Consumption Dynamics, Asset Pricing, and Welfare Effects under Information Processing Constraints*

Yulei Luo†

Princeton University

Job Market Paper: December 2004

Abstract

This paper studies consumption dynamics, asset returns and optimal portfolio choice, and welfare losses under information processing constraints (it is also called "rational inattention" (RI) in Sims (2003).) in two canonical macroeconomic models: the permanent income hypothesis model (PIH) and the consumption-based capital asset pricing model (CCAPM). It is shown that incorporating RI into these otherwise standard macroeconomic models can provide an additional propagation mechanism and largely affect the intertemporal allocation of consumption, which makes the models better explain the data in some important aspects.

The main contributions of this paper are: first, we propose a tractable analytical approach to solve the multivariate LQ or approximate LQ models with information constraints; second, given the closed-form solutions, we address a variety of consumption and asset pricing puzzles, and show how incorporating RI may resolve these puzzles in the correct direction; third, we find that the utility costs due to deviating from the first best instantly adjusted path are very trivial, which can rationalize a key assumption in Sims (2003) that consumers only devote small fractions of their capacity in observing and processing information; finally, we compare the RI model with the habit formation model.

In addition, we consider the extension to the risk-sensitive PIH model, in which the risk-sensitive preference combined with labor income uncertainty generates precautionary savings motive and then allows us to examine the effects of RI on both precautionary savings and the marginal propensity to consume out of income.

Keywords: *Rational Inattention, Stochastic Optimal Control under Imperfect Observations, Consumption Decisions, Risk Premium, Welfare Effects, Risk-sensitive LQG*

JEL Classification Codes: *C61, D81, E21, G12, G11*

* I am grateful to Christopher A. Sims for his excellent advice, encouragement, and many conversations throughout the project. I also thank Per Krusell, Jonathan Parker, Ricardo Reis, Lars Svensson, and Wei Xiong for their numerous comments and suggestions. I also benefited from comments by participants at Princeton macro seminar, Inter-University Student Conference (2004), and the Thesis Writers' macro/international finance workshop. Any remaining errors are solely mine. The updated version of this paper and a technical appendix can be downloaded from http://www.princeton.edu/~yluo/

† Department of Economics, Princeton University, Princeton, NJ 08544, email address: yluo@princeton.edu.

1. Introduction

Canonical macroeconomic theory endows agents with the ability to process as much information as they need instantaneously and having perfect information about the state. However, Shannon (1948) showed that measuring a real-valued stochastic process without error implies an infinite amount of information processing capacity. Hence, the assumption of perfect observations is not consistent with the reality since ordinary people only have limited information capacity. In his pioneering works, Sims (1998, 2003) derived a tractable framework for studying agents having information processing constraints[1] in the sense that they only have limited capacity to process information when making economic decision. Moreover, he solved the Linear Quadratic Gaussian (henceforth, LQG) permanent income (henceforth, PIH) models numerically and showed that consumption responded to income shocks with delay and gradually. A key assumption in his paper is that information imperfection emerges endogenously because agents only have finite (Shannon) capacity to process information. Therefore, the nature of the noise due to imperfect observations is determined endogenously and/or optimally when agents need to allocate their limited information capacity across various sources. This endogeneity of noise distinguishes our RI model from other LQG models with imperfect state observations that are widely studied in control theory (e.g., Whittle 1982, 1996).

It is well known that the standard full-information instantly-adjusted rational expectation models (e.g. the PIH model and the CCAPM model) can not explain some important macroeconomic phenomena and are not consistent with a host of empirical evidence. For example, the excess smoothness puzzle, the excess sensitivity puzzle, and the equity premium puzzle. Furthermore, some VAR evidence and casual micro survey evidence also reject the predictions of the standard models. For example, Sims (1998, 2003) argued that introducing adjustment cost mechanisms pervasively in the conventional DSGE models still can not explain two apparent facts from the VAR evidences: most cross-variable relationships among macroeconomic time series are smooth and delayed, but these variables respond quickly to their own shocks; Fuhrer (2000) showed that aggregate consumption exhibits gradual and "hump-shaped" responses to monetary shocks, while the standard PIH models imply consumption should jump in response to shocks; Dynan and Maki (2000) documented that consumers may not adjust their consumption in response to variation in their equity and lagged equity returns will affect future consumption growth. In this paper we suggest that allowing for information processing constraints can help us overcome some of the deficiencies in the standard full-information models. Specifically, we introduce these information constraints into two otherwise standard macroeconomic models: the PIH model and the CCAPM model and examine how incorporating RI into these models may better explain the data in some important aspects. The reasons that we adopt these two models to explore the impacts of RI are because they are two most widely studied and tested models in macroeconomics and finance, and also they are the baseline models of the theory of the intertemporal allocation of consumption that is at the heart of macroeconomics and finance.

Hall (1978) showed that under some assumptions[2] consumption process is a martingale, that is, changes in consumption could not be predictable over time because of rational expectations. However, many papers after Hall (1978) have found convincing evidences about the deviations of aggregate consumption from a martingale. Two most important deviations are the excess sensitivity of consumption to past and current changes in income (Flavin, 1982) and the excess smoothness of consumption to permanent income shock

[1] It is labelled "Rational Inattention" in Sims (2003). In this paper, we will use "information processing constraints" and "Rational Inattention" (henceforth "RI") interchangeably.

[2] E.g., an infinite horizon representative agent, time-separable preferences, nondurability, and a constant return on nonhuman wealth equal to the time discount rate.

(Campbell and Deaton, 1989, Deaton, 1992). To match the data better, some theoretical models have been developed. For example, Campbell and Mankiw (1989, 1991) proposed an alternative model composed of "rule of thumb" consumers who do not behave according to the PIH model and thus cannot dissave in order to smooth consumption; Deaton (1992) discussed the effects of a hypothesis about habit formation in consumption; Carroll (2001) examined the buffer-stock savings hypothesis; and Reis (2003) adopted inattentiveness and cost of planning.

In the first model of this paper, we study the implications of RI for the intertemporal optimal allocation of consumption in the LQG PIH framework. The main difference between this model and the model examined in Sims (2003) is that here we propose an analytical approach to solve the PIH model with any plausible income processes including both the trend-stationary process and the difference-stationary process[3]. This analytical solution gives us clear economic insight about the nature of the RI model and it also greatly facilitates comparative statics analysis. This is particularly useful to understand how introducing RI help us resolve the excess smoothness puzzle and the excess sensitivity puzzle in the consumption literature. Furthermore, it can be used to compare with other PIH models with friction (e.g., the internal additive habit-formation model). The main findings in this model are: (1) both individual and aggregate consumption respond with delay and gradually to various income innovations and changes in consumption can be characterized by MA(∞) processes, (2) consumption is sensitive to past shocks to income, and aggregate consumption is smooth relative to permanent income even if current income follows a difference-stationary process, that is, RI can be a potential explanation for the Deaton puzzle, (3) the welfare loss due to limited channel capacity may be negligible under plausible assumptions[4]. This provides some evidence that the assumption of low channel capacity seems to be quite reasonable: it is not worth collecting and digesting information to improve consumption decisions, (4) optimal choice of channel capacity is dependent on the model parameters including the interest rate, the nature of income processes, and the marginal cost of capacity, (5) consumption, savings, and wealth dynamics derived from the RI model is quite similar to that from the habit formation model.

The main mechanism behind these results is that in the RI model measurement error in observing the state is unavoidable and rational. Hence, with limited Shannon capacity, consumers can not respond instantly and without error to changes in wealth. Consequently, the responses of consumption with respect to the shocks to wealth are smooth and delayed, and changes in consumption can be predicted by past known income shocks. Furthermore, the responses to income shocks will rise gradually and eventually reach a higher flat asymptote as compared with the case without information processing constraints, because the initially undetected income shocks accumulate interest before consumption reacts fully to them.

