{18}$$

subject to the distorted transition equation (i.e., the worst-case model), (16), where $\vartheta > 0$ is the Lagrange multiplier on the constraint specified in (17) and controls how bad the error can be. (18) is a standard dynamic programming problem and can be easily solved using the standard procedure.[20]

The following proposition summarizes the solution to the RB-RI model, under a mild parameter restriction akin to the "breakdown condition" from Hansen and Sargent (2007).

Proposition 1. *Suppose $R\omega_\eta^2 < \vartheta$. Given ϑ and κ, the consumption function under RB and RI is*

[18] Luo et al. (2012) used this approach to study the joint dynamics of consumption, income, and the current account in emerging and developed countries.

[19] Formally, this setup is a game between the decision-maker and a malevolent nature that chooses the distortion process w_t.

[20] There is a one-to-one correspondence between ψ_0 in (17) and ϑ in (18).

$$c_t = \frac{R-1}{1-\Pi}\widehat{s}_t - \frac{\Pi\overline{c}}{1-\Pi} \quad (19)$$

with $\Pi < 1$*, the mean of the worst-case shock is*

$$\omega_\eta w_t = \frac{(R-1)\Pi}{1-\Pi}\widehat{s}_t - \frac{\Pi\overline{c}}{1-\Pi}, \quad (20)$$

and $\widehat{s}_t$ *is governed by*

$$\widehat{s}_{t+1} = \rho_s\widehat{s}_t + \frac{\Pi\overline{c}}{1-\Pi} + \eta_{t+1} \quad (21)$$

under the approximation model, where $\rho_s = \frac{1-R\Pi}{1-\Pi} \in (0,1)$*,*

$$\Pi = \frac{R\omega_\eta^2}{\vartheta} \in (0,1), \quad (22)$$

η_{t+1} *and* ω_η^2 *are defined in (13) and (15), respectively, and* $\theta = 1 - 1/\exp(2\kappa)$.

Proof. See Appendix A.1. $\Pi < 1$ can be obtained because the second-order condition for the optimization problem is

$$\frac{R(R-1)}{2\left(1 - R\omega_\eta^2/\vartheta\right)} > 0, \text{ i.e., } \Pi < 1. \qquad \square$$

It is worth noting that (19) can also be obtained using multiplier preferences to represent a fear of model misspecification:

$$\widehat{v}(\widehat{s}_t) = \max_{c_t}\left\{-\frac{1}{2}(c_t - \overline{c})^2 + \beta \min_{m_{t+1}} E_t\left[m_{t+1}\widehat{v}(\widehat{s}_{t+1}) + \vartheta m_{t+1}\ln(m_{t+1})\right]\right\}, \quad (23)$$

where m_{t+1} is the likelihood ratio, $E_t\left[m_{t+1}\ln(m_{t+1})\right]$ is defined as the relative entropy of the distribution of the distorted model with respect to that of the approximating model, and $\vartheta > 0$ is the shadow price of capacity that can reduce the distance between the two distributions, i.e., the Lagrange multiplier on the constraint:

$$E_t\left[m_{t+1}\ln(m_{t+1})\right] \le \eta,$$

where $\eta \ge 0$ defines an entropy ball of the distribution of the distorted model with respect to that of the approximating model. Following the same procedure adopted in Hansen and Sargent (2007), we can also obtain the corresponding value function:

$$\widehat{v}(\widehat{s}_t) = \Omega\left(\widehat{s}_t - \frac{\overline{c}}{R-1}\right)^2 + \rho, \quad (24)$$

where $\Omega = -\frac{R(R-1)}{2(1-\Pi)}$ and $\rho = \frac{\vartheta}{2(R-1)}\ln\left(1 - \frac{(R-1)\Pi}{1-\Pi}\right)$. Although the two-player minmax game and multiplier preferences lead to the same consumption-saving decisions, they have different asset pricing implications. (See Section 5 for a detailed discussion.) Furthermore, given the quadratic value functions under RB, we can show that the loss function due to RI is also quadratic and consequently the optimality of the ex post Gaussianity of the state still holds in the RI-RB model. (See Appendix A.2 for a proof.)

Equations (19) and (22) determine the effects of model uncertainty due to RB and state uncertainty due to RI on the marginal propensity to consume out of perceived permanent income ($MPC_\eta \equiv \frac{R-1}{1-\Pi}$) and the constant precautionary saving premium ($PS \equiv \frac{\Pi \bar{c}}{1-\Pi}$). Since Π is increasing with the degrees of both RB (smaller ϑ) and RI (smaller κ and θ), it is straightforward to show that either RB or RI leads to more constant precautionary savings and higher marginal propensity to consume, holding other factors constant and given that $\Pi < 1$:

$$\frac{\partial \left(MPC_\eta\right)}{\partial \vartheta} < 0 \text{ and } \frac{\partial \left(PS\right)}{\partial \vartheta} < 0.$$

We now present the intuition about the effects of robustness (ϑ) on precautionary savings. 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. 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 θ gets smaller.

RB and RI affect consumption and precautionary savings through distinct channels. RI affects Π by increasing the variance of the innovation to the perceived state, ω_η^2, whereas RB affects Π via changing the structure of the response of consumption to income shocks. Furthermore, if we consider the marginal propensity to consume out of true permanent income,

$$MPC_\zeta \equiv \frac{R-1}{1 - R\theta / \left[\vartheta \left(1 - (1-\theta) R^2\right)\right] \omega_\zeta^2} \theta, \tag{25}$$

we can immediately see that

$$\frac{\partial \left(MPC_\zeta\right)}{\partial \vartheta} < 0, \frac{\partial \left(MPC_\zeta\right)}{\partial \theta} > 0.$$

That is, both an increase in the demand for robustness and an increase in inattention increases the marginal propensity to consume out of true (but unobserved) permanent income.

To examine the relative importance of the two informational frictions in determining the consumption function and precautionary savings, we compare the effects from proportionate shifts in ϑ governing RB and κ governing RI. Specifically, the marginal effects on Π from an increase in ϑ and κ are given by

$$\frac{\partial \Pi}{\partial \kappa} = \frac{R\left(1 - R^2\right)\exp\left(-2\kappa\right)}{\vartheta\left[1 - \exp\left(-2\kappa\right) R^2\right]^2} \omega_\zeta^2 \text{ and } \frac{\partial \Pi}{\partial \vartheta} = -\frac{R\omega_\eta^2}{\vartheta^2},$$

respectively. Therefore, the marginal rate of transformation between proportionate changes in ϑ and changes in κ can be written as

$$\text{MRT} = -\frac{\partial \Pi / \partial \kappa}{\left(\partial \Pi / \partial \vartheta\right) \vartheta} = \frac{2\left(R^2 - 1\right)\exp\left(-2\kappa\right)}{\left(1 - \exp\left(-2\kappa\right)\right)\left(1 - \exp\left(-2\kappa\right) R^2\right)} > 0. \tag{26}$$

This expression gives the proportionate reduction in ϑ (i.e., a stronger preference for RB) that compensates, at the margin, for a decrease in κ (i.e., more inattentive) – in the sense of preserving the same effect on the consumption function for a given $\widehat{s}_t$. Equation (26) shows that this compensating change depends on the interest rate (R) and the degree of inattention (κ). Fig. 1 clearly shows that MRT is decreasing with κ for any given R. Since $\partial\left(\text{MRT}\right)/\partial\kappa < 0$, consumers with

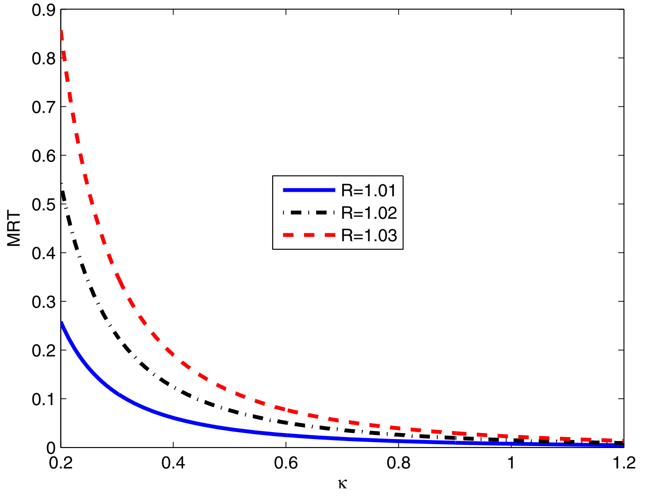

Fig. 1. MRT between RB and RI.

lower capacity will ask for higher compensation in an proportionate increase in model uncertainty facing them for an increase in capacity. For example, when $R = 1.03$, $\mathrm{MRT} = 0.256$ when $\kappa = 0.5$ bits, while $\mathrm{MRT} = 0.054$ when $\kappa = 1$ bit. In other words, to maintain the same effect on the consumption function, a decrease in κ by 50 percent (from 1 bit to 0.5 bits) matches up approximately with a proportional decline in ϑ of 2.7 percent. We will show later that there is a model-independent procedure for estimating ϑ; the trade-off here could in principle be used to discipline the choice for κ.[21]

It is also instructive to examine exactly what agents "fear" – that is, what are the dynamics of total resources under the worst-case model? Substituting (19) and (20) into (16) yields the law of motion for $\widehat{s}_t$ under the worst-case model:

$$\widehat{s}_{t+1} = \widehat{s}_t + \eta_{t+1} = (1 - \theta R)\widehat{s}_t + \theta R s_t + \theta\left(\zeta_{t+1} + \xi_{t+1}\right) \tag{27}$$

as compared to the actual process

$$\widehat{s}_{t+1} = \left(\frac{1 - R^2\omega_\eta^2/\vartheta}{1 - R\omega_\eta^2/\vartheta} - \theta R\right)\widehat{s}_t + \frac{\left(R\omega_\eta^2/\vartheta\right)\overline{c}}{1 - R\omega_\eta^2/\vartheta} + \theta R s_t + \theta\left(\zeta_{t+1} + \xi_{t+1}\right). \tag{28}$$

The key difference between the two processes is the autocorrelation parameter; since $\frac{1-R^2\omega_\eta^2/\vartheta}{1-R\omega_\eta^2/\vartheta} < 1$, the worst case model is more persistent than the true process. As noted in Kasa (2006), the most destructive distortions are low-frequency ones, so naturally the agents in the model design their decision rules to be robust against precisely those kinds of processes. ϑ does not appear in (27), as it only determines the size of the distortion process $\{w_t\}$ needed to achieve the worst-case model.[22]

[21] κ (or θ) is difficult to estimate outside the model; the literature on processing information provides estimates of the total ability of humans, but little guidance on how much of that ability would be dedicated to monitoring economic data. Obviously it would not be feasible to model all the competing demands for attention.

[22] If $\theta = 1$ (so that $\widehat{s}_t = s_t$) then the worst-case model is a random walk.

3.2. Robust Kalman filter gain

Another source of robustness could arise from the Kalman filter gain. In Section 3.1, we assumed that the agent distrusts the innovation to the perceived state but trusts the regular Kalman filter gain. Following Hansen and Sargent (2005; 2007, Chapter 17), in this section we consider a situation in which the agent pursues a robust Kalman gain and faces the commitment on the part of the minimizing agent to previous distortions.[23] Specifically, assume that at t the agent observes the noisy signal

$$s_t^* = s_t + \xi_t, \tag{29}$$

where s_t is the true state and ξ_t is the iid endogenous noise. The variance of the noise term, $\Lambda \equiv \text{var}(\xi_t)$, is $\left(\omega_\zeta^2 + R^2\Sigma\right)\Sigma/\left[\omega_\zeta^2 + \left(R^2-1\right)\Sigma\right]$, and $\Sigma = \omega_\zeta^2/\left(\exp(2\kappa) - R^2\right)$ is the steady state conditional variance. Given the budget constraint,

$$s_{t+1} = Rs_t - c_t + \zeta_{t+1}, \tag{30}$$

we consider the following time-invariant robust Kalman filter equation,

$$\widehat{s}_{t+1} = (1-\theta)(R\widehat{s}_t - c_t) + \theta(s_{t+1} + \xi_{t+1}), \tag{31}$$

where $\widehat{s}_{t+1}$ is the estimate of the state using the history of the noisy signals, $\left\{s_j^*\right\}_{j=0}^{t+1}$. We want θ to be robust to unstructured misspecifications of Equations (29) and (30). To obtain a robust Kalman filter gain, the agent considers the following distorted model:

$$s_{t+1} = Rs_t - c_t + \zeta_{t+1} + \omega_\zeta \nu_{1,t+1}, \tag{32}$$

$$s_{t+1}^* = s_{t+1} + \xi_{t+1} + \varrho\nu_{2,t+1}, \tag{33}$$

where $\varrho = \sqrt{\Lambda}$ and $\nu_{1,t+1}$ and $\nu_{2,t+1}$ are distortions to the conditional means of the two shocks, ζ_{t+1} and ξ_{t+1}, respectively.

Combining (30), (31), (32) with (33) gives the following dynamic equation for the estimation error:

$$e_{t+1} = (1-\theta)Re_t + (1-\theta)\zeta_{t+1} - \theta\xi_{t+1} + (1-\theta)\omega_\zeta\nu_{1,t+1} - \theta\varrho\nu_{2,t+1}, \tag{34}$$

where $e_t = s_t - \widehat{s}_t$.[24] We can then solve for the robust Kalman filter gain corresponding to this problem by solving the following deterministic optimal linear regulator problem:

$$e_0^T P e_0 = \max_{\{\nu_{t+1}\}} \sum_{t=0}^{\infty} \left(e_t^T e_t - \vartheta \nu_{t+1}^T \nu_{t+1}\right), \tag{35}$$

subject to

$$e_{t+1} = (1-\theta)Re_t + D\nu_{t+1}, \tag{36}$$

[23] Hansen and Sargent (2007) also discussed robust filtering without commitment. It is still debatable that which approach, with commitment or without commitment, is more appealing and tractable for modeling robust filtering under RI. The former applies the separation principle and can thus allow us to solve the robust control and filtering problems separately using a two-stage procedure, while the latter solves the robust estimation problem implied by solving the lifetime utility maximization problem. In this paper, for tractability we only consider the robust filtering problem with commitment and leave the problem without commitment for future research.

[24] Note that control variable, c, does not affect the estimation error equation.

where $D = \left[\ (1-\theta)\,\omega_\zeta \ \ -\theta\varrho\ \right]$ and $\nu_{t+1} = \left[\ \nu_{1,t+1} \ \ \nu_{2,t+1}\ \right]^T$. We can compute the worst-case shock by solving the corresponding Bellman equation and obtain

$$\nu_{t+1}^{*} = Qe_t, \tag{37}$$

where I is the identity matrix, P is the value function matrix, and $Q = \left(\vartheta I - D^T P D\right)^{-1} \times D^T P\,(1-\theta)\,R$. Note that here Q depends on robustness (ϑ) and channel capacity (κ).

For arbitrary Kalman filter gain θ, using (37), (34) can be written as

$$e_{t+1} = \left\{(1-\theta)\,R + \left[(1-\theta)\,\omega_\zeta - \theta\varrho\right] Q\right\} e_t + (1-\theta)\,\zeta_{t+1} - \theta\xi_{t+1}. \tag{38}$$

Taking unconditional mean on both sides of (38) gives

$$\Sigma_{t+1} = \left\{(1-\theta)\,R + \left[(1-\theta)\,\omega_\zeta - \theta\varrho\right] Q\right\} \Sigma_t + (1-\theta)^2\,\omega_\zeta^2 + \theta^2\omega_\xi^2, \tag{39}$$

where $\Sigma_{t+1} = E\left[e_{t+1}^2\right]$. From (39), it follows directly that in the steady state

$$\Sigma\,(\theta;\,Q) = \frac{(1-\theta)^2\,\omega_\zeta^2 + \theta^2\omega_\xi^2}{1-\chi^2},$$

where $\chi = (1-\theta)\,R + \left[(1-\theta)\,\omega_\zeta - \theta\varrho\right] Q$, and the robust Kalman filter gain $\theta\,(\vartheta,\kappa)$ minimizes the variance of e_t, $\Sigma\,(\theta;\,Q)$:

$$\theta\,(\vartheta,\kappa) = \arg\min\left(\Sigma\,(\theta;\,Q\,(\vartheta,\kappa))\right). \tag{40}$$

The upper panel of Fig. 2 illustrates how robustness (measured by ϑ) and inattention (measured by κ) affect the robust Kalman gain when $R = 1.02$ and $\omega_\zeta^2 = 4.29$.[25] It clearly shows that holding the degree of attention (i.e., channel capacity κ) fixed, increasing robustness (reducing ϑ) increases the Kalman gain (θ). In addition, for given robustness (ϑ), the Kalman gain is increasing with capacity. For example, when $\log(\vartheta) = 3$, the robust Kalman gain will increase from 60.17 percent to 77.35 percent when capacity κ increases from 0.6 bits to 1 bit; when $\kappa = 0.6$ bits, the robust Kalman gain will increase from 58.31 percent to 60.17 percent if ϑ falls from $\log(\vartheta) = 4$ to 3.[26]

After obtaining the robust Kalman gain $\theta\,(\vartheta,\kappa)$, we can solve the Bellman equation proposed in Section 3.1 using the Kalman filtering equation with robust θ. The following proposition summarizes the solution to this problem:

Proposition 2. *Given ϑ and κ, the consumption function is*

$$c_t = \frac{R-1}{1-\Pi}\widehat{s}_t - \frac{\Pi\overline{c}}{1-\Pi}, \tag{41}$$

where $\Pi = \frac{R\omega_\eta^2}{\vartheta} \in (0,1)$,

$$\omega_\eta^2 = \mathrm{var}\,(\eta_{t+1}) = \frac{\theta\,(\vartheta,\kappa)}{1-(1-\theta\,(\vartheta,\kappa))\,R^2}\omega_\zeta^2, \tag{42}$$

and $\widehat{s}_t$ *is governed by*

[25] See Section 4.1 for estimating the value of ω_ζ^2. We use the program `rfilter.m` provided in Hansen and Sargent (2007) to compute the robust Kalman filter gain $\theta\,(\vartheta,\kappa)$.

[26] This result is consistent with that obtained in a continuous-time filtering problem discussed in Kasa (2006).

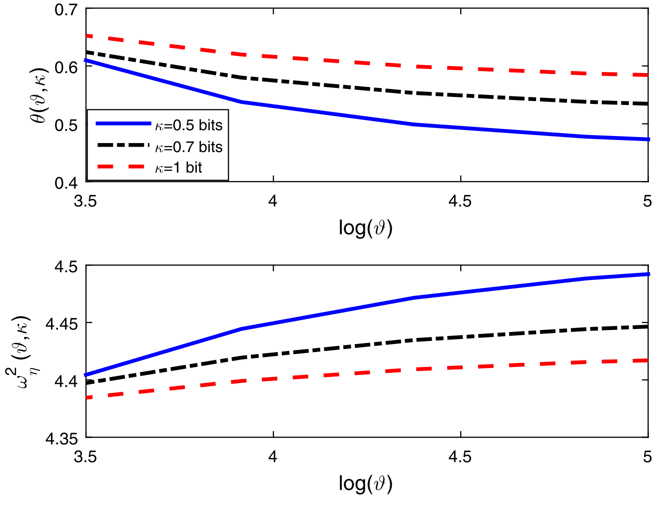

Fig. 2. Effects of RB and RI on robust Kalman gain θ and ω^2_η.

$$\widehat{s}_{t+1} = \rho_s \widehat{s}_t + \eta_{t+1}, \tag{43}$$

where $\rho_s = \frac{1-R\Pi}{1-\Pi} \in (0, 1)$.

Proof. The proof is the similar to that provided in Appendix A.1. Here we just need to replace $\theta(\kappa) = 1 - \exp(-2\kappa)$ with $\theta(\vartheta, \kappa)$. □

Note that here θ is a function of both ϑ (concerns about Kalman gain) and κ (channel capacity), rather than simply $1 - 1/\exp(2\kappa)$ as obtained in Section 3.1. In this case the agent is not only concerned about disturbances to the perceived permanent income, but also concerned about the Kalman gain. It is clear from (41) and (42) that the preference for robustness has opposing effects on both the marginal propensity to consume out of permanent income, i.e., the responsiveness of c_t to $\widehat{s}_t$ $\left(MPC_\eta = \frac{R-1}{1-\Pi}\right)$ and precautionary savings, i.e., the intercept of the consumption profile $\left(PS = \frac{\Pi\bar{c}}{1-\Pi}\right)$.[27] Specifically, if we temporarily shut down the concern about disturbances to perceived permanent income, we can see from (41) that the smaller the value of ϑ the lower the MPC and the smaller the precautionary saving increment

$$\frac{\partial\left(MPC_\eta\right)}{\partial\vartheta} > 0 \text{ and } \frac{\partial\left(PS\right)}{\partial\vartheta} > 0$$

because $\frac{\partial\omega^2_\eta}{\partial\vartheta} > 0$, $\frac{\partial\omega^2_\eta}{\partial\theta} < 0$, $\frac{\partial\omega^2_\eta}{\partial\kappa} < 0$, and $\frac{\partial\theta}{\partial\vartheta} < 0$. From (41), we can see that the precautionary savings increment in the RB-RI model is determined by the interaction of three factors: labor income uncertainty, preferences for robustness (RB), and finite information-processing capacity (RI). The lower panel of Fig. 2 also illustrates how robustness (ϑ) and channel capacity (κ) affect ω^2_η. We now provide some intuition about the effects of robustness (ϑ) on precautionary

[27] Note that given the consumption function Π has the same effect on the marginal propensity to consume and precautionary savings.

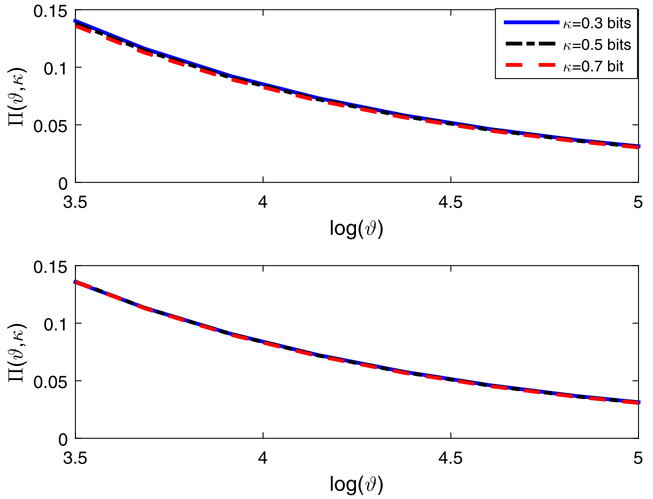

Fig. 3. Effects of RB and RI on Π.

savings in this case. An increase in robustness (a reduction in ϑ) will increase the Kalman gain θ, which leads to lower ω_η^2 and then low precautionary savings. We can see that under certain conditions a greater reaction to the shock can either be interpreted as an increased concern for robustness in the presence of model misspecification, or an increase in information-processing ability when agents only have finite channel capacity.

It is clear that both Type I and Type II misspecification affect Π. We now evaluate their relative importance in determining Π. The upper and lower panels of Fig. 3 illustrate Π as functions of ϑ for different values of κ when only Type I misspecification and both types of misspecification are considered, respectively. It is clear from this figure that Type I misspecification significantly dominates Type II misspecification via its direct impact on the value of Π; consequently, Type II misspecification only has very tiny impact on affecting Π and then the model's dynamics. (The upper and lower panels of Fig. 3 are indistinguishable.) To keep the model more tractable, we only consider Type I misspecification when we examine the implications of induced uncertainty on the consumption-income dynamics and the market price of uncertainty.

3.3. Comparison with risk-sensitive control and filtering

Risk-sensitivity (RS) was first introduced into the LQ-Gaussian framework by Jacobson (1973) and extended by Whittle (1990). Exploiting the recursive utility framework of Epstein and Zin (1989), Hansen and Sargent (1995) introduced discounting into the RS specification and show that the resulting decision rules are time-invariant. In the RS model agents effectively compute expectations through a distorted lens, increasing their effective risk aversion by overweighting negative outcomes. The resulting decision rules depend explicitly on the variance of the shocks, producing precautionary savings, but the value functions are still quadratic functions of the states.[28] HST (1999) and Hansen and Sargent (2007) interpreted RS preferences in terms

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

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. In the corresponding risk-sensitive filtering LQ problem, the problem is that when the state cannot be observed perfectly, is the classical Kalman filter that minimizes the expected loss function still optimal? In our LQ-PIH model setting, we can easily see that the regular Kalman filter is still optimal given the quadratic forms of the utility function and the value function.[29] In this section we will explore how the RS filtering affects consumption dynamics and precautionary savings and show that the OE between RB and RS is no longer linear, but takes a more complicated non-linear form. The RI version of risk-sensitive control based on recursive preferences with an exponential certainty equivalence function can be formulated as

$$\widehat{v}\left(\widehat{s}_t\right) = \max_{c_t}\left\{-\frac{1}{2}\left(c_t - \overline{c}\right)^2 + \beta\mathcal{R}_t\left[\widehat{v}\left(\widehat{s}_{t+1}\right)\right]\right\} \tag{44}$$

subject to the Kalman filter equation (12).[30] The distorted expectation operator is now given by

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

where $s_0 | \mathcal{I}_0 \sim N\left(\widehat{s}_0, \overline{\sigma}^2\right)$. It is worth noting that given that the value function in the RS model is quadratic, the regular Kalman filter is still optimal because the objective function in the filtering problem is the square of the estimation error.

Following the same procedure used in Hansen and Sargent (1995) and Luo and Young (2010), we can solve this risk-sensitive control problem explicitly. The following proposition summarizes the solution to the RI-RS model when $\beta R = 1$:

Proposition 3. *Given finite channel capacity κ and the degree of risk-sensitivity α, the consumption function of a risk-sensitive consumer under RI is*

$$c_t = \frac{R-1}{1-\Pi}\widehat{s}_t - \frac{\Pi\overline{c}}{1-\Pi}, \tag{45}$$

where

$$\Pi = R\alpha\omega_\eta^2 \in (0,1), \tag{46}$$

ω_η^2 is defined in (15).