The second topic of this paper is about RI, asset pricing, and portfolio choices. According to the canonical economic model of asset pricing, the CCAPM model, the risk of a portfolio of stocks depends on the expected covariance of equity returns with consumption: high covariance implies high risk and thus high expected return. But after two decades of research, it is well known that consumption risk measured by the contemporaneous covariance of returns with consumption in the standard CCAPM model is too small to explain the observed equity premium. In other words, the canonical model can not explain two important

[3]In Sims (2003), he solved the PIH model with i.i.d. income process analytically, while solved the multivariate state PIH model numerically since in that case consumers need to allocate their limited capacity efficiently across various states. Here we will show that in the multivariate state case what matters for consumption decision may be just a linear combination of all elements in the state vector instead of the state vector itself, and thus the multivariate state case will be reduced to the univariate state case that can be solved analytically.

[4]We use several criterions to measure welfare loss due to RI and in most cases the welfare losses are very tiny. Hence, it would be reasonable to assume that individuals would use low capacity in processing information since increasing capacity only bring them very tiny welfare improvement.

phenomena simultaneously, that is, the contemporaneous low covariance between aggregate consumption and equity returns[5] and high equity premium if given the variability of exogenous shocks to asset returns. As documented in Mehra and Prescott (1985), Campbell (1999), and others, in quarterly or annual US data the average equity premium is very high, while real consumption growth is very smooth. Hence, the standard CCAPM can not rationalize the high equity premium by using the contemporaneous consumption risk. Following the same procedure used in the PIH model, we can solve an otherwise standard CCAPM model with RI approximately[6] and then evaluate how RI affects asset returns and portfolio choice by changing consumption risk. Parker (2001, 2003) and Parker and Jullard (2004) evaluated the central insight of the CCAPM but allowed the *possibilities* that households do not instantaneously and completely adjust consumption to the innovations to financial wealth. They present and review evidence that the ultimate consumption risk is a better measure for consumption risk than the contemporaneous consumption risk. Hence, incorporating RI into the canonical CCAPM model can be one of the *possibilities*. First, we show that incorporating RI into this simple CCAPM model can generate low volatility of aggregate consumption and low contemporaneous covariance between aggregate consumption growth and equity returns. Second, it is shown that if we measure the riskiness of an asset by its ultimate impacts on consumption instead of contemporaneous impacts, the model could generate high risk premium, which is around a factor of $\frac{1}{\theta}$ larger than[7] that if using contemporaneous consumption risk. The intuition is that in the RI model the contemporaneous covariance of consumption growth and asset returns understates the risk since consumption only adjusts slowly with respect to innovations to returns. Furthermore, we also show that RI would reduce the optimal allocation in risky asset if consumption risk is measured correctly. Finally, we introduce labor income risk into the CCAPM model and show that RI reduces the hedging demand of risky asset in the presence of uninsurable labor income risk.

Finally, we also consider the extension to the risk-sensitive PIH model. In the first two models, the LQ or approximate LQ specification (CRRA specification) implies that certainty equivalence holds and thus the nature of income risks does not affect the consumption function. Consequently, there is no precautionary saving motive in such models. However, as documented by numerous empirical studies, precautionary motive is very important for wealth accumulation and aggregate savings. Therefore, it is worthwhile to investigate the relationship among RI, risk aversion, and income uncertainty. Hence, in the third model, we provide such a framework which is called "the risk-sensitive LQG model" or "the Linear Exponential Quadratic Gaussian (henceforth, LEQG)"[8], and introduce increased risk aversion[9] and hence precautionary savings motive by an exponential transformation of the standard LQG problem[10]. Following Whittle (1981, 1990) and Hansen and Sargent (1995), we solve for a closed-form solution and examine the interaction among income uncertainty, risk aversion, and RI. Moreover, we also discuss some implications for consumption and saving dynamics, as well as welfare losses due to RI.

Recently, there have been some papers that incorporate information frictions into a variety of theoretical

[5] Also, the smoothness of aggregate consumption.

[6] We use "approximately" here because the CCAPM model has to be locally approximated to fit it into the same Gaussian-error framework as in the LQ setup.

[7] θ is the observing weight in the Kalman equation and it is equal to $1-1/\exp(2\kappa)$, where κ is the amount of channel capacity measured by 'nats' or 'bits'. We will discuss them in details in section 2.

[8] We will use these terms interchangeably throughout this paper.

[9] Acutually, this model is a special case of Epstein-Zin's non-expected utility and then can disentangle the elasticity of intertemporal substitution and risk aversion. Consequently, we can hold EIS constant and examine the interaction between increased risk aversion, RI, consumption, and savings.

[10] See detailed discussions for the risk-sensitive LQG models, see Whittle (1981, 1982, 1990, and 1996). For the applications of this framework in economics, see Hansen and Sargent (1995), Hansen, Sargent, and Tallarini (1999), Tallarini (2001), and others.

models, and that explore how imperfect information acquiring and processing affect the optimal decision rules of consumers, firms, and investors, as well as its implications for equilibrium outcomes. For example, Woodford (2001), Ball et. al. (2003), Adam (2004), and Gumbau-Brisa (2004) analyzed the effects of imperfect common knowledge on monetary policy and inflation dynamics; Peng and Xiong (2001) discussed how information capacity constraints affect the dynamics of asset return volatility; Moscarini (2003) derived optimal time-dependent adjustment rules from the information constraints; Luo (2004) examined the effects of RI on the amplification and propagation of aggregate shocks, asset returns, and the welfare costs of business cycles in a stochastic growth model; and Peng (2004) explored the effects of information constraints on the equilibrium asset price dynamics and consumption behavior in a continuous-time model. A number of recent papers have also explored the potential of inattentiveness from another attack line. For example, Gabaix and Laibson (2001) assumed that investors update their portfolio decisions infrequently and show that this can better explain the risk premium puzzle; Mankiw and Reis (2002) examined the effects of inattentiveness of firms on the dynamics of output and inflation; and Reis (2003) derived the optimal decision rules for inattentive consumers and then discussed the implications of inattentiveness for individual and aggregate consumption behaviors[11].

This paper is organized as follows. In Second 2, we solve a LQG PIH model with RI explicitly and discuss some implications of RI for the dynamics of consumption, as well as welfare. In section 3, a CCAPM model with RI is solved and its implications for asset returns and portfolio choices are discussed. Section 4 describe the extension to the risk-sensitive LQG PIH model with RI and examine the effects of increased risk aversion and RI on consumption dynamics and precautionary savings. In section 5, we conclude with a summary of our findings and some further extensions. Section 6 is appendix.