Proof. The proof is straightforward and is similar to that in HST (1999) and Luo and Young (2010). □

---

29 As shown in Moore et al. (1997), even if the agent has risk-sensitive preferences when filtering,

$$\min \ln E_t\left\{\exp\left[-\vartheta\left(s_t - \widehat{s}_t^{RS}\right)^2\right]\right\},$$

the risk-sensitive estimate $\widehat{s}_t^{RS}$ is identical to the minimum variance estimate $\widehat{s}$ obtained from solving

$$\min E_t\left[\left(s_t - \widehat{s}_t\right)^2\right].$$

30 Given the quadratic form of the value function, introducing risk-sensitivity does not change the optimality of the *ex post* Gaussianity of the true state and the induced noise; see Luo and Young (2010) for more discussion.

Comparing (19) obtained from the model with only concerns about the innovation to the perceived state (i.e., without robust Kalman filtering) in Section 3.1 with (45), it is straightforward to show that RB and RS under RI are indistinguishable using only consumption-savings decisions if

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

Note that (47) is exactly the same as the observational equivalence condition obtained in the full-information RE model (see Backus et al., 2004). That is, under the assumption that the agent trusts the Kalman filter equation, the OE result obtained under full-information RE still holds under RI.[31]

HST (1999) show that as far as the quantity observations on consumption and savings are concerned, the robustness version ($\vartheta > 0$ or $\alpha > 0$, $\widetilde{\beta}$) of the PIH model is observationally equivalent to the standard version ($\vartheta = \infty$ or $\alpha = 0$, $\beta = 1/R$) of the PIH model for a unique pair of discount factors.[32] 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.[33] Alternatively, holding all parameters constant except the pair (α, β), the RI version of the PIH model with RB consumers ($\vartheta > 0$ and $\beta R = 1$) is observationally equivalent to the standard RI version of the model ($\vartheta_0 = \infty$ and $\widetilde{\beta} > 1/R$). To do so, we compare the consumption function obtained from the RI model ($\vartheta = \infty$ and $\widetilde{\beta} > 1/R$), $c_t = \left(R - \frac{1}{\beta R}\right)\widehat{s}_t - \frac{1}{R-1}\left(1 - \frac{1}{\beta R}\right)\overline{c}$, with (41) and (45), and find that when $\widetilde{\beta} = \frac{1}{R}\frac{1-R\omega_\eta^2/\vartheta}{1-R^2\omega_\eta^2/\vartheta} = \frac{1}{R}\frac{1-R\alpha\omega_\eta^2}{1-R^2\alpha\omega_\eta^2} > \frac{1}{R}$, consumption and savings are identical in the RI, RB-RI, and RS-RI models.

However, if we compare (41) obtained from the model with both concerns about the innovation to the perceived state and concerns about Kalman gain with (45), it is obvious that the observational equivalence between RB and RS under RI, (47), no longer holds. Given the same value of κ, the Kalman gain only depends on κ in the RS model, whereas it depends on both κ and ϑ (the preference for robust Kalman gain) in the RB model. The two Kalman gains are therefore different for any finite value of ϑ. If we allow for different values of κ, the models are observationally equivalent when $\alpha = \vartheta^{-1}$ and

$$\theta\left(\vartheta, \kappa_{RB}\right) = 1 - \frac{1}{\exp(2\kappa_{RS})}.$$

Fig. 4 illustrates how κ_{RS} varies with ϑ and κ_{RB} when the OE between RB and RS holds under RI. It clearly shows that given the level of ϑ, κ_{RS} is increasing with κ_{RB}.

3.4. Optimal attention under RB

As argued in Sims (2010), instead of using fixed channel capacity to model finite information-processing ability, one could assume that the marginal cost of information-processing (i.e., the

[31] The states are different, of course – s_t vs. $\widehat{s}_t$.

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

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

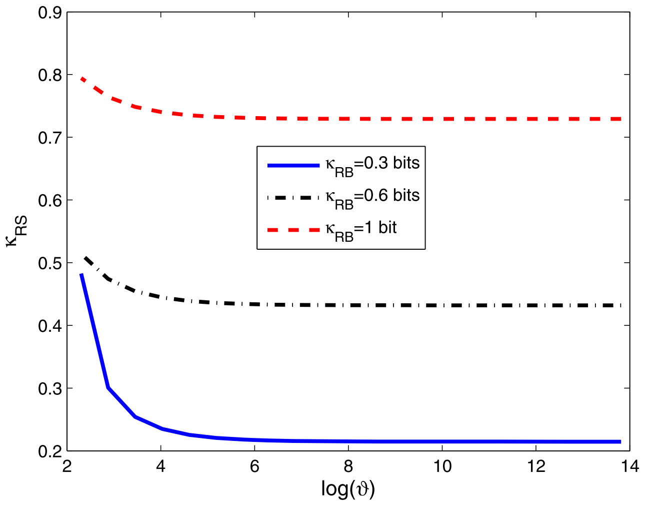

Fig. 4. The OE between RB and RS.

shadow price of information-processing capacity) is fixed. That is, the Lagrange multiplier on (9) is constant, and the resulting endogenous capacity that minimizing the expected loss function due to RI is elastic. In the univariate RB-RI model proposed above, we also assume that the agent faces a marginal cost of attention, $\lambda > 0$, and chooses optimal attention (κ) to monitor the true state s_t (or equivalently, the conditional variance, Σ_t) to minimize the following expected loss function:

$$\max_{\Sigma_t} \left\{ -\sum_{t=0}^{\infty} \beta^t \Omega_0 \Sigma_t \right\},$$

subject to (9), where $\Omega_0 = -\frac{R(R-1)}{2(1-\Pi_0)}$ and $\Pi_0 = \frac{R\omega_\zeta^2}{\vartheta}$. (See Appendix A.2 for the derivation of the loss function.) Solving this problem leads to the following optimal conditional variance of the true state:

$$\Sigma^* = \frac{-\left[R(R-1)\lambda + \Omega_0\omega_\zeta^2\right] - \sqrt{\left[R(R-1)\lambda + \Omega_0\omega_\zeta^2\right]^2 - 4\lambda\Omega_0 R^2\omega_\zeta^2}}{2\Omega_0 R^2}, \tag{48}$$

where λ is the Lagrange multiplier corresponding to the information-processing constraint, (9).[34]

Proposition 4. *The two RI modeling strategies are observationally equivalent in the sense that they lead to the same steady state conditional variance if the following equality holds:*

[34] It is straightforward to show that as λ goes to 0, $\Sigma^* = 0$; and as λ goes to ∞, $\Sigma^* = \infty$.

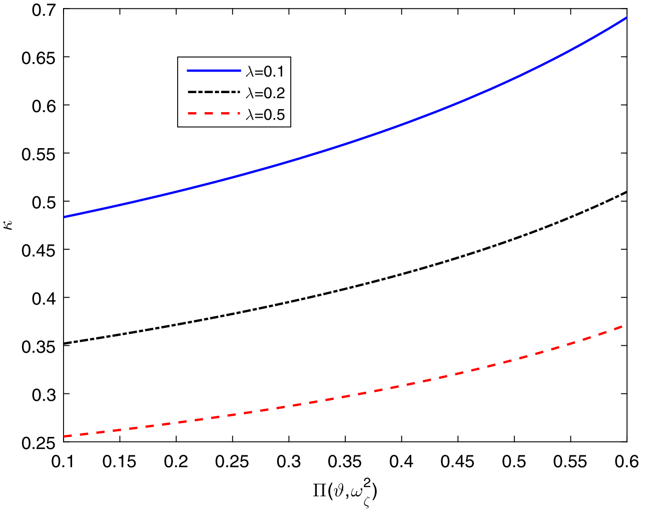

Fig. 5. Optimal attention under RB.

$$\kappa\left(\lambda,\omega_\zeta^2\right) = \ln R + \frac{1}{2}\ln\left(1 - \frac{2\Omega_0\omega_\zeta^2}{R(R-1)\lambda + \Omega_0\omega_\zeta^2 + \sqrt{\left[R(R-1)\lambda + \Omega_0\omega_\zeta^2\right]^2 - 4\lambda\Omega_0 R^2\omega_\zeta^2}}\right). \tag{49}$$

Proof. The proof is straightforward by comparing (48) with $\Sigma = \omega_\zeta^2/\left(\exp(2\kappa) - R^2\right)$ (the fixed-capacity case). □

It is obvious that κ converges to its lower limit $\underline{\kappa} \equiv \ln R \approx (R-1)$ as λ goes to ∞; and it converges to ∞ as λ goes to 0. In other words, using this RI modeling strategy, the consumer is allowed to adjust the optimal level of capacity in such a way that the marginal cost of information-processing for the problem at hand remains constant. Fig. 5 shows that optimal attention is increasing with the degree of RB and the amount of fundamental uncertainty for different values of λ. The reason for this result is that agents with strong preference for RB are more sensitive to the risk they face and thus choose to devote more capacity to monitoring the evolution of the state.

We use the following two-stage optimization procedure to solve the joint optimal control-filtering model under RB and RI:

1. Given that κ is constant channel capacity, guess that the ex post Gaussian distribution of the true state and additive iid Gaussian noise due to RI are still optimal when the agent has a preference for RB. Given the optimality of ex post Gaussianity and Gaussian noise, we can apply the standard Kalman filter and use the robust LQG dynamic programming to solve the RI-RB model explicitly. Using the loss function derived from the value functions under

RI-RB and RB, we can verify that our guess about the optimality of ex post Gaussianity and Gaussian noise is correct.
2. Minimizing the same loss function due to the information-processing constraint obtained in Stage 1 and fixed marginal cost leads to optimal conditional variance (Σ^*) and endogenous attention (κ), which verifies that the assumption of constant channel capacity we used in Stage 1 is correct.

Given this relationship between λ and κ, in the following analysis we just use the value of κ to measure the degree of attention. It is worth noting that although the above two RI modeling strategies, inelastic and elastic capacity, are observationally equivalent in the "static" sense, they have distinct implications for the model's propagation mechanism if the economy is experiencing regime switching (e.g., before and after the great moderation). With inelastic capacity, the propagation mechanism governed by the Kalman gain is fixed regardless of changes in fundamental uncertainty, while with elastic capacity the propagation mechanism will change in response to changes in fundamental uncertainty.

4. Dynamics of consumption, income, and wealth

The section studies the implications of induced uncertainty due to the interaction of RI and RB on the joint dynamics of consumption, income, and wealth. We will explore how incorporating RI and RB can help the model better fit the empirical evidence on the cross-sectional dispersions of consumption and income.

4.1. Empirical evidence

In order to estimate the household income process and measure the relative consumption dispersion in the data defined as the ration of the standard deviation of consumption growth to that of income growth, $\frac{\mathrm{sd}(\Delta c)}{\mathrm{sd}(\Delta y)}$, we need to have a panel data set which contains both consumption and income at the household level. Following Blundell et al. (2008) (henceforth BPP), LNWY (2015) constructed a panel of household income and consumption based on the Panel Study of Income Dynamics (PSID) and the Consumer Expenditure Survey (CEX), and estimated the household income process as specified by Equation (4). The estimated values of the persistence coefficient (ϕ) and the standard deviation of the income innovation (σ) are 0.88 and 0.29, respectively.[35] In addition, LNWY (2015) documented that the average empirical value of the relative volatility of consumption growth to income growth is 0.209, and the minimum and maximum values of the empirical relative volatility are 0.159 and 0.285, respectively, in the sample period, 1980–1996.

To explore how induced uncertainty due to RB and RI affects the market price of uncertainty, we adopt the calibration procedure outlined in HSW (2002), AHS (2003), and Hansen and Sargent (2007, Chapter 9) to calibrate the value of Π that summarizes the interaction between RB and RI. Specifically, we calibrate Π by using the notion of a model detection error probability (DEP) that is based on a statistical theory of model selection. (See Appendix A.3 for the detailed calibration process using DEP.) Fig. 6 illustrates how Π varies with the value of DEP (p) for different values of θ when $R = 1.02$ and $\overline{c} = 3y_0$. We can see from the figure that for a given value of θ, the stronger the preference for robustness (higher ϑ), the less the p is. For example,

[35] For the details of the constructed data set and the estimation of the income process, see LNWY (2015).

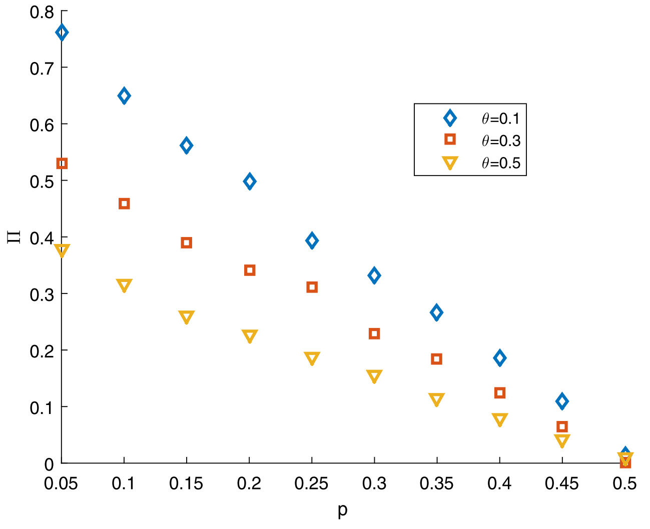

Fig. 6. Relationship between $\Pi\left(\vartheta, \omega_{\zeta}^{2}\right)$ and p.

when $\theta = 0.3$, $p = 40\%$ when $\log(\vartheta) = 3.5$, while $p = 15\%$ when $\log(\vartheta) = 2$. Both values of p are reasonable as argued in AHS (2003), HSW (2002), Maenhout (2004), and Hansen and Sargent (2007, Chapter 9).

4.2. *Sensitivity and volatility of consumption to income*

We will now discuss the effect of RI-RB on the joint dynamics of consumption and income. Combining (41) with (43) yields an expression for individual consumption in the RI-RB economy:

$$c_t = \frac{1 - R\Pi}{1 - \Pi} c_{t-1} + \frac{(R-1)\Pi}{1-\Pi}\overline{c} + \frac{R-1}{1-\Pi}\left[\frac{\theta\zeta_t}{1-(1-\theta)R\cdot L} + \theta\left(\xi_t - \frac{\theta R\xi_{t-1}}{1-(1-\theta)R\cdot L}\right)\right], \tag{50}$$

where L is the lag operator and we assume that $(1-\theta)R < 1$. This expression implies that consumption growth is a weighted average of all past permanent income and noise shocks. In addition, it is also clear from (50) that the propagation mechanism of the model is determined by the Kalman filter gain, θ (or equivalently κ). Fig. 7 illustrates that consumption in the RB-RI model reacts gradually to income shocks, with monotone adjustments to the corresponding RB asymptote when $\log(\vartheta) = 3.5$ and $\kappa = 0.6$. This case is illustrated by the dash-dotted line in Fig. 7. Similarly, the dotted line corresponds to the case in which $\log(\vartheta) = 5$ and $\kappa = 0.6$. With a stronger preference for robustness, the precautionary savings increment is larger and thus an income shock that is initially undetected would have larger effects on consumption during the adjustment process.[36]

[36] Estimating the process (50) on quarterly real nondurable and service consumption, logged and linearly detrended with $R = 1.01$, produces a nonstationary process (the autoregressive roots are 1.3038 and −0.3027, which delivers an

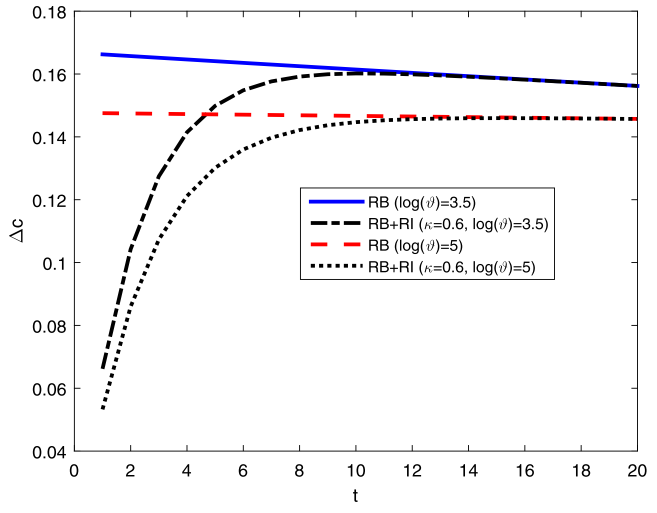

Fig. 7. Impulse responses of consumption to income shock.

Using (50), we can obtain the expression for the relative volatility of consumption growth relative to income growth.[37] The following proposition provides the expression of this relative volatility.

Proposition 5. *The relative volatility of consumption growth relative to income growth is*

$$\mu \equiv \frac{\mathrm{sd}\,(\Delta c_t)}{\mathrm{sd}\,(\Delta y_t)} = \frac{(R-1)\sqrt{1+\rho}}{R-\rho}\frac{1}{1-\Pi}\sqrt{\frac{1}{1+\rho_s}}\sqrt{\frac{\theta}{1-\rho_\theta R}}, \tag{51}$$

where we use the fact that $\omega_\xi^2 = \mathrm{var}\,(\xi_t) = \frac{1-\theta}{\theta\left[1-(1-\theta)R^2\right]}\omega_\zeta^2$, $\rho_s = \frac{1-R\Pi}{1-\Pi} \in (0,1)$, *and* $\rho_\theta = (1-\theta)\,R \in (0,1)$.

Proof. The proof is straightforward by taking unconditional variances on both sides of (50). □

In the standard full-information RE model where $\kappa = \infty$ and $\vartheta = \infty$, it is straightforward to show that $\mu = \frac{R-1}{R-\rho}\sqrt{\frac{1+\rho}{2}} = 0.139$, which is well below its empirical counterpart, 0.209, reported in LNWY (2015). In contrast, when we consider the RI model without RB, the relative volatility of consumption growth to income growth is increasing with the degree of inattention, i.e., $\partial\mu/\partial\theta < 0$. The main reason for this result is that rational inattention affects consumption volatility via two channels: (i) the gradual and smooth responses to income shocks (i.e., the $1-(1-\theta)R\cdot L$ term in (50)) and (ii) the RI-induced noises (ξ_t). Specifically, a reduction in capacity κ decreases the Kalman gain θ, which strengthens the smooth responses to income

eigenvalue just slightly larger than 1), so we do not pursue this direction further. We note in passing that the estimate for θ is close to 0 and the estimate for Π is close to 1. Lower values for R move these values closer to 0 and 1, so we cannot deliver stationarity easily by simply reducing R.

[37] Here we follow the consumption literature and use $\frac{\mathrm{sd}(\Delta c_t)}{\mathrm{sd}(\Delta y_t)}$ instead of $\frac{\mathrm{sd}(c_t)}{\mathrm{sd}(y_t)}$ to measure the relative volatility of the consumption process. Note when $\vartheta = \infty$ (no RB), $\mathrm{sd}\,(y_t)$ or $\mathrm{sd}\,(c_t)$ are not well defined.

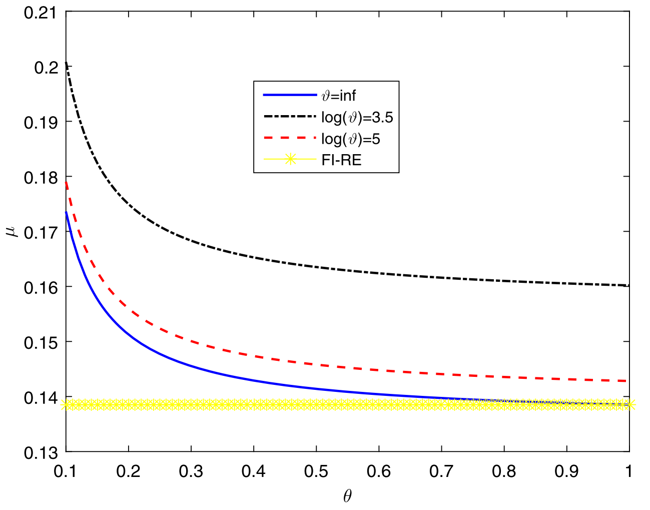

Fig. 8. Relative volatility of consumption to income under RB-RI.

shock and increases the volatility of the RI-induced noise. Luo (2008) showed that the noise effect dominates the smooth response effect, and the volatility of consumption growth decreases with κ. However, as illustrated in Fig. 8, in the presence of RB (i.e., $\vartheta = \infty$), the relative volatility of consumption to income is still significantly lower than the empirical counterpart even if channel capacity is very low. For example, $\mu = 0.146$ when $\theta = 30\%$, and $\mu = 0.151$ when $\theta = 20\%$.

Fig. 8 illustrates how RB (ϑ) increases the relative volatility of consumption growth to income growth for different values of κ. For any given value of κ, μ is decreasing with ϑ. That is, the stronger the preference for RB, the higher the relative volatility. For example, when $\log\vartheta = 3$ and $\theta = 20\%$, $\mu = 0.2$, which is consistent with its empirical counterpart.[38] The reason is that reducing ϑ increases Π. To explore the intuition behind this result, we consider the perfect-state-observation case in which $\kappa = \infty$. In this case, the relative volatility of consumption growth to income growth can be written as

$$\mu = \frac{(R-1)\sqrt{1+\rho}}{R-\rho}\frac{1}{1-\Pi}\sqrt{\frac{1}{1+\rho_s}}, \tag{52}$$

which clearly shows that ϑ increases the relative volatility via two channels. First, a higher ϑ increases the marginal propensity to consume out of permanent income $\left(\frac{R-1}{1-\Pi}\right)$, and second, it increases consumption volatility by reducing the persistence of permanent income measured by ρ_s: $\frac{\partial \rho_s}{\partial \Pi} < 0$. It is worth noting that although $\theta = 20\%$ is a very low number and is well below the total information-processing ability of human beings, it is not unreasonable in practice for ordinary consumers because they also face many other competing demands on capacity. For an extreme case, a young worker who accumulates balances in his retirement savings account (e.g., 401(k)) might pay no attention to the behavior of the stock market until he retires. In this case,

[38] The corresponding values of Π and the detection error probability (DEP) are 0.254 and 30%, respectively. See Appendix A.3 for a detailed discussion on the DEP calibration procedure.

they fail to recognize any fungibility between these assets and his preretirement consumption. In addition, in our model for simplicity we only consider one shock to total resources s, while in reality consumers face substantial idiosyncratic shocks that we do not model in this paper. Therefore, the exogenous capacity given in our model can be regarded as a shortcut to small fractions of consumers' total capacity used to monitor their total resources hit by the innovation to total resources.

Furthermore, if we allow the consumers to choose optimal attention (i.e., elastic capacity as in Sims, 2010), κ and θ are increasing functions of the fundamental uncertainty $\left(\omega_\zeta^2\right)$. In this case, for given ϑ and λ (the constant marginal information-processing cost), consumption will be smoother when income becomes more volatile because greater income uncertainty makes consumers devote more attention to monitoring the state. This theoretical result might provide a potential explanation for the empirical evidence documented in BPP (2008) and LNWY (2015) that income and consumption inequality diverged over the sampling period they study. To explore this issue in our model, we do the following exercise. First, following LNWY (2015), we divide the full sample into two sub-samples (1980–1987 and 1988–1996; or 1980–1988 and 1989–1996) and apply the same estimation procedure to re-estimate ω and ρ (see Table 3 of LNWY (2015) for the estimation results). Household income is more volatile in late sub-periods than earlier ones. For example, ω and ρ are 0.258 and 0.861 respectively in the sub-sample (1980–1987), while they are 0.290 and 0.838 in the sub-sample (1988–1996). The average values of μ are 0.24 and 0.19 in the first and second sub-samples, respectively. In the elastic capacity case, using the estimated income processes in the first sub-sample, we first use $\mu_{cy} = 0.24$ to calibrate that $\lambda = 1.2$. Using this calibrated value of λ, we find that $\mu = 0.2$, which matches the empirical counterpart almost perfectly.[39]

4.3. Implications for wealth dynamics

Combining the original budget constraint, $b_{t+1} = Rb_t + y_t - c_t$, with the consumption function (41), we can obtain the following expression for individual saving d_t:

$$d_t \equiv b_{t+1} - b_t = \frac{-\Pi (R-1)}{1-\Pi}\left(b_t - \overline{b}\right) + \left(1 - \frac{R-1}{(1-\Pi)(R-\rho)}\right)(y_t - \overline{y}) + \varsigma_{t+1}, \quad (53)$$

where $\varsigma_{t+1} = \frac{R-1}{1-\Pi}(s_t - \widehat{s}_t)$ is determined by the estimation error, $s_t - \widehat{s}_t = \frac{(1-\theta)\zeta_t}{1-(1-\theta)R\cdot L} - \frac{\theta\xi_t}{1-(1-\theta)R\cdot L}$ and $\overline{b} = \frac{\overline{c}-\overline{y}}{R-1}$ is the steady state value of b_t. From (53), it is straightforward to show that the unconditional mean of individual saving is zero. That is, induced uncertainty due to the interaction of RB and RI does not affect the amount of individual saving on average.[40]

Furthermore, using (53), we can compute the relative volatility of individual savings to income. The following proposition provides the expression of this ratio.