2. The LQ PIH Model with Rational Inattention

The first model discussed in this paper is a simple LQG PIH model with RI. Consider a representative consumer who maximizes his utility function subject to both the usual flow budget constraint and the information processing constraints that will be specified later. The decision problem of this consumer can be characterized by the following optimization problem

$$\max_{C_t} E_0 \sum_{t=0}^{\infty} \beta^t u(C_t)$$

subject to:

$$W_t = R(W_{t-1} - C_{t-1}) + Y_t, \tag{2.1}$$

$$S_{t+1}|\mathcal{I}_{t+1} \sim \mathcal{D}_{t+1} \tag{2.2}$$

$$S_t|\mathcal{I}_t \sim \mathcal{D}_t \tag{2.3}$$

and the requirement that the rate of information flow at $t+1$ implicit in the specification of the distributions, $\mathcal{D}_t$ and $\mathcal{D}_{t+1}$ (Later we will show that both are normal distribution), be less than channel capacity, where S_t is the state variable that can be W_t if income process is i.i.d., $[W_t, Y_t]$ for more general income process, or

[11]'RI' modeled in Sims (2003) and this paper is based on Shannon channel capacity, while Reis modeled 'inattentiveness' by assuming and justifying the existence of decision costs that induces agents to only infrequently update their decisions. As shown below, although the two assumptions are based on distinct mechanisms, they may generate similar aggregate dynamics.

a linear combination of W_t and Y_t, $\mathcal{I}_t$ is the information available at time t, $u(C_t) = C_t - \frac{1}{2}C_t^2$ is the utility function, (2.1) is the usual flow budget constraint for this consumer, C_t is individual consumption, W_t is individual wealth, β is the discount factor, R is the constant gross interest rate (for simplicity, we impose $\beta R = 1$), Y_t is individual income process that may be composed of several components, and the innovations to income process are Gaussian (they will be specified in section 2.1). The expectation is formed under the assumption that $\{C_t\}_0^\infty$ are chosen under the information processing constraints.

The consumer faces the information processing constraints in the sense that he only devotes limited channel capacity in observing the relevant state that includes both individual state and aggregate state. Following Sims (2003), we also use the concept of entropy from information theory to characterize the rate of information flow and then use the reduction in entropy as a measure for information[12]. With the finite capacity, the consumer will choose a signal that reduces the uncertainty of the state. Formally, this idea can be described by the following information constraint

$$\mathcal{H}(S_{t+1}|\mathcal{I}_t) - \mathcal{H}(S_{t+1}|\mathcal{I}_{t+1}) \leq \kappa \tag{2.4}$$

where κ is the consumer's information channel capacity[13] that imposes an upper bound on the amount of information that can be transmitted via the channel[14], $\mathcal{H}(S_{t+1}|\mathcal{I}_t)$ denotes the entropy of the state prior to observing the new signal at $t+1$, and $\mathcal{H}(S_{t+1}|\mathcal{I}_{t+1})$ the entropy after observing the new signal. Given the LQG framework, we suppose that the initial state S_{-1} given $\mathcal{I}_{-1}$ is distributed $\mathcal{D}_{-1} = N(\widehat{S}_{-1}, \Sigma_{-1})$[15].Then by induction (based on the updating recursions of the conditional mean and variance), we have $S_t|\mathcal{I}_t \sim N(\widehat{S}_t, \Sigma_t)$. Therefore, (2.4) can be rewritten as

$$\log_2 |\Psi_t| - \log_2 |\Sigma_{t+1}| \leq 2\kappa \tag{2.5}$$

where $\Sigma_{t+1} = Var_{t+1}[S_{t+1}]$ and $\Psi_t = Var_t[S_{t+1}]$ are the posterior and the prior variance-covariance matrices of the state vector.Note that here we use the fact that the entropy of a Gaussian random variable is equal to half of its logarithm variance plus some constant term. Sims (2003) shows that in the one-dimensional state case (e.g., income is i.i.d.), this information constraint completes the characterization of the optimization problem and everything can be solved analytically, while for the multivariate state case, we need another information constraint, that is, $\Psi_t \succeq \Sigma_{t+1}$. This constraint is used to rule out the possibility that $\Psi_t - \Sigma_{t+1}$ might not be positive semi-definite because information flow cannot be kept low by forgetting some existing information, and trading this off for increased precision about other elements in the state vector. As will be shown below, in the PIH model with any plausible income process we are able to reduce any multivariable state case to the univariate state case. Hence, (2.5) is enough for our analysis. We then assume that with a finite capacity κ the optimizing consumer will choose a signal that reduces the conditional variance of S_{t+1} by a maximum (limited) amount. Hence, in this case information imperfections emerge endogenously since

[12]Entropy is defined as a measure of the uncertainty about a random variable. See Shannon (1948) and Cover and Thomas (1991) for details.

[13]If the base for logarithms is 2, the unit used to measure information flow is called 'bits', and if we use the natural logarithm e, the unit is called 'nats'. Hence, 1 nat is equal to $\log_2 e = 1.433$ bits.

[14]We can regard it as a simple technology parameter. It can be modeled exogeneously or endogeneously. Note that information channels only limit the overall amount of information flowing to agents, while the agents might also need to allocate the capacity efficiently across different elements in the state vector, that is, decide which element of the state vector to observe with what precision.

[15]As shown in Sims (2003), in the static problem, minimizing of LQ losses subject to information processing constraint implies that the conditional distribution of S_{-1} is normal.

consumers are assumed to choose the nature of signal and then receive information about the state through the information channel that is contaminated with the endogenous noise. As a result, this idiosyncratic noise could generate private information about the common shock (e.g., the business-cycle shocks) hitting the economy.

It is straightforward to show that in this univariate case the equation (2.5) has a steady state Σ. In steady state the consumer is assumed to behave as if observing a noisy measurement which is $S^*_{t+1} = S_{t+1} + \xi_{t+1}$, where ξ_{t+1} is the endogenous noise and its variance $\Lambda_t = Var_t[\xi_{t+1}]$ is determined by the usual updating formula of the variance of a Gaussian distribution based on a linear observation:

$$\Sigma_{t+1} = \Psi_t - \Psi_t(\Psi_t + \Lambda_t)^{-1}\Psi_t. \tag{2.6}$$

Note that in steady state $\Sigma = \Psi - \Psi(\Psi + \Lambda)^{-1}\Psi$,which can be solved as $\Lambda = (\Sigma^{-1} - \Psi^{-1})^{-1}$.

We can also write the updating recursions of the conditional mean $\widehat{S}_t$, that is, the Kalman filtering, as follows

$$\widehat{S}_t = F(\widehat{S}_{t-1}, S_t + \xi_t) \tag{2.7}$$

where the linear function F is determined by the nature of income process and then the definition of the relevant state S_t, as well as the value of Λ and Σ. Once we specify them, we can easily obtain this function form.

The fact that optimal control does not affect state estimation procedures under LQ assumptions and that optimal controls are certainty equivalent versions of optimal deterministic controls under LQG assumptions is referred to as *the separation principle*. This means that optimization of state estimation and control can be decoupled under these assumptions. See Whittle (1982, 1996) for detailed discussions. Hence, in our RI model once we derive the full-information linear optimal consumption rule $C_t = G(S_t)$ and the optimal state estimation $\widehat{S}_t$ from (2.6) and (2.7), we can obtain the optimal consumption rule in the RI model: $C_t = G(\widehat{S}_t)$. This optimal consumption rule, the flow budget constraint (2.1), and the Kalman filter equation (2.7) constitute a dynamic system that can characterize our RI economy completely.

To investigate the implications of RI for consumption dynamics, it is helpful to review the main predictions from the standard full-information PIH model. The following consumption function represents the permanent income theory:

$$C_t = PI_t = \frac{R-1}{R}[W_t + \sum_{j=1}^{\infty} R^{-j} E_t Y_{t+j}] \tag{2.8}$$

where PI_t is permanent income defined as the annuity value of the consumer's net worth including the present discounted value of expected future labor income as well as non-human wealth. Hence, consumption is determined by permanent income instead of current income. Furthermore, the first difference of the above consumption function can be written as

$$\Delta C_t = \Delta PI_t = (R-1)\sum_{j=1}^{\infty} R^{-j}(E_t - E_{t-1})Y_{t+j-1} \tag{2.9}$$

Hence, the change in consumption is equal to the change in permanent income and depends only on the revision in expectations of future labor income. In other words, under the PIH, consumption process is a martingale, and the change in consumption depends neither on the past history of labor income nor on anticipated changes in labor income. This is the well-known result of Hall (1978). However, a number of empirical studies have shown that aggregate consumption does not behave like a martingale and there is positive serial correlation on the change in aggregate consumption. Furthermore, excess smoothness of consumption to income innovations and excess sensitivity to lagged income are also found in the data. As argued in Deaton (1992), excess smoothness and excess sensitivity are not different, but two aspects of same phenomenon. If changes in consumption are orthogonal to lagged information, then they must be equal to changes in permanent income, and they cannot be too smooth. See Deaton (1992) for a recent review on these issues. Although these empirical findings have led to a large number of potential explanations, here we offer RI as an alternative potential explanation for these phenomena.

It is assumed that the consumer in the PIH model with RI cannot observe the state(s) perfectly because observing the state perfectly requires an unlimited rate of information transferred, which is at odds with reality since ordinary people only have a limited channel capacity. Consequently, noise emerges endogenously and the nature of noise is determined endogeneously. Furthermore, in most multivariate state cases people need to determine how to allocate their limited capacity optimally over a variety of states[16]. Another difficulty emerging from imperfect observations is that the *effective* state in this type of optimal control problems is not the traditional state variable, but the so-called *information* state which is defined as the distribution of the state variable conditional on the information set available at time t, $\mathcal{I}_t$. In other words, it makes an expansion of state space to the space of distributions on the state vector S. This means a substantial increase in dimensionality when we use dynamic programming to characterize this optimization problem. Consequently, it makes the model with information processing constraints very difficult to solve, which is called "the curse of dimensionality" in the literature[17]. Fortunately, the Linear-Quadratic-Gaussian assumptions simplify the dimensionality problem greatly because the conditional distribution of the state vector is Gaussian, that is., the first two moments, the conditional mean $\widehat{S}_t$ and the conditional covariance matrix Σ_t, are enough to characterize the *effective* state. Actually, in some cases the problem could be even simpler in that Σ_t evolutes over time by determinant rules which are independent of optimal policy or Σ_t converges to a constant matrix in the steady state.

2.1. Consumption Dynamics

In this subsection, we will first derive the expression of the change in consumption in terms of income innovations and noise, and then examine the implication of RI for both individual consumption dynamics and aggregate consumption dynamics.

2.1.1. Individual Consumption Behavior

We assume that income process is composed by two components: one is permanent and the other is transitory i.i.d.. This decomposition of income into permanent and transitory components is widely adopted in the

[16] Sims (2003) and Peng (2004) provide various criterions to determine the allocation of capacity efficiently in the multivariate case.

[17] Another set of models which also have the problem of "the curse of dimensionality" is the heterogeneous-agent model with both idiosyncratic shock and aggregate shock. In those models, the measure becomes a state variable and thus the model have infinite states. Krusell and Smith (1998) and Rios-Rull (1999) solved this problem by approximating the measure by using a small number of constituent moments and these moments can be used to predict future prices. And it is shown in their papers that using aggregate (average) capital stock as an approximation for the distribution is enough to solve the models numerically with very small errors.

literature[18]. For example, Moffitt and Gottschalk (1995) found that an ARMA $(1,1)$ process with the autocorrelation is close to 0 can characterize the transitory component in income best, and Storesletten et. al (2001) estimated individual income process and found that they cannot reject the hypothesis that the persistent income shock is permanent. Formally,

$$Y_t = Y_t^P + Y_t^T \tag{2.10}$$

where

$$Y_t^P = Y_{t-1}^P + \varepsilon_t \text{ and } Y_t^T - \overline{Y}^T = \eta_t \tag{2.11}$$

where Y_t^P is permanent component and Y_t^T is transitory component. It is assumed that the consumer can separately identify the two components, so that consumption can respond to each. Permanent innovations to income are consolidated into the income base and it is possible that the consumer can recognize which is which. All innovations are assumed to have zero mean. Furthermore, we assume that the all shocks are uncorrelated over time and uncorrelated with each other and the variance of ε_t and η_t are ω^2 and ν^2, respectively.

Substituting the specified income process into the consumption function (2.8) gives

$$C_t = \frac{R-1}{R}(W_t + \frac{1}{R-1}Y_t^P) + \frac{1}{R}\overline{Y}^T \tag{2.12}$$

Sims (2003) proposed a numerical procedure to solve the multivariate state model with RI, here we will show that there exist analytical solutions for this kind of PIH model with RI given the ARMA(1,1) income process. From the above consumption function, it is obvious that what determines consumption is a *linear combination* of two individual state variables, W_t and Y_t^P, so we can regard this linear combination as a new state variable, that is,

$$M_t = W_t + \frac{1}{R-1}Y_t^P$$

and the consumption function becomes $C_t = \frac{R-1}{R}M_t + \frac{1}{R}\overline{Y}^T$. Furthermore, it is straightforward to prove that the dynamics of this new state is consistent with the original flow budget constraint because

$$W_{t+1} + \frac{1}{R-1}Y_{t+1}^P = R[(W_t + \frac{1}{R-1}Y_t^P) - C_t] + \frac{R}{R-1}\varepsilon_{t+1} + \eta_{t+1} + \overline{Y}^T,$$

which can be rewritten in terms of M_t :

$$M_{t+1} = R(M_t - C_t) + \zeta_{t+1} + \overline{Y}^T \tag{2.13}$$

where $\zeta_{t+1} = \frac{R}{R-1}\varepsilon_{t+1} + \eta_{t+1}$.

Introducing RI in this model yields the following consumption function since certainty equivalence holds in the LQ model

$$C_t = \frac{R-1}{R}\widehat{M}_t + \frac{1}{R}\overline{Y}^T \tag{2.14}$$

[18] For simplicity, here we do not separate aggregate (common) components from idiosyncratic components in both permanent and transitory components. In the next subsections where we aggregate over consumers, we assume that aggregation smooths away individual idiosyncracies and only aggregate components left.

where $\widehat{M_t} = E[M_t|\mathcal{I}_t]$.

Given the model specification, the conditional distribution of M_t given information at time t, $\mathcal{I}_t$ is $N(\widehat{M_t}, \sigma^2_{M,t})$[19]. Then the new flow budget constraint implies that

$$E_t[M_{t+1}] = \widehat{M_t} \text{ and } Var_t[M_{t+1}] = R^2\sigma^2_{M,t} + (\frac{R}{R-1})^2\omega^2 + \nu^2$$

Now we can characterize the nature of the endogenous noise and then the evolution of the effective state $\widehat{M_t}$ as follows. First, the first information constraint (2.5) in this model implies

$$\kappa = \frac{1}{2}[\log_2((\frac{R}{R-1})^2\omega^2 + \nu^2 + R^2\sigma^2_{M,t}) - \log_2(\sigma^2_{M,t+1})]$$

and it has a steady state

$$\overline{\sigma}^2_M = \frac{(\frac{R}{R-1})^2\omega^2 + \nu^2}{\exp(2\kappa) - R^2} \tag{2.15}$$

The consumer behaves as if observing a noisy measurement $M^*_{t+1} = M_{t+1} + \xi_{M,t+1}$ in steady state and the noise $\xi_{M,t+1}$ is independent with $Var[\xi_{M,t+1}] = \frac{[(\frac{R}{R-1})^2\omega^2+\nu^2+R^2\overline{\sigma}^2_M]\overline{\sigma}^2_M}{(\frac{R}{R-1})^2\omega^2+\nu^2+(R^2-1)\overline{\sigma}^2_M}$. Finally, we obtain the following recursive Kalman equation about $\widehat{M_t}$

$$\widehat{M}_{t+1} = (1-\theta)\widehat{M_t} + \theta(M_{t+1} + \xi_{M,t+1}) \tag{2.16}$$

where $\theta = \overline{\sigma}^2_M / Var[\xi_{M,t+1}] = 1 - 1/\exp(2\kappa)$ is the optimal weight on observation.Note that when $\theta = 1$ (i.e., channel capacity $\kappa = +\infty$), the RI model reduces to the standard PIH model.