Proposition 6. *The relative volatility of individual savings is*

$$\mu_d \equiv \frac{\mathrm{sd}(\Delta b)}{\mathrm{sd}(\Delta y)} = \sqrt{\frac{1+\rho}{2(R-\rho)^2}}\sqrt{\begin{array}{c}\frac{1-\rho}{1+\rho} + \frac{\Gamma^2}{1-\rho_s^2} + \left(\frac{R-1}{1-\Pi}\right)^2\frac{1-\theta}{1-R^2(1-\theta)} + \frac{2(1-\rho)\Gamma}{1-\rho\rho_s} \\ + \left(\frac{R-1}{1-\Pi}\right)\frac{2(1-\rho)(1-\theta)}{1-\rho\rho_\theta} + \left(\frac{R-1}{1-\Pi}\right)\frac{2\Gamma(1-\theta)}{1-\rho_s\rho_\theta}\end{array}}, \quad (54)$$

39 Choosing other sub-samples (1980–1988 and 1989–1996) does not change the main result.

40 LNWY (2015) argued that this result continues to hold in general equilibrium.

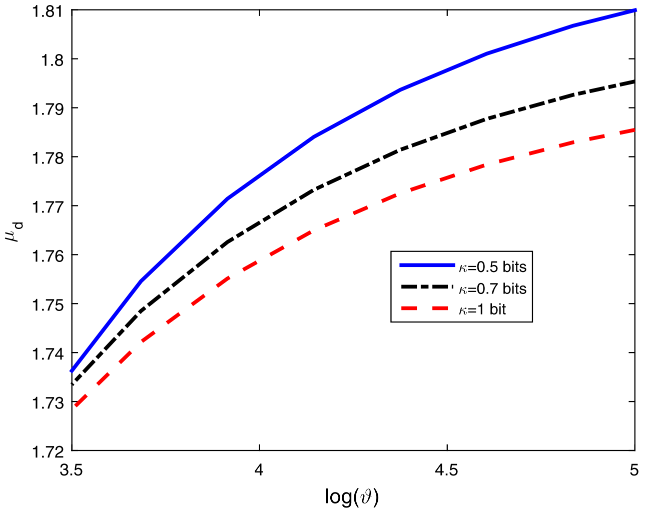

Fig. 9. Relative volatility of individual saving.

where $\Gamma = -\frac{(R-1)\Pi}{1-\Pi} < 0$, $\rho_s = \frac{1-R\Pi}{1-\Pi}$, *and* $\Pi = \frac{\theta}{1-(1-\theta)R^2}\frac{R\omega_\zeta^2}{\vartheta}$.

Proof. The proof is straightforward by taking unconditional variances on both sides of (53). □

The complexity of this expression prevents us from obtaining clear results about how RI and RB affect the variance of individual savings. Fig. 9 illustrates the effects of RB (ϑ) on the relative volatility of individual savings for different values of κ, when $\omega_\zeta^2 = 4.29$ and $\rho = 0.88$. We can see from the figure that μ_d is decreasing with channel capacity (κ) and is increasing with robustness (ϑ). The main reason for the first result is that saving is treated as a residual in the consumption-saving problem with a given income process. The RI-induced noises not only affect the consumption function but also affect the saving function via the residual term, which increases both the volatility of consumption and the volatility of saving. In contrast, the second result that RB reduces the volatility of saving is because it reduces the volatility of consumption and RB by itself does not generate any additional terms that have the potential to increase the volatility of saving. Furthermore, if we now consider an aggregate economy with a continuum of ex ante identical inattentive consumers with the same preference for robustness and each of them has the consumption function (41), then the total saving demand in the economy is equal to zero. The intuition is simple. The saving function can be expressed as a combination of different types of income and noise shocks: ς, ε or ζ, and ξ, and all of these shocks are idiosyncratic. These idiosyncratic shocks cancel out after aggregating across consumers and therefore have no effect on aggregate savings.

4.4. *Welfare and policy implications under RI-RB*

Since information-processing constraints cannot help in individuals' optimization, the average welfare difference between the RI-RB and RB 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 consumption path? Alternatively, we ask what an agent would pay to increase channel capacity κ to ∞ (so that the optimal choice is $\theta = 1$). Furthermore, we can also examine how the preference for RB affects the RI-induced welfare losses in the RI-RB model.[41]

To compute the welfare losses of consumers who are concerned about model misspecification as proposed in Sections 3.1 and 3.2 from RI due to deviating from the full-information RE path, we need the costs of deviating from the full-information benchmark. Specifically, following Sims (2003) and Luo and Young (2010), we first define the average loss function due to limited information-processing capacity as:

$$L \equiv E\left[v\left(s_t\right) - \widehat{v}\left(\widehat{s}_t\right)\right], \tag{55}$$

where $\widehat{v}\left(\widehat{s}_t\right) = \Omega\left(\widehat{s}_t - \frac{\overline{c}}{R-1}\right)^2 + \rho$ is the value function under RI-RB, $v\left(s_t\right) = \Omega_0\left(s_t - \frac{\overline{c}}{R-1}\right)^2 + \rho_0$ is the corresponding value function when $\kappa = \infty$ (i.e., $\theta = 1$), $\Omega_0 = -\frac{R(R-1)}{2(1-\Pi_0)}$, $\rho_0 = \frac{\vartheta}{2(R-1)}\ln\left(1 - \frac{(R-1)\Pi_0}{1-\Pi_0}\right)$, $\Pi_0 = \frac{R\omega_\zeta^2}{\vartheta}$, $\Omega = -\frac{R(R-1)}{2(1-\Pi)} < \Omega_0$, and $\rho = \frac{\vartheta}{2(R-1)}\ln\left(1 - \frac{(R-1)\Pi}{1-\Pi}\right)$, and $\Pi = \frac{R\omega_\eta^2}{\vartheta}$. Using the value functions under FI-RE and RI, the expression for the expected welfare loss function can be written as:

$$\begin{aligned} L = {} & \frac{R\left(R-1\right)}{2}\left(\frac{1}{1-\Pi}\operatorname{var}\left[\widehat{s}_t\right] - \frac{1}{1-\Pi_0}\operatorname{var}\left[s_t\right]\right) \\ & + \frac{R\left(R-1\right)}{2}\frac{\left(\Pi-\Pi_0\right)}{\left(1-\Pi\right)\left(1-\Pi_0\right)}\left(E\left[s_t\right]\right)^2 \\ & + \frac{\overline{c}R\left(\Pi_0-\Pi\right)}{\left(1-\Pi\right)\left(1-\Pi_0\right)}E\left[s_t\right] - \frac{R\overline{c}^2}{2\left(R-1\right)}\left(\frac{1}{1-\Pi_0} - \frac{1}{1-\Pi}\right) \\ & + \left[\frac{\vartheta}{2\left(R-1\right)}\ln\left(1 - \frac{\left(R-1\right)\Pi_0}{1-\Pi_0}\right) - \frac{\vartheta}{2\left(R-1\right)}\ln\left(1 - \frac{\left(R-1\right)\Pi}{1-\Pi}\right)\right]. \end{aligned}$$

To do quantitative welfare analysis in the models we need to know the level of $E\left[s_t\right]$. First, denote by γ the local coefficient of relative risk aversion, which equals $\gamma \equiv \frac{E[y]}{\overline{c}-E[y]}$ for the utility function $u\left(\cdot\right)$ 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 the PSID and CEX data by LNWY (2015). Specifically, we set the mean income level $E\left[y\right] = 1$, $\omega = 0.29$, and $\phi = 0.88$, and then find the value of the bliss point $\overline{c}$ that generates reasonable relative risk aversion γ. For example, if γ is equal to 1, $\overline{c} = 2E\left[y\right] = 2$. Furthermore, assume that the ratio of mean financial wealth to mean labor income is 5, that is, $E\left[b\right]/E\left[y\right] = 5$.[42] Since $s_t = b_t + \frac{1}{R-\rho}y_t + \frac{1-\rho}{(R-\rho)(R-1)}\overline{y}$, we have $E\left[s_t\right] = 55$. Following Cochrane (1989) and Pischke (1995), we use a money metric to measure the welfare cost of deviating from the infinite capacity case. Specifically, dividing the expected welfare losses $L \equiv E\left[v\left(s_t\right) - \widehat{v}\left(\widehat{s}_t\right)\right]$ by the marginal utility of a dollar at time t and converting it to dollars per period yields:

$$\text{\$ Loss/year: } \Delta^{RI} = \left(R-1\right)\frac{L}{u'\left(\overline{y}\right)}. \tag{56}$$

41 Luo and Young (2010) found that the welfare costs of RI are fairly small in the risk-sensitive environment.

42 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.

Table 1
Welfare losses due to RI for different values of ϑ.

	\$ Loss/year (Δ^{RI})	$(\vartheta, \Pi_0) = (20, 0.213)$	$(15, 0.284)$	$(10, 0.426)$	$(8, 0.533)$
$\overline{c} = 2$	$\theta = 10\%$	0.18	0.28	0.89	3.44
	20%	0.05	0.08	0.17	0.35
	30%	0.03	0.04	0.09	0.16
	50%	0.01	0.02	0.03	0.06
$\overline{c} = 3$	$\theta = 10\%$	0.29	0.50	1.66	6.57
	20%	0.08	0.13	0.32	0.67
	30%	0.04	0.07	0.16	0.31
	50%	0.02	0.03	0.06	0.12
$\overline{c} = 4$	$\theta = 10\%$	0.42	0.75	2.51	9.95
	20%	0.12	0.19	0.49	1.02
	30%	0.06	0.10	0.24	0.47
	50%	0.02	0.04	0.09	0.17

Table 1 reports welfare costs for several values of θ and ϑ for different values of $\overline{c}$. It is clear from Table 1 that the welfare losses due to RI are trivial and are increasing with the preference for RB. For example, for $\vartheta = 10$ (i.e., $\Pi_0 = 0.426$) and $\theta = 10\%$, (that is, 10% of the uncertainty is removed upon the receipt of a new signal), the loss only amounts to 1.66 dollars per year when $\overline{c} = 3$.[43] When ϑ increases from 10 to 20, the welfare loss reduces to only 29 cents. This result is similar to the findings by Pischke (1995), Luo (2008), and Luo and Young (2010), and is robust to changes in the bliss point ($\overline{c}$) and the ratio of mean financial wealth to mean income $\left(E[b]/E[y]\right)$.[44] This thus provides some evidence that it is reasonable for agents to devote low channel capacity to observing and processing information even if the agents have very strong preferences for robustness 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.

Finally, it is worth noting that we can also infer counter-cyclical policies (stabilization policies) implications under RI-RB. The traditional way to calculate welfare costs of business cycles 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 the RB model and thus amplifies the impacts of RB on total uncertainty faced by the consumers. Since stabilization policies which can reduce aggregate fluctuations affect our model economy by reducing ω_ζ^2, RB-RI consumers gain more welfare from countercyclical policies than do FI-RE consumers and RI consumers.[45]

43 In the calibration exercise in Section 5, we will show that the values of ϑ we used in the welfare calculations are plausible in the sense that they correspond to plausible values of the detection error probabilities.

44 Pischke (1995) found that in most cases the utility losses due to no information about aggregate income shocks are less than \$1 per quarter.

45 Note that ω_ζ^2 may include 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.

5. Market price of induced uncertainty

The PIH model presented in Section 3.1 is usually regarded as a partial equilibrium model. However, as noted in Hansen (1987) and HST (1999), it can be interpreted as a general equilibrium model with a linear production technology and an exogenous income process. Given the expression of optimal consumption in terms of the state variables derived from the robust version of the PIH model with inattentive agents, we can price assets by treating the process of aggregate consumption that solves the model as though it were an endowment process. In this setup, equilibrium prices are shadow prices that leave the agent content with that endowment process. HST (1999) studied how robustness and risk-sensitivity affect the predicted market price of uncertainty within a PIH model with shocks to both labor income and preferences, and find that RB or RS significantly alter the model's predictions on the market price of uncertainty and thus provides an alternative explanation for the observed market price of risk and the equity premium puzzle.

In this section, using the optimal consumption and saving decisions derived in the previous sections, we will explore how induced uncertainty due to the interactions of RB and RI with income shocks affect the market price of uncertainty. We first consider the single-period asset pricing case. In this case, we assume that the agent purchases a security at period t at a price q_t, holds it for one period, and then sells it at $t+1$ for a total payoff ϕ_{t+1} in terms of the consumption good after collecting the dividend. Under this assumption, the following Euler equation holds:

$$q_t = \widetilde{E}_t\left[\left(\beta\frac{u'(\widehat{s}_{t+1})}{u'(\widehat{s}_t)}\right)\phi_{t+1}\right], \tag{57}$$

where $\beta\frac{u'(\widehat{s}_{t+1})}{u'(\widehat{s}_t)}$ is the stochastic discount factor (SDF) and $\widetilde{E}_t[\cdot]$ is the distorted conditional expectations operator. Note that here SDF depends on the perceived states because optimal consumption is a linear function of perceived permanent income; because $\widehat{s}_{t+1}$ is a function of the true state s_{t+1}, s_t will also affect the SDF. The corresponding formula for q_t in terms of the original conditional expectations operator can be written as

$$q_t = E_t\left[m_{t,t+1}\phi_{t+1}\right], \tag{58}$$

where $m_{t,t+1}$ depends not only on the usual SDF but also on robustness. As has been shown in HST (1999), RB or RS are reflected in the usual measure of the SDF being scaled by a random variable with conditional mean 1. They also show that this multiplicative adjustment to the SDF increases the volatility of the SDF and thus drives up the risk premium. To explore the effects of induced uncertainty on the market price of uncertainty, we write (58) as

$$q_t = E_t\left[\phi_{t+1}\right]E_t\left[m_{t,t+1}\right] + \text{cov}_t\left(m_{t,t+1},\phi_{t+1}\right),$$

which leads to the following price bound:

$$q_t \geq E_t\left[\phi_{t+1}\right]E_t\left[m_{t,t+1}\right] - \text{sd}_t\left(m_{t,t+1}\right)\text{sd}_t\left(\phi_{t+1}\right),$$

where $\text{sd}_t(\cdot)$ denotes the conditional standard deviation. If we define the market price of uncertainty (MPU) as $\text{MPU} \equiv \frac{\text{sd}_t(m_{t,t+1})}{E_t[m_{t,t+1}]}$, the Hansen–Jagannathan (HJ) bound can be rewritten as

$$\text{MPU} \geq \frac{E_t\left[\phi_{t+1}/q_t\right]}{\text{sd}_t\left(\phi_{t+1}/q_t\right)},$$

where the RHS is the Sharpe ratio and is above 0.2 for most industrial countries.[46] In the standard full-information state- and time-separable utility model, the value of MPU is an order of magnitude lower than what is required for this inequality to be satisfied, which is just another manifestation of the equity premium puzzle: consumption growth is smooth, uncorrelated with returns, and has near zero autocorrelation, leading to a small cost of bearing uncertainty.[47]

5.1. MPU under RB and RI

The SDF, $m_{t,t+1}$, can be decomposed into

$$m_{t,t+1} = m^f_{t,t+1} m^{rb}_{t,t+1},$$

where $m^f_{t,t+1} \equiv \beta \frac{u'(\widehat{s}_{t+1})}{u'(\widehat{s}_t)}$ is the "familiar" stochastic discount factor ($\vartheta = \infty$) and $m^{rb}_{t,t+1}$ is the Radon–Nikodym derivative, or the likelihood ratio of the distorted conditional probability of $\widehat{s}_{t+1}$ with respect to the approximating conditional probability. Under the two-player game specification of RB, (18), asset prices are computed using the pessimistic view of the next period's shock: $\widetilde{\eta}_{t+1} = \omega_\eta \widetilde{\epsilon}_{t+1} = \omega_\eta (\epsilon_{t+1} + w_t)$, where $\widetilde{\epsilon}_{t+1}$ is a normally distributed variable with mean $\omega_\eta w_t$ and variance ω_η^2. In this case, the Radon–Nikodym derivative can be written as

$$m^{rb}_{t,t+1} \equiv \frac{\exp\left(-\left(\widetilde{\epsilon}_{t+1} - w_t\right)^2/2\right)}{\exp\left(-\widetilde{\epsilon}^2_{t+1}/2\right)} = \exp\left(\widetilde{\epsilon}_{t+1} w_t - w_t^2/2\right).$$

By construction, we obtain $E_t\left[m^{rb}_{t,t+1}\right] = 1$. By straightforward calculations, we obtain the following conditional second moment of $m^{rb}_{t,t+1}$ as a means for computing its conditional variance:

$$E_t\left[\left(m^{rb}_{t,t+1}\right)^2\right] = \exp\left(w_t^2\right). \tag{59}$$

The following proposition summarizes the result on how induced uncertainty affects the market price of uncertainty.

Proposition 7. *The expression for the market price of induced uncertainty is*

$$\text{sd}_t\left(m^{rb}_{t,t+1}\right) = \sqrt{\exp\left(w_t^2\right) - 1} \cong |w_t| \tag{60}$$

for small distortions, where w_t is the mean of the worse-case shock:

$$w_t = \Theta\left[(R-1)\widehat{s}_t - \overline{c}\right] \tag{61}$$

and $\Theta = \frac{\Pi/\omega_\eta}{1-\Pi}$.

Proof. Given $E_t\left[m^{rb}_{t,t+1}\right] = 1$, we can obtain (60) using (59). It is also straightforward to derive (61) using (20). □

[46] In the U.S. data presented in Campbell (2003), the Sharpe ratio is about 0.52 (annualized) during 1947–1998. Using a longer annual U.S. time series put together by Shiller yields a similar value of the Sharpe ratio.

[47] The standard deviation of aggregate consumption growth is 0.84 percent, the correlation with real returns on the S&P500 Index is 0.22, and the autocorrelation is 0.08, using nondurables and services deflated by the PCE deflator.

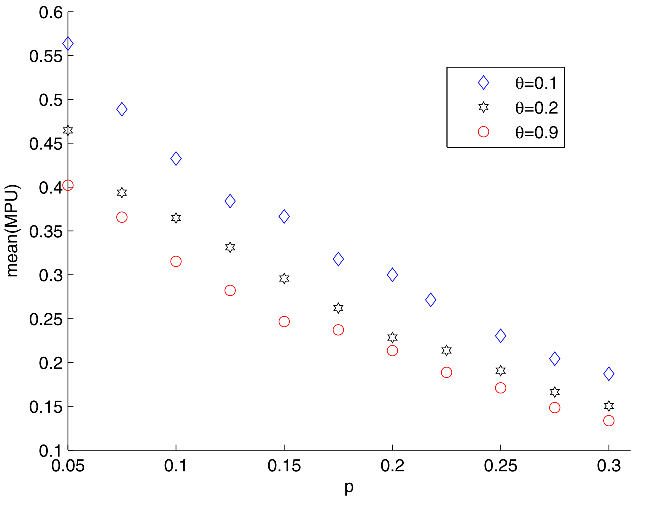

Fig. 10. Effects of p on the mean of MPU for different θ.

Expression (60) clearly shows that the amount of market price of uncertainty contributed by the Radon–Nikodym derivative is approximately equal to the norm of the mean of the worse-case shock (w).[48] Note that Θ can be used to measure the importance of RB and RI in determining the market price of uncertainty for given $\widehat{s}_t$ and $\partial\Theta/\partial\Pi > 0$. Specifically, both ϑ and κ influence $m^{rb}_{t,t+1}$ through their effects on Π.[49] Lowering ϑ strengthens the preference for robustness and then drive $m^{rb}_{t,t+1}$ away from 1 by increasing Π. Lowering κ reduces the Kalman gain, and then increases ω_η and Π. However, since the evolution of $\widehat{s}_t$ is also affected by ω_η and Π, we have to take both the Θ term and the $(R-1)\widehat{s}_t - \overline{c}$ term in (61) into account when evaluating how the interaction of RB and RI affects the market price of uncertainty.

To fully explore how induced uncertainty due to RB and RI affects the market price of uncertainty, here we also adopt the DEP to calibrate the value of Π that summarizes the interaction between RB and RI. Following the consumption and saving literature, we set $R = 1.02$, $\omega/y_0 = 0.1$, $\rho = 0.9$, and $\overline{c} = 4y_0$.[50] Using these parameter values, Fig. 10 shows how p affects the mean of MPU under RB and RI.[51] Using either the data set documented in Campbell (2003) or that provided by Shiller, the estimated Sharpe ratio for the postwar U.S. time series is greater than 50 percent a year. Fig. 11 plots the Hansen–Jagannathan (HJ) bound under RB and RI when $\theta = 0.1$ using the Shiller data set. It is clear from this figure that the model's predicted MPU can enter the HJ bound when both the preference for RB and the degree of RI are strong, i.e., when $p = 0.05$ or $p = 0.1$ and $\theta = 0.1$. As we have discussed in Section 4.2, although $\theta = 10\%$ is a very low number and is well below the total information-processing ability of human beings, it

48 In other words, $|w_t|$ is an upper bound on the approximate enhancement to the market price of uncertainty caused by the interaction of RB and RI.

49 When $\kappa = \infty$, i.e., no RI, (61) reduces to $w_t = \frac{\Pi/\omega_\eta}{1-\Pi}\left[(R-1)s_t - \overline{c}\right]$. Without RB, $w_t = 0$.

50 The number of periods used in the simulation, T, is set to be the actual length of the data we study. For example, if we consider the post-war U.S. annual time series data provided by Shiller from 1946–2010, $T = 65$.

51 Because the effects of RB and RI on the mean and median of MPU are quite similar, we focus on the mean of MPU in our subsequent analysis.

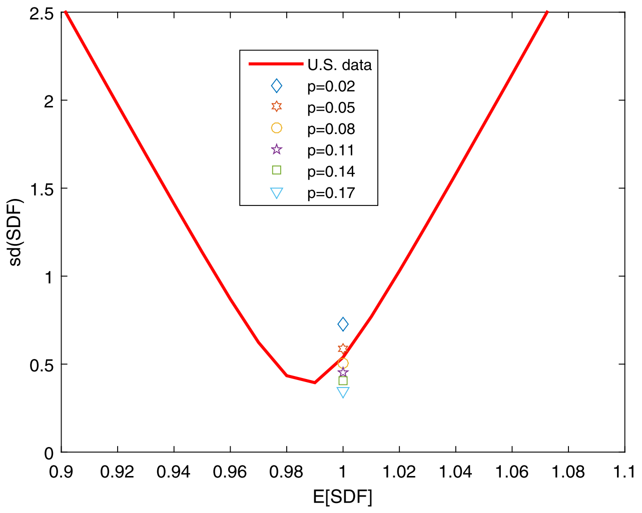

Fig. 11. HJ bound under RB and RI ($\theta = 0.1$).

is not unreasonable in practice for ordinary consumers because they also face many other competing demands on capacity. The capacity given in our model can thus be regarded as a shortcut to small fractions of consumers' total capacity used to monitor their total resources hit by the innovation to total resources. The low capacity can also be rationalized by the fact that the welfare losses due to limited capacity are very tiny. (See Table 1.) In addition, although one could argue that the DEP between (5%, 10%) implies excessive pessimism, the data were actually being generated in real time and agents did not have access to the full sample when they made decisions and when the data were actually being generated. As an informal correction for this, the effective time horizon should be shorter than the actual time horizon (T) and the resulting calibrated DEP is higher.[52]

If we consider the international developed-country data set in Campbell (2003), the Sharpe ratio is between 15 percent and 20 percent for Australia and Italy, between 20 percent and 30 percent for Canada and Japan, and above 30 percent for all the other countries. In the long-run annual data sets the lower bound on the standard deviation exceeds 30 percent for all three countries. From Fig. 11, we can see that the theoretical MPU satisfies the HJ bound for higher values of p. Our results are robust to reasonable changes in the values of R, ω/y_0, ρ, and $\overline{c}$.

Since θ is a critical parameter, we show explicitly in Fig. 12 how the predictions for each value of p vary when $\theta = 0.5$. It is clear that a low value of θ is important for satisfying the HJ bounds; as we reduce θ we reduce the required level of p, so that the worst-case model can move closer to the approximating model (that is, we can reduce the required fear of model misspecification).

[52] To match the observed volatility of six U.S. dollar exchange rates (the Australian dollar, the Canadian dollar, the Danish dollar, the Japanese yen, the Swiss franc, and the British pound), Djeutem and Kasa (2013) showed that the detection error probability should be set between 7.5% to 13.1%. (See Table 5 in their paper.)

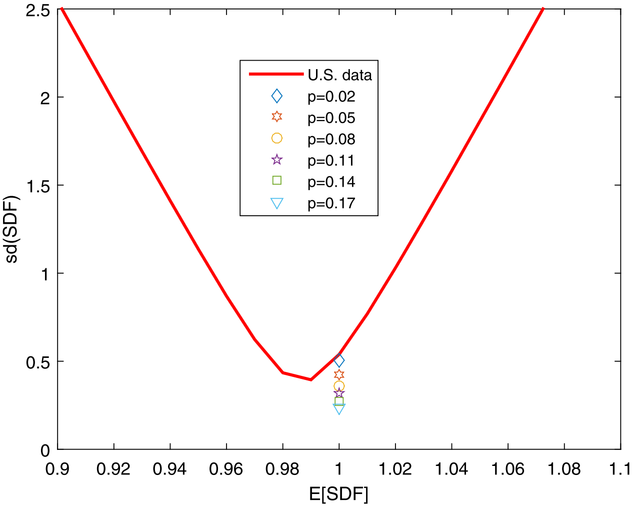

Fig. 12. HJ bound under RB and RI ($\theta = 0.5$).

5.2. *MPU under RS and RI*

Under the risk-sensitivity specification of RB, (23), we can also compute the corresponding market price of uncertainty. The intertemporal marginal rate of substitution between t and $t+1$ can be written as

$$m_{t,t+1} = m^f_{t,t+1} m^{rs}_{t,t+1},$$

where $m^f_{t,t+1} \equiv \beta \frac{u'(\widehat{s}_{t+1})}{u'(\widehat{s}_t)}$ and $m^{rs}_{t,t+1} \equiv \frac{\exp(-v_{t+1}/\vartheta)}{E_t[\exp(-v_{t+1}/\vartheta)]} = \frac{\exp\left(-\left(\Omega\,\widehat{s}^2_{t+1}+\rho\right)/\vartheta\right)}{\exp(-(\widehat{\Omega}\widehat{s}_t^2+\widehat{\rho})/\vartheta)}$. Using a formula found in Jacobson (1973) and used in HST (1999), we have $E_t\left[\left(m^{rs}_{t,t+1}\right)^2\right] = \exp(-2[(\widetilde{\Omega} - \widehat{\Omega})\widehat{s}_t^2 + (\widetilde{\rho} - \widehat{\rho})]/\vartheta)$, because $E_t\left[\exp\left(-2\left(\Omega\widehat{s}^2_{t+1} + \rho\right)/\vartheta\right)\right] = \exp\left(-2\left(\widetilde{\Omega}\widehat{s}_t^2 + \widetilde{\rho}\right)/\vartheta\right)$. The following proposition summarizes the result on how induced uncertainty affects the amount of market price of uncertainty under the RS specification of RB.