Equation (2.10), (2.13), (2.16), and (2.14) form a dynamic system. Based on them, we can obtain the impulse response functions of consumption. Figure 1 below shows the responses of consumption with respect to two components in income and one error shock, with channel capacity 1. The two horizontal lines accompanied with the the responses to both permanent and transitory shocks represent the levels of the flat responses of consumption in the absence of information capacity constraints. It is obvious in this figure that the impulse responses of consumption to two income shocks exhibit delay and gradually reach a flat asymptote. As pointed out in Sims (2003), this flat asymptote is above the horizontal line representing the responses in the PIH model without RI because consumption does not react fully to income shocks initially and then the share of undetected income shocks goes to savings and accumulates interest before the consumer digests the shocks fully. The property that consumption responds gradually and with delay to the shocks to wealth is an important potential for explaining not only individual and aggregate consumption behavior but also other observed business cycle and asset returns phenomena. For example, many monetary DSGE models imply that both real spending and inflation jump immediately in response to shocks, in contradiction to a host of empirical evidence showing that both price and real variables exhibit gradual and hump-shaped responses to real and monetary shocks. Hence, once RI mechanism can be introduced into these models in a reasonable way, the model will fit the data better[20].

[19]It follows from minimizing the quadratic loss function subject to the information constraint.

[20]In Luo (2004), RI is introduced into an otherwise standard RBC model with stochastic growth and it is shown that RI can be an important propagation mechanism for aggregate shocks in the sense that main real macroeconomic variables exhibit gradual and hump-shaped responses to aggregate shocks and the autocorrelation function of output growth is significantly positive in the first several periods.

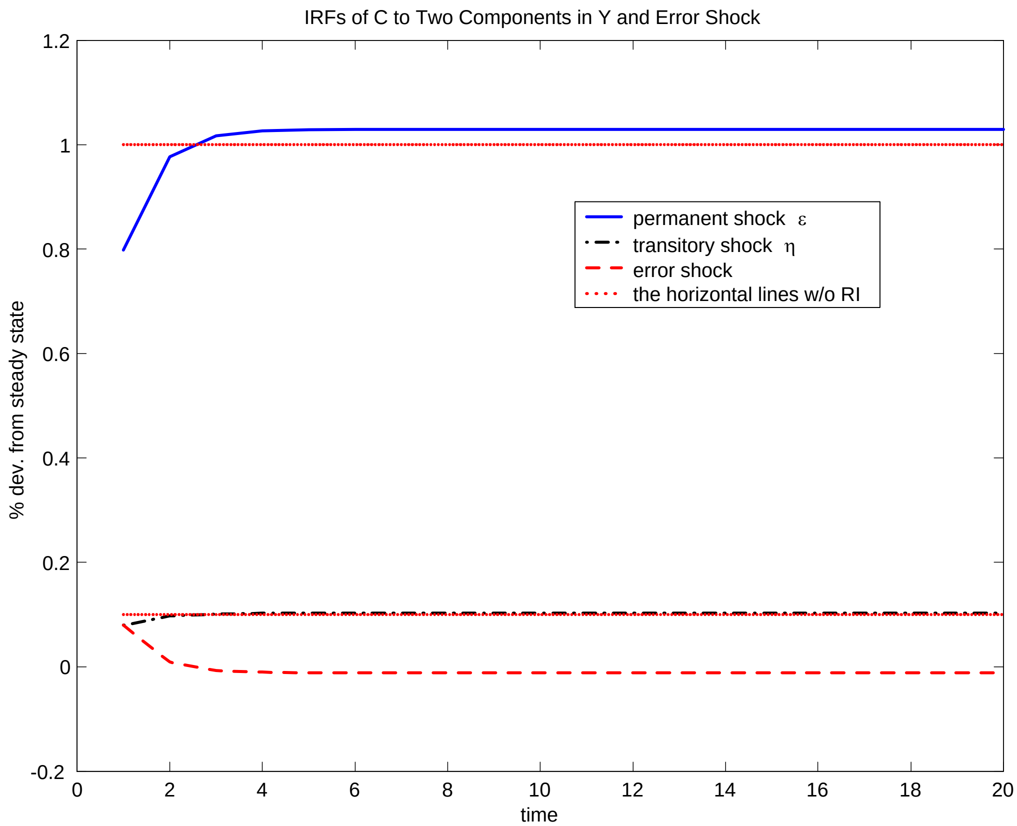

Figure 1

Alternatively, to gain some further insights about the effects of RI on the dynamics of consumption and savings and to make the model testable, it is useful to derive some analytical expressions of the change in consumption or savings in terms of income shocks. Fortunately, given our model specifications, we do have such analytical expressions. Combining equation (2.14) and (2.16) yields the following consumption evolution equation

$$C_t = (1-\theta)C_{t-1} + [1-(1-\theta)](H_1 M_t + H_0) + \theta H_1 \xi_t \tag{2.17}$$

where $H_1 = \frac{R-1}{R}$, $H_0 = \frac{1}{R}\overline{Y}^T$, and $H_1 M_t + H_0$ is the permanent income. We can read from the above equation that current consumption can be expressed as a weighted average of past consumption and permanent income; the more the inattention, the more weight will be assigned to the past consumption.

Substituting the specified income process into expression (2.9), the change in consumption in the absence of RI, yields

$$\Delta C_t = H_1 \varepsilon_t + \eta_t \tag{2.18}$$

that is, the change in consumption is not sensitive to past income shocks and jumps immediately in response to income shock. By contrast, the change in consumption in the model with RI can be written as

$$\Delta C_t = H_1 \theta \zeta_t + (R-1)\theta \frac{(1-\theta)\zeta_{t-1} - \theta \xi_{t-1}}{1-(1-\theta)R \cdot L} + \theta H_1 \xi_t \tag{2.19}$$

where we use the fact that $M_t - \widehat{M}_t = \frac{(1-\theta)\zeta_t - \theta\xi_t}{1-(1-\theta)R\cdot L}$. Note that given the expression of θ and $\overline{\sigma}_M^2 > 0$, i.e., $\exp(2\kappa) - R^2 > 0$, it is straightforward to prove that $(1-\theta)R < 1$. This expression reflects how the change in consumption responds to all current and past shocks to income, and it is also consistent with what we see from Figure 1. We define $I(i)$ as the coefficients attached to all income shocks ζ_{t-i} and $J(i)$ as the coefficients attached to endogenous noise ξ_{t-i}, and

$$I(i)=\begin{cases} H_1\theta \text{ if } i=0 \\ (R-1)\theta(1-\theta)[(1-\theta)R]^{i-1} \text{ if } i\geq 1 \end{cases};J(i)=\begin{cases} \theta H_1 \text{ if } i=0 \\ -(R-1)\theta^2[(1-\theta)R]^{i-1} \text{ if } i\geq 1 \end{cases}$$

Equation (2.19) shows that the growth of consumption is a MA(∞) process with decreasing coefficients and it implies that consumption adjust slowly and gradually to income shocks, with reactions that build up over time. Note that when $\theta=1$, i.e.,$\kappa=\infty$, the above expression reduces to (2.18). In other words, this expression of the change in individual consumption implies that individual consumption should be sensitive to both current and past income shocks, as well as endogenous noise ξ. Parker (1999) and Souleles (1999) provided the evidence that consumption responses to past news on after-tax income. Parker (1999) examined the impact of Social Security tax withholding, and Souleles (1999) considers income tax refunds. The news in both cases are unpredictable and consumption reacts to the news with delay.

2.1.2. Aggregation and Implications for Consumption Smoothness

In this subsection, we will discuss the effects of RI on aggregate consumption dynamics. Specifically, we will examine if RI can be an alternative potential explanation for the excess sensitivity puzzle and the excess smoothness puzzle in the consumption literature. Consider an economy composed of a continuum of consumers (measure 1). They are distinguished by their channel capacity κ_i or the optimal observation weight θ_i since it is determined solely by κ_i, where $i=1,2,....$ And we assume λ_i is the fraction of the individuals with channel capacity κ_i. We denote the consumption expenditure of type i individuals and the average/aggregate consumption at period t by $C_{i,t}$ and C_t, respectively. Hence, the growth of aggregate consumption can be defined as

$$\Delta C_t=\int_i \lambda_i \Delta C_{i,t} di$$

From the preceding subsection, we know that

$$\Delta C_{i,t}=[H_1\theta_i\zeta_{i,t}+(R-1)\theta_i\frac{(1-\theta_i)\zeta_{i,t-1}}{1-(1-\theta_i)R\cdot L}]+[H_1\theta\xi_{i,t}-(R-1)\theta_i^2\frac{\xi_{i,t-1}}{1-(1-\theta_i)R\cdot L}], \tag{2.20}$$

where the subscript i in the shock terms means that consumer i is hit by the shocks which may be common or idiosyncratic. Aggregating it over all consumers yields

$$\begin{aligned}\Delta C_t=&[H_1(\int_i\theta_i\chi_i di)\zeta_t+H_1\sum_{j=1}^{\infty}(\int_i\theta_i(1-\theta_i)^jR^j\chi_i di)\zeta_{t-j}] \\ &+[H_1(\int_i\theta_i\varpi_i di)\xi_t-(R-1)\sum_{j=1}^{\infty}(\int_i\theta_i^2(1-\theta_i)^{j-1}R^{j-1}\varpi_i di)\xi_{t-j}]\end{aligned} \tag{2.21}$$

where ζ_t is the average level of income innovations, χ_i is the ratio of income innovation of type i individuals to the average income innovation. Similarly, ξ_t is the average level of error terms and ϖ_i is the ratio of error of type i individuals to the average level of error. Note that

$$\zeta_t = \int_i \lambda_i \zeta_{i,t}\, di,\ \chi_i = \frac{\lambda_i \zeta_{i,t}}{\zeta_t},\ \xi_t = \int_i \lambda_i \xi_{i,t} di,\ \varpi_i = \frac{\lambda_i \xi_{i,t}}{\xi_t}$$

For simplicity, we assume that consumers have identical channel capacity $\kappa_i = \kappa$, and then the above expression for the change in aggregate consumption can be simplified to

$$\Delta C_t = [H_1 \theta \zeta_t + H_1 \sum_{j=1}^{\infty} \theta (1-\theta)^j R^j \zeta_{t-j}] + [H_1 \theta \xi_t - (R-1) \sum_{j=1}^{\infty} \theta^2 (1-\theta)^{j-1} R^{j-1} \xi_{t-j}] \tag{2.22}$$

where $\zeta_t = \int_i \zeta_{i,t}\, di$ and $\xi_t = \int_i \xi_{i,t} di$. Hence, if the average inattention in the economy is not high, aggregate consumption also display slow adjustment to shocks. Reis (2003) provided a simple analysis of the response of aggregate consumption by estimating a structure vector autoregression (VAR) on consumption and income growth. Figure 2 reported in his paper showed aggregate consumption has a delayed adjustment to the shock. Furthermore, the figure shows that the adjustment is not very delayed: most of the adjustment is finished with one year, which is consistent with Figure 1 above where the impulse response reaches the flat asymtote.

To evaluate the effects of RI on aggregate consumption, we consider two cases. In the first extreme case, we assume that the idiosyncratic endogenous noise due to RI will vanish when aggregate over all consumers, and in the second case, we assume that the noises are totally common[21].

Case 1: For the first case, the terms in the second bracket in equation (2.22) vanish, so (2.22) simplifies to

$$\Delta C_t = H_1 \theta \zeta_t + H_1 \theta \sum_{j=1}^{\infty} (1-\theta)^j R^j \zeta_{t-j} \tag{2.23}$$

where for simplicity, we assume that all idiosyncratic components in income process are cancelled out and then ζ_t only includes aggregate component $Y_t^{P,a}$ which we assume is permanent and included in the component Y_t^P specified above

$$Y_t^{P,a} = Y_{t-1}^{P,a} + \varepsilon_t^a \tag{2.24}$$

where ε_t^a is i.i.d. with mean 0 and variance ω_a^2.

Hence, in the standard PIH model, $\Delta C_t = H_1 \zeta_t = \varepsilon_t^a$. In other words, the volatility of consumption growth is the same as the volatility of the innovation to income. As in the consumption literature, we can compare our model with other PIH models by looking at their different predictions for the shape of the normalized power spectrum of consumption growth equation (2.23). Following Gali (1991), Deaton (1992) and Reis (2003), we can derive the excess smoothness ratio as $\mu = 1/\sqrt{2\pi f_{\Delta C}(0)}$ that is just the ratio of standard derivation of the change in consumption to the standard deviation of the change in permanent income. Hence, the standard PIH model predicts that $\mu = 1$, while $\mu < 1$ implies that consumption is

[21] Of course, the assumption is not realistic, but as argued in Sims (2003), to some extent, people rely on common sources of coded information and thus there is a considerable common component in individuals' reactions to the shocks to wealth.

excessively smooth relative to income. Note the formula for the power spectrum is

$$h_{\Delta C}(\varpi) = \frac{1}{2\pi}[\gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k \cos(\varpi k)] \tag{2.25}$$

Given the equation of changes in consumption (2.23), the corresponding autocovariance function is

$$\gamma_k = \sum_{i=0}^{\infty}[\omega^2 I(k+i)I(i)] \tag{2.26}$$

Substituting the expression of γ_k in (2.25) and normalizing it by γ_0 yields

$$f_{\Delta C}(\varpi) = \frac{1}{2\pi}[1 + 2\frac{\sum_{k=1}^{\infty}\sum_{i=0}^{\infty}[\omega^2 I(k+i)I(i)]\cos(\varpi k)}{\sum_{i=0}^{\infty}[\omega^2 I(i)^2]}] \tag{2.27}$$

Evaluating it at $\varpi = 0$ yields

$$\mu = \sqrt{\frac{\sum_{i=0}^{\infty} I(i)^2}{[\sum_{i=0}^{\infty} I(i)]^2}} \tag{2.28}$$

Since $I(i) > 0$ for $i = 1, ...\infty$ in our RI model, the values of μ is always less than 1. In other words, RI generates excess smoothness in consumption relative to income. Deaton (1992) pointed out that in United States data, the hypothesis that output growth is positively serially correlated is difficult to reject statistically. If so, then it is a puzzle why consumption does not move more dramatically in response to output changes than it does in the data. More precisely, if income is difference-stationary, innovations in income generate changes in permanent income that is larger than the innovations, so that permanent income theory actually magnifies innovations in actual income. Figure 2 below shows the relationship between the channel capacity κ and the excess smoothness ratio μ, and it is obvious that the excess smoothness increases with information capacity.