Proposition 8. *Under RS-RI, the Radon–Nikodym derivative is*

$$E_t\left[\left(m^{rs}_{t,t+1}\right)^2\right] = \exp\left(\Xi^{rs}\left[(R-1)\widehat{s}_t - \overline{c}\right]^2\right)\Upsilon, \tag{62}$$

where $\Xi^{rs} \equiv \frac{R\Pi(1-R\Pi)}{\vartheta_0(1-\Pi)^2[1-(2R-1)\Pi]}$ *and* $\Upsilon \equiv \frac{1-(R-1)\Pi/(1-\Pi)}{\sqrt{1-2(R-1)\Pi/(1-\Pi)}}$. *The market price of induced uncertainty is*

$$\mathrm{sd}_t\left(m^{rs}_{t,t+1}\right) = \sqrt{\exp\left(\Xi^{rs}\left((R-1)\widehat{s}_t - \overline{c}\right)^2\right)\Upsilon - 1} \cong \left|\sqrt{\Xi^{rs}}\left((R-1)\widehat{s}_t - \overline{c}\right)\right| \tag{63}$$

Proof. See Appendix A.4. □

Denote $\Xi^{rb} = \Theta^2$, we have

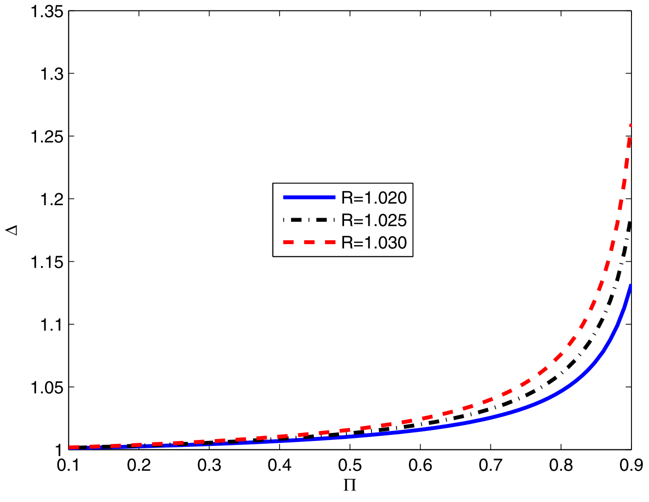

Fig. 13. Difference of MPU under RB+RI and RS+RI.

$$\Delta \equiv \sqrt{\frac{\Xi^{rs}}{\Xi^{rb}}} = \sqrt{\frac{1 - R\Pi}{1 - (2R - 1)\,\Pi}}, \tag{64}$$

which is close 1 when R is close 1 and Π is not too large. Using (64), it is straightforward to show that $\frac{\partial \Delta}{\partial \Pi} > 0$ and $\frac{\partial \Delta}{\partial \theta} < 0$, i.e., the stronger the degree of RI, the larger the difference of the market price of uncertainty under the RB and RS specifications. Fig. 13 illustrates how Δ varies with Π for given values of R. It is clear that theoretically the difference of the market price of uncertainty between RB and RS under RI can be very significant. For example, when $R = 1.02$, $\Delta = 1.25$ when $\Pi = 0.93$. In other words, the MPU under the RS-RI specification is 25 percent higher than that under the RB-RI specification when the two models are observationally equivalent. However, after calibrating an empirically-plausible ϑ (and Π) using the DEP, $\mathrm{sd}_t\left(m^{rb}_{t,t+1}\right)$ and $\mathrm{sd}_t\left(m^{rs}_{t,t+1}\right)$ are very close because R is close to 1 and the calibrated values of Π are between 0.1 and 0.2 for $p = 0.05$.[53] Consequently, the two specifications have similar effects on the mean of for MPU and the HJ bound for different values of p and θ.[54]

6. Conclusion

This paper has provided a characterization of the consumption-savings behavior of agents who have a preference for robustness (worries about model misspecification) and limited information-processing ability. After obtaining the optimal individual decisions, we explore how two types of induced uncertainty, state uncertainty due to RI and model uncertainty due to RB, affect consumption and saving decisions and consumption and income inequality as well as the market prices of uncertainty. Specifically, we show that concerns about different types of model misspecification – (i) disturbances to the perceived permanent income and (ii) the Kalman gain – can have opposite effects on consumption and savings via interacting with finite capacity but the

[53] The calibrated values of Π are lower for higher values of p.

[54] The detailed results are available upon request from the corresponding author.

first type of model misspecification dominates the second one in determining the propagation mechanism in the control and filtering problem. In addition, we show that once allowing RB consumers to use the robust Kalman filter to update the perceived state, the simple observational equivalence (OE) between RB and RS obtained in HST (1999) no longer holds; instead, we find a more complicated OE between RB and RS. We also find that if channel capacity is elastic, RB increases the optimal level of attention.

We then apply our framework to two questions – the relative volatility of individual consumption versus income and the equity premium. Our model is capable of producing a much higher relative volatility of consumption than other frameworks, and also can match the observed decline over time (the decoupling of consumption and income inequality). Our model also produces a risk-return trade-off that lies inside the Hansen–Jagannathan bounds without relying on excessive pessimism, provided information flow is relatively low; since we also find that the costs of low information capacity are small, we view these outcomes as supporting our framework.

Appendix A

A.1. Solving the two-player game version of the robust model

To solve the Bellman equation (18) subject to (16), $\widehat{s}_{t+1} = R\widehat{s}_t - c_t + \omega_\eta w_t + \eta_{t+1}$, we conjecture that

$$\widehat{v}(\widehat{s}_t) = -C - B\widehat{s}_t - A\widehat{s}_t^2, \tag{65}$$

where A, B, and C are constants to be determined. Substituting this guessed value function into the Bellman equation (18) gives

$$-C - B\widehat{s}_t - A\widehat{s}_t^2 = \max_{c_t}\min_{w_t}\left\{-\frac{1}{2}(c_t - \overline{c})^2 + \beta E_t\left[\widetilde{\vartheta}_0 w_t^2 - C - B\widehat{s}_{t+1} - A\widehat{s}_{t+1}^2\right]\right\}, \tag{66}$$

where $\widetilde{\vartheta} = \vartheta/2$. We can do the min and max operations in any order, so we choose to do the minimization first. The first-order condition for w_t is

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

which means that

$$w_t = \frac{B + 2A(R\widehat{s}_t - c_t)}{2\left(\widetilde{\vartheta} - A\omega_\eta^2\right)}\omega_\eta. \tag{67}$$

Substituting (67) back into (66) gives

$$\begin{aligned} &- A\widehat{s}_t^2 - B\widehat{s}_t - C \\ &= \max_{c_t}\left\{-\frac{1}{2}(\overline{c} - c_t)^2 + \beta E_t\left[\widetilde{\vartheta}\left[\frac{B + 2A(Rs_t - c_t)}{2\left(\vartheta - A\omega_\eta^2\right)}\omega_\eta\right]^2 - As_{t+1}^2 - Bs_{t+1} - C\right]\right\}, \end{aligned}$$

where $\widehat{s}_{t+1} = R\widehat{s}_t - c_t + \omega_\eta w_t + \eta_{t+1}$. The first-order condition for c_t is

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

Using the solution for ν_t the solution for consumption is

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

Substituting the above expressions into the Bellman equation gives

$$-A\widehat{s}_t^2-B\widehat{s}_t-C$$
$$=-\frac{1}{2}\left(\frac{2A\beta R}{1-A\omega_\eta^2/\widetilde{\vartheta}+2\beta A}\widehat{s}_t+\frac{-2\beta A\overline{c}+\beta B}{1-A\omega_\eta^2/\widetilde{\vartheta}+2\beta A}\right)^2$$
$$+\frac{\beta\widetilde{\vartheta}\omega_\eta^2}{\left(2\left(\widetilde{\vartheta}-A\omega_\eta^2\right)\right)^2}\left[\frac{2AR\left(1-A\omega_\eta^2/\widetilde{\vartheta}\right)}{1-A\omega_\eta^2/\widetilde{\vartheta}+2\beta A}\widehat{s}_t+B-\frac{2\overline{c}\left(1-A\omega_\eta^2/\widetilde{\vartheta}\right)A+2\beta AB}{1-A\omega_\eta^2/\widetilde{\vartheta}+2\beta A}\right]^2$$
$$-\beta A\left\{\left[\frac{R}{1-A\omega_\eta^2/\widetilde{\vartheta}+2\beta A}\widehat{s}_t-\frac{-B\omega_\eta^2/\widetilde{\vartheta}+2c+2B\beta}{2\left(1-A\omega_\eta^2/\widetilde{\vartheta}+2\beta A\right)}\right]^2+\omega_\eta^2\right\}$$
$$-\beta B\left[\frac{R}{1-A\omega_\eta^2/\widetilde{\vartheta}+2\beta A}\widehat{s}_t-\frac{-B\omega_\eta^2/\widetilde{\vartheta}+2c+2B\beta}{2\left(1-A\omega_\eta^2/\widetilde{\vartheta}+2\beta A\right)}\right]-\beta C.$$

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

$$A=\frac{\beta R^2-1}{2\beta-\omega_\eta^2/\widetilde{\vartheta}},\ B=\frac{\left(\beta R^2-1\right)\overline{c}}{(R-1)\left(\omega_\eta^2/\widetilde{\vartheta}-2\beta\right)},$$
$$C=\frac{R\left(\beta R^2-1\right)}{\left(2\beta R-R\omega_\eta^2/\widetilde{\vartheta}\right)(R-1)^2}\left((R-1)\,\omega_\eta^2+\overline{c}^2\right),$$

where $\widetilde{\vartheta}=\vartheta/2$. When $\beta R=1$, we obtain the consumption function (19) in the text.

A.2. *Optimality of ex post Gaussianity under RB*

Following Sims (2003, Section 5), we first define the expected loss function due to limited information-processing capacity as follows:

$$L_t=E_t\left[v\left(s_t\right)-\widehat{v}\left(\widehat{s}_t\right)\right],\tag{68}$$

where $\widehat{v}\left(\widehat{s}_t\right)=\Omega\left(\widehat{s}_t-\frac{\overline{c}}{R-1}\right)^2+\rho$ is the value function under RI-RB, $v\left(s_t\right)=\Omega_0\left(s_t-\frac{\overline{c}}{R-1}\right)^2+\rho_0$ is the corresponding value function when $\kappa=\infty$ (i.e., $\theta=1$), $\Omega_0=-\frac{R(R-1)}{2(1-\Pi_0)}$, $\rho_0=$

$\frac{\vartheta}{2(R-1)}\ln\left(1-\frac{(R-1)\Pi_0}{1-\Pi_0}\right)$, $\Pi_0=\frac{R\omega_\zeta^2}{\vartheta}$, $\Omega=-\frac{R(R-1)}{2(1-\Pi)}<\Omega_0$, and $\rho=\frac{\vartheta}{2(R-1)}\ln\left(1-\frac{(R-1)\Pi}{1-\Pi}\right)$. It is straightforward to show that:

$$\begin{aligned}
&\min E_t\left[v\left(s_t\right)-\widehat{v}\left(\widehat{s}_t\right)\right]\\
&\quad=\min E_t\left[\Omega_0\left(s_t-\frac{\overline{c}}{R-1}\right)^2+\rho_0-\left[\Omega\left(\widehat{s}_t-\frac{\overline{c}}{R-1}\right)^2+\rho\right]\right]\\
&\Longleftrightarrow \min E_t\left[\Omega_0\left(s_t^2-\frac{2\overline{c}}{R-1}s_t\right)-\Omega\left(\widehat{s}_t^2-\frac{2\overline{c}}{R-1}\widehat{s}_t\right)-\left(\Omega-\Omega_0\right)\left(\widehat{s}_t^2-\frac{2\overline{c}}{R-1}\widehat{s}_t\right)\right]\\
&\Longleftrightarrow \min E_t\left[\Omega_0\left(s_t^2-\widehat{s}_t^2\right)-\Omega_0\frac{2\overline{c}}{R-1}\left(s_t-\widehat{s}_t\right)\right]\\
&\Longleftrightarrow \min E_t\left[\Omega_0\left(s_t-\widehat{s}_t\right)^2\right]\\
&\Longleftrightarrow \min \Omega_0\Sigma_t,
\end{aligned}\tag{69}$$

where we use the fact that $\widehat{s}_t=E_t\left[s_t\right]$. When $\vartheta=\infty$ (or $\Pi_0=0$), this loss function reduces to $-\frac{R(R-1)}{2}\Sigma_t$ which is just that obtained in Sims (2003, 2010). Since the only difference in these two settings is just in the constant coefficient in the loss function, RB does not affect the optimality of ex post Gaussianity in the linear-quadratic-Gaussian setting.