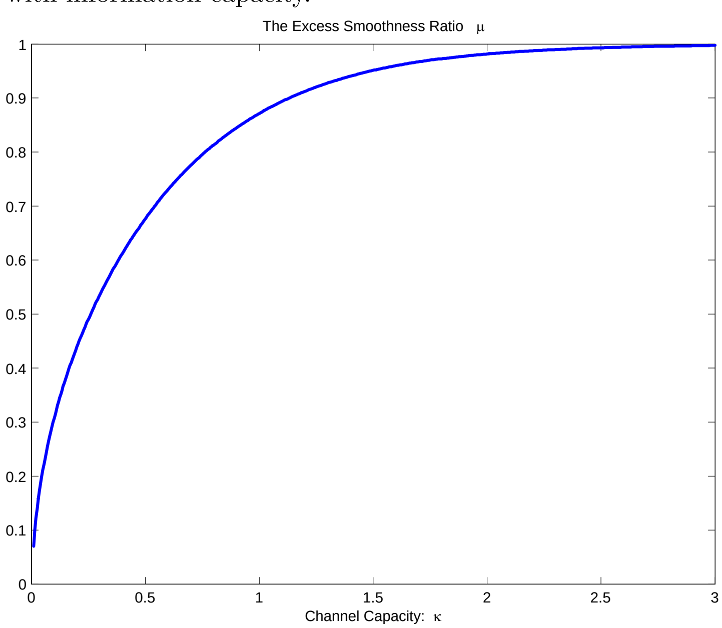

Figure 2

Case 2: In the second extreme case, $\xi_t = \xi_{i,t}$. Then given equation (2.22), the corresponding autocovariance function is now

$$\gamma_k = \sum_{i=0}^{\infty} [\omega^2 I(k+i) I(i) + \upsilon^2 J(k+i) J(i)] \tag{2.29}$$

where $\upsilon^2 = Var(\xi_t)$. Hence, the excess smoothness ratio can be written as

$$\mu = \sqrt{\frac{\sum_{i=0}^{\infty} [\omega^2 I(i)^2 + \upsilon^2 J(i)^2]}{[\sum_{i=0}^{\infty} \omega I(i)]^2 + [\sum_{i=0}^{\infty} \upsilon J(i)]^2}} \tag{2.30}$$

This case is more complicated than case 1 in that the entire series of the endogenous noise terms will affect changes in consumption. However, given the nice structure of the MA(∞) process and the property of noise, it is quick to evaluate the values of μ for different value of channel capacity. Given the above parameter values, we plot figure 2 that shows the relationship between the channel capacity κ and the excess smoothness ratio μ.

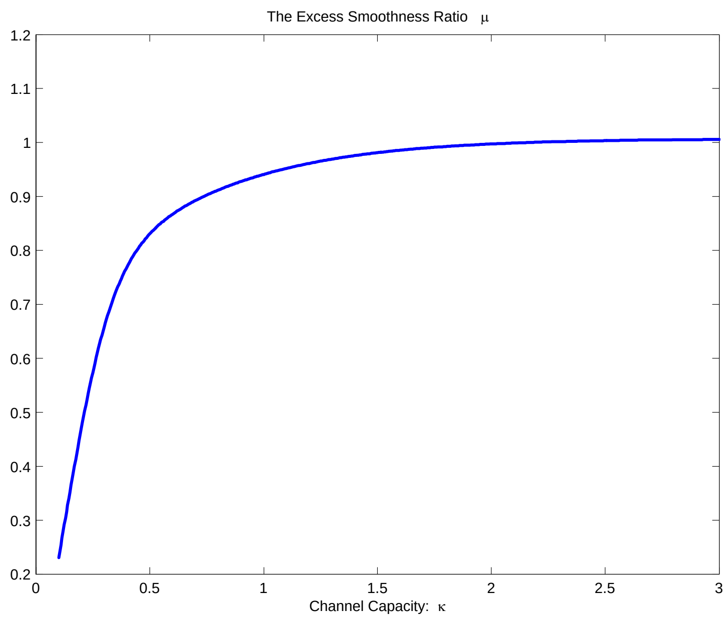

Figure 3

Proposition 1. *In the RI model, changes in consumption is both excess smoothness $\mu < 1$ and excess sensitivity $I(i) > 0$ for all i. Furthermore, the excess smoothness ratio is increasing with channel capacity and it gradually rises to $\mu = 1$ which is just predicted by the standard PIH model.*

2.1.3. Aggregate Stylized Facts, the Model Predictions, and Calibration

In this subsection, we will first report some stylized facts about income and consumption processes in aggregate data. This has two purposes. First, we can establish some simple time series aggregate income process. Combined it with the findings of the previous section allows us to obtain predictions from the RI model. Second, we will also report the results on aggregate consumption and use them to calibrate the main structure parameter in the RI model.

The following table is borrowed from Pischke (Table V and VI, Pischke (1995)). The income series refers to labor income and consumption expenditure includes nondurables and services. As in Deaton (1992), Pischke (1995) also estimated AR(1) income process by OLS and found that this specification fits data well.

In the table, it is clear that the regression coefficient of consumption changes on lagged income changes is around 0.11 and significant. In other words, the standard PIH model can not predict this feature in the data since consumption is martingale and does not depend on lagged income in the standard PIH model. Furthermore, the excess smoothness ratio in the table is only around 0.578 which is well below that predicted by the standard PIH model.

Table 1

Aggregate Stylized Facts on ΔY_t and ΔC_t (quarterly)

(standard errors in parentheses)

sample period	AR(1) coef.of ΔY_t	s.d. of income innovations	
1954 – 1990	0.288 (0.08)	\$50 (mean income is \$7000)	
sample period	coef. of ΔC_t on $\Delta Y_{t-1}(\widehat{\beta})$	AR(1) coef. of ΔC_t	excess smoothness ratio
1954 – 1990	0.110 (0.045)	0.200 (0.082)	0.578 (0.055)

We now report the predictions of the RI model and then use the reported stylized facts to calibrate the model structure parameter, κ. For simplicity, here we follow Pischke (1995) and assume that all consumers have identical income processes while each agent faces different realizations of this process, and this process is composed of two parts:

$$Y_t^i = Y_t + Y_t^{i,T}$$

where aggregate component Y_t is difference-stationary and positively serially correlated in first-differences[22]:

$$Y_{t+1} - Y_t = \rho(Y_t - Y_{t-1}) + \epsilon_{t+1},$$

where $\rho > 0$, and the second components $Y_t^{i,T}$ is individual specific and i.i.d.: $Y_t^{i,T} - \overline{Y}^{i,T} = \eta_t^i$. The idiosyncratic innovation η_t^i, will sum to zero when aggregating over all consumers.