A.3. Calibrating the robustness parameter

Using the model detection error probability (DEP), we can then infer what values of the RB parameter imply reasonable fears of model misspecification for empirically-plausible approximating models. The DEP denoted by p is a measure of how far the distorted model can deviate from the approximating model without being discarded; low values for this probability mean that agents are unwilling to discard very many models, implying that the cloud of models surrounding the approximating model is large (since agents want errors to be rare, they push the two models very far apart). The value of p is determined by the following procedure. Let model A denote the approximating model, (12), and model B be the distorted model, (16). Define p_A as

$$p_A=\text{Prob}\left(\ln\left(\frac{L_A}{L_B}\right)<0\,\middle|\,A\right),\tag{70}$$

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

$$p_B=\text{Prob}\left(\log\left(\frac{L_A}{L_B}\right)>0\,\middle|\,B\right).\tag{71}$$

Given initial priors of 0.5 on each model and that the length of the sample is T, the DEP, p, is defined as the average of p_A and p_B:

$$p\left(\vartheta;\Pi;T\right)=\frac{1}{2}\left(p_A+p_B\right),\tag{72}$$

where ϑ is the robustness parameter used to generate model B. Given this definition, we can see that $1-p$ measures the probability that econometricians can distinguish the approximating model from the distorted model. Now we show how to compute the model detection error probability in the RB model. The general idea of the calibration exercise is to find a value of (or Π) such that $p(\vartheta;\Pi)$ equals a given value (for example, 5 percent or 10 percent) after simulating model A, (12), and model B, (16).

A.4. Computing the market price of uncertainty

Given the value function we obtained in Section 3.1,

$$v(\widehat{s}_t)=\Omega\left(\widehat{s}_t-\frac{\overline{c}}{R-1}\right)^2+\rho, \tag{73}$$

where $\Omega=-\frac{R(R-1)}{2(1-\Pi)}$ and $\rho=\frac{\vartheta}{2(R-1)}\ln\left(1-\frac{(R-1)\Pi}{1-\Pi}\right)$, it follows from Jacobson (1973) and HST (1999) that the risk-sensitivity operator can be written as

$$\mathcal{R}_t\left[\widehat{v}(\widehat{s}_{t+1})\right]=-\vartheta\log E_t\left[\exp\left(-\widehat{v}(\widehat{s}_{t+1})/\vartheta\right)\right]=\widehat{\Omega}\left(\widehat{s}_t-\frac{\overline{c}}{R-1}\right)^2+\widehat{\rho},$$

where

$$\widehat{\Omega}=\rho_s^2\Omega\left(1-\frac{2}{\vartheta}\Omega\omega_\eta^2\left(1+\frac{2}{\vartheta}\Omega\omega_\eta^2\right)^{-1}\right)=-\frac{R(R-1)(1-R\Pi)}{2(1-\Pi)^2}$$

and

$$\widehat{\rho}=\rho+\frac{\vartheta}{2}\ln\left(1+\frac{2}{\vartheta}\Omega\omega_\eta^2\right)=\frac{\vartheta}{2(R-1)}\ln\left(1-\frac{2(R-1)\Pi}{1-\Pi}\right)+\frac{\vartheta}{2}\ln\left(1-\frac{(R-1)\Pi}{(1-\Pi)}\right),$$

where we assume that $\frac{2(R-1)\Pi}{1-\Pi}<1$.

Given that

$$m_{t,t+1}^{rs}\equiv\frac{\exp(-v_{t+1}/\vartheta)}{E_t\left[\exp(-v_{t+1}/\vartheta)\right]}=\frac{\exp\left(-\left(\Omega\widehat{s}_{t+1}^2+\rho\right)/\vartheta\right)}{\exp\left(-\left(\widehat{\Omega}\widehat{s}_t^2+\widehat{\rho}\right)/\vartheta\right)},$$

we have

$$\left(m_{t,t+1}^{rs}\right)^2=\frac{\exp\left(-2\left(\Omega\widehat{s}_{t+1}^2+\rho\right)/\vartheta\right)}{\exp\left(-2\left(\widehat{\Omega}\widehat{s}_t^2+\widehat{\rho}\right)/\vartheta\right)}.$$

Multiplying the numerator and denominator by the time t conditional mean of the exponential term in the numerator, $E_t\left[\exp\left(-2\left(\Omega\widehat{s}_{t+1}^2+\rho\right)/\vartheta\right)\right]$, gives

$$\left(m_{t,t+1}^{rs}\right)^2=\frac{E_t\left[\exp\left(-2\left(\Omega\widehat{s}_{t+1}^2+\rho\right)/\vartheta\right)\right]}{\exp\left(-2\left(\widehat{\Omega}\widehat{s}_t^2+\widehat{\rho}\right)/\vartheta\right)}\frac{\exp\left(-2\left(\Omega\widehat{s}_{t+1}^2+\rho\right)/\vartheta\right)}{E_t\left[\exp\left(-2\left(\Omega\widehat{s}_{t+1}^2+\rho\right)/\vartheta\right)\right]},$$

where the exponential term, $E_t\left[\exp\left(\sigma\left(\Omega\widehat{s}_{t+1}^2+\rho\right)\right)\right]$, can be computed using a formula found in Jacobson (1973):

$$E_t\left[\exp\left(-2\left(\Omega\widehat{s}_{t+1}^2+\rho\right)/\vartheta\right)\right]=\exp\left(-2\left(\widetilde{\Omega}\widehat{s}_t^2+\widetilde{\rho}\right)/\vartheta\right),$$

where $\widetilde{\Omega}=\rho_s^2\Omega\left(1-\frac{4}{\vartheta}\Omega\omega_\eta^2\left(1+\frac{4}{\vartheta}\Omega\omega_\eta^2\right)^{-1}\right)=\frac{R\rho_s^2\Omega}{R+4\Omega\Pi}$ and $\widetilde{\rho}=\rho+\frac{\vartheta}{4}\ln\left(1+\frac{1}{\vartheta}\Omega\omega_\eta^2\right)=\rho+\frac{\vartheta}{4}\ln\left(1-\frac{2(R-1)\Pi}{(1-\Pi)}\right)$. Therefore, we obtain

$$E_t\left[\left(m_{t,t+1}^{rs}\right)^2\right] = \exp\left(-2\left(\left(\widetilde{\Omega}-\widehat{\Omega}\right)\widehat{s}_t^2+\left(\widetilde{\rho}-\widehat{\rho}\right)\right)/\vartheta\right),$$

which yields (62) in the main text.

References

Anderson, Evan W., Hansen, Lars Peter, Sargent, Thomas J., 2003. A quartet of semigroups for model specification, robustness, prices of risk, and model detection. J. Eur. Econ. Assoc. 1 (1), 68–123.

Backus, David K., Routledge, Bryan R., Zin, Stanley E., 2004. Exotic preferences for macroeconomists. NBER Macroecon. Annu. 2004, 319–414.

Blundell, Richard, Pistaferri, Luigi, Preston, Ian, 2008. Consumption inequality and partial insurance. Am. Econ. Rev. 98 (5), 1887–1921.

Caballero, Ricardo, 1990. Consumption puzzles and precautionary savings. J. Monet. Econ. 25 (1), 113–136.

Cagetti, Marco, Hansen, Lars Peter, Sargent, Thomas J., Williams, Noah, 2002. Robustness and pricing with uncertain growth. Rev. Financ. Stud. 15, 363–404.

Campbell, John, 2003. Consumption-based asset pricing. In: Constantinides, George, Harris, Milton, Stultz, Rene (Eds.), Handbook of the Economics of Finance, vol. 1B. North-Holland, Amsterdam, pp. 803–887.

Campbell, John, Deaton, Angus, 1989. Why is consumption so smooth? Rev. Econ. Stud. 56 (3), 357–374.

Chen, Zengjing, Epstein, Larry, 2002. Ambiguity, risk, and asset returns in continuous time. Econometrica 70 (4), 1403–1443.

Cochrane, John, 1989. The sensitivity of tests of the intertemporal allocation of consumption to near-rational alternatives. Am. Econ. Rev. 9 (3), 319–337.

Djeutem, Edouard, Kasa, Kenneth, 2013. Robustness and exchange rate volatility. J. Int. Econ. 91 (1), 27–39.

Epstein, Larry G., Wang, Tan, 1994. Intertemporal asset pricing under Knightian uncertainty. Econometrica 62 (2), 283–322.

Epstein, Larry G., Zin, Stanley E., 1989. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. Econometrica 57 (4), 937–969.

Epstein, Larry G., Schneider, Martin, 2010. Ambiguity and asset markets. Annu. Rev. Financ. Econ. 2 (1), 315–346.

Flavin, Marjorie A., 1981. The adjustment of consumption to changing expectations about future income. J. Polit. Econ. 89 (5), 974–1009.

Gilboa, Itzhak, Schmeidler, David, 1989. Maxmin expected utility with non-unique prior. J. Math. Econ. 18 (2), 141–153.

Hall, Robert E., 1978. Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence. J. Polit. Econ. 86 (6), 971–987.

Hansen, Lars Peter, 1987. Calculating asset prices in three example economies. In: Bewley, Truman F. (Ed.), Advances in Econometrics, Fifth World Congress. Cambridge University Press.

Hansen, Lars Peter, Sargent, Thomas J., 1995. Discounted linear exponential quadratic Gaussian control. IEEE Trans. Autom. Control 40, 968–971.

Hansen, Lars Peter, Sargent, Thomas J., 2005. Robust estimation and control under commitment. J. Econ. Theory 124 (9), 258–301.

Hansen, Lars Peter, Sargent, Thomas J., 2007. Robustness. Princeton University Press.

Hansen, Lars Peter, Sargent, Thomas J., Tallarini Jr., Thomas D., 1999. Robust permanent income and pricing. Rev. Econ. Stud. 66 (4), 873–907.

Hansen, Lars Peter, Sargent, Thomas J., Wang, Neng, 2002. Robust permanent income and pricing with filtering. Macroecon. Dyn. 6 (1), 40–84.

Jacobson, David H., 1973. Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. IEEE Trans. Autom. Control 18, 124–131.

Ju, Nengjiu, Miao, Jianjun, 2012. Ambiguity, learning, and asset returns. Econometrica 80 (2), 559–591.

Kasa, Kenneth, 2006. Robustness and information processing. Rev. Econ. Dyn. 9 (1), 1–33.

Luo, Yulei, 2008. Consumption dynamics under information processing constraints. Rev. Econ. Dyn. 11 (2), 366–385.

Luo, Yulei, 2016. Robustly strategic consumption-portfolio rules with informational frictions. Manuscript. The University of Hong Kong.

Luo, Yulei, Young, Eric R., 2010. Risk-sensitive consumption and savings under rational inattention. Am. Econ. J. Macroecon. 2 (4), 281–325.

Luo, Yulei, Nie, Jun, Young, Eric R., 2012. Robustness, information-processing constraints, and the current account in small open economies. J. Int. Econ. 88 (1), 104–120.

Luo, Yulei, Nie, Jun, Wang, Gaowang, Young, Eric R., 2015. What we don't know doesn't hurt us: rational inattention and the permanent income hypothesis in general equilibrium. Working paper 14-14. Federal Reserve Bank of Kansas City.

Maccheroni, Fabio, Marinacci, Massimo, Rustichini, Aldo, 2006. Ambiguity aversion, robustness, and the variational representation of preferences. Econometrica 74 (6), 1447–1498.

Maćkowiak, Bartosz, Wiederholt, Mirko, 2009. Optimal sticky prices under rational inattention. Am. Econ. Rev. 99 (3), 769–803.

Maenhout, Pascal J., 2004. Robust portfolio rules and asset pricing. Rev. Financ. Stud. 17 (4), 951–983.

Mondria, Jordi, 2010. Portfolio choice, attention allocation, and price comovement. J. Econ. Theory 145 (5), 1837–1864.

Moore, John B., Elliott, Robert J., Dey, Subhrakanti, 1997. Risk-sensitive generalizations of minimum variance estimation and control. J. Math. Syst. Estim. Control 7 (1), 1–15.

Peng, Lin, 2004. Learning with information capacity constraints. J. Financ. Quant. Anal. 40, 307–330.

Pischke, Jörn-Steffen, 1995. Individual income, incomplete information, and aggregate consumption. Econometrica 63, 805–840.

Shafieepoorfard, Ehsan, Raginsky, Maxim, 2013. Rational inattention in scalar LQG control. Proc. IEEE Conf. Decis. Control 52, 5733–5739.

Sims, Christopher A., 2003. Implications of rational inattention. J. Monet. Econ. 50 (3), 665–690.

Sims, Christopher A., 2010. Rational inattention and monetary economics. In: Friedman, Benjamin J., Woodford, Michael (Eds.), Handbook of Monetary Economics, pp. 155–181.

Strzalecki, Tomasz, 2011. Axiomatic foundations of multiplier preferences. Econometrica 79, 47–73.

Wang, Neng, 2003. Caballero meets Bewley: the permanent-income hypothesis in general equilibrium. Am. Econ. Rev. 93 (3), 927–936.

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

Van Nieuwerburgh, Stijn, Veldkamp, Laura, 2009. Information immobility and the home bias puzzle. J. Finance 64 (3), 1187–1215.

Van Nieuwerburgh, Stijn, Veldkamp, Laura, 2010. Information acquisition and under-diversification. Rev. Econ. Stud. 77 (2), 779–805.