Following the same procedure as above, we can easily derive the closed-form expression of ΔC_t in the RI model. The key step is defining a new state variable

$$M_t = W_t + \frac{R + (R-1)\rho}{(R-1)(R-\rho)} Y_t - \frac{\rho R}{(R-1)(R-\rho)} Y_{t-1}.$$

and then the flow budget constraint can be written in terms of M_t

$$M_{t+1} = R(M_t - C_t) + \zeta_{t+1} + \overline{Y}^T$$

where $\zeta_{t+1} = \frac{R^2}{(R-1)(R-\rho)}\epsilon_{t+1} + \eta_{t+1}^i$. Hence, we can express the change in aggregate consumption as follows

[22] Note that given this income process the full-information PIH model predicts that the change in aggregate consumption is

$$\Delta C_t = \Delta PI = \frac{R}{R-\rho}\epsilon_t > \epsilon_t,$$

that is, the excess smoothness ratio should be greater than 1 and thus the implied consumption is too volatile relative to current income.

$$\Delta C_t = H_1\theta\overline{\zeta}_t + \frac{(R-1)\theta(1-\theta)\overline{\zeta}_{t-1}}{1-(1-\theta)R\cdot L} \tag{2.31}$$

where $\overline{\zeta}_t = \frac{R^2}{(R-1)(R-\rho)}\epsilon_{t+1}$ and $H_1 = \frac{R-1}{R}$. This MA (∞) process can be rewritten as a AR(1) process

$$\Delta C_t = (1-\theta)R\Delta C_{t-1} + H_1\theta\overline{\zeta}_t \tag{2.32}$$

This equation has a number of interesting implications. The aggregate consumption in this equation is not a random walk as the standard PIH model predicts. Consumption now follows an AR(1) in first differences. This intuition behind this is simple. Suppose that an aggregate shock hits the model economy, the consumers can not digest the current state completely due to limited channel capacity. As a result, they will change their consumption but not by as much as the shock calls for. Since the shock is persistent, in the next period their income is higher than expected and they will increase their consumption further and so on.

We can then use these aggregate facts reported in Table 1 to calibrate κ. Specifically, the three main facts: the AR(1) coefficient of ΔC_t, the regression coefficient of ΔC_t on ΔY_{t-1}, and the excess smoothness ratio, will be used to calibrate κ. We first set $\rho = 0.288$ in aggregate income process. R is set to 1.01 quarterly. From equation (2.32) for the change in aggregate consumption, we can calibrate θ as follows: $(1-\theta)R = 0.2$, where the standard error of the estimated AR(1) coefficient of ΔC_t, 0.2, is 0.082. Hence, for the confidence interval around $[0.2-0.082, 0.2+0.082]$, the range of θ is around $[0.72, 0.88]$, that is, κ is in the range $[0.9 \text{ bits}, 1.5 \text{ bits}]$. Similarly, for the interval around $[0.2-2\cdot 0.082, 0.2+2\cdot 0.082]$, the range of θ is around $[0.64, 0.96]$, that is, κ is in the range $[0.7 \text{ bits}, 2.4 \text{ bits}]$.

We may also calibrate κ based on the fact about excess sensitivity, that is, the coefficient of ΔC_t on ΔY_{t-1}. Suppose that the econometrician estimates the following equation:

$$\Delta C_t = \alpha + \beta\Delta Y_{t-1} + e_{t+1} \tag{2.33}$$

If the data is generated by equation (2.33), the expected value of β is

$$\begin{aligned} \widehat{\beta} &= \frac{Cov[\Delta C_t, \Delta Y_{t-1}]}{Var[\Delta Y_{t-1}]} \\ &= \frac{Cov[\frac{(R/R-\rho)\theta\epsilon_t}{1-(1-\theta)R\cdot L}, \frac{\epsilon_{t-1}}{1-\rho\cdot L}]}{\omega_\epsilon^2/(1-\rho^2)} = \frac{R^2}{R-\rho}\frac{(1-\rho^2)\theta(1-\theta)}{1-R\rho(1-\theta)} \end{aligned} \tag{2.34}$$

and then we can use this expression to help to calibrate κ. Given $R = 1.01$ and $\rho = 0.288$, for the range $[\widehat{\beta}-0.045, \widehat{\beta}+0.045]$, the range of θ is around $[0.86, 0.95]$[23], that is, $\kappa \in [1.4 \text{ bits}, 2.1 \text{ bits}]$. Similarly, for the range $[\widehat{\beta}-2\cdot 0.045, \widehat{\beta}+2\cdot 0.045]$, the range of θ is around $[0.81, 0.98]$, that is, $\kappa \in [1.2 \text{ bits}, 2.8 \text{ bits}]$.

Finally, if we define the excess smoothness of aggregate consumption μ as the ratio of the standard deviation of consumption changes to the standard deviation of aggregate income innvoations, we have

$$\mu = \frac{\sigma(\Delta C_t)}{\sigma(\epsilon_t)} = \frac{\frac{R}{R-\rho}\theta}{\sqrt{1-[(1-\theta)R]^2}}. \tag{2.35}$$

[23] Actually, solving the quadratic equation for θ yields two values of θ. However, since the range $[0.05, 0.14]$ is at odds with the calibrated θ from the ΔC_t process, we eliminate it and only remain the range $[0.86, 0.95]$.

This ratio can also be used to calibrate κ. Specifically, for the range $[\mu - 0.055, \mu + 0.055]$, the range of θ is around $[0.2, 0.3]$. For the range $[\mu - 2 \cdot 0.055, \mu + 2 \cdot 0.055]$, the range of κ is around $[0.16, 0.4]$.

2.2. Welfare Effects of Information Processing Constraints

In this subsection, we examine the welfare effects of income fluctuations under RI. Specifically, we investigate this issue from two aspects: First, we calculate how much utility consumer will lose if the actual consumption path deviates from the first-best instantly adjusted consumption path due to RI. Second, what is the welfare costs of business cycle fluctuations in the RI economy. To examine the welfare effects of RI, we first need to establish a welfare criterion. As usual, we use the value function to represent individuals' intertemporal welfare.

In the standard full-information PIH model, we have the consumption rule: $C_t = H_0 + H_1 M_t$. And we can guess that the value function has the following form

$$V(M_t) = A_0 + A_1 M_t + A_2 M_t^2 \tag{2.36}$$

where A_0, A_1, and A_2 are undetermined coefficients. Following the standard procedure, we can pin down them as follows

$$A_1 = \frac{1}{1-\beta} H_1 (1 - H_0); A_2 = \frac{1}{1-\beta}(-\frac{1}{2} H_1^2); A_0 = \frac{1}{1-\beta}(H_0 - \frac{1}{2} H_0^2 + \beta A_2 \omega_\zeta^2)$$

By contrast, in the RI economy, we have the following Bellman equation

$$\widehat{V}(\widehat{M}_t) = \max_{C_t} E_t[u(C_t) + \beta \widehat{V}(\widehat{M}_{t+1})] \tag{2.37}$$

where the expectation is formed under the assumption that the current and future consumption are chosen under information processing constraints. Similarly, we can guess that $\widehat{V}(\widehat{M}_t) = B_0 + B_1 \widehat{M}_t + B_2 \widehat{M}_t^2$ and pin down the coefficients B_0, B_1, and B_2 as follows[24]:

$$B_1 = \frac{1}{1-\beta} H_1 (1 - H_0); B_2 = \frac{1}{1-\beta}(-\frac{1}{2} H_1^2); B_0 = \frac{1}{1-\beta}[H_0 - \frac{1}{2} H_0^2 + \beta B_2 ((R^2 - 1)\sigma_M^2 + \omega_\zeta^2)]$$

where $\omega_\zeta^2 = (\frac{R}{R-1})^2 \omega^2 + \nu^2$ and $\sigma_M^2 = \frac{\omega_\zeta^2}{\exp(2\kappa) - R^2}$. Note that $A_1 = B_1$ and $A_2 = B_2$.

Since information processing constraints cannot help in individual's optimization, the average welfare difference between the two economies should be greater than 0. One way to evaluate the effect of RI on welfare is to compute the difference of the two value functions around[25] $\overline{M}_t = E_t[M_{t+1}] = E_t[\widehat{M}_{t+1}]$:

$$\Delta \overline{V} = V(\overline{M}_t) - \widehat{V}(\overline{M}_t) = \frac{1}{2}(\frac{R^2 - 1}{R})\sigma_M^2. \tag{2.38}$$

Note that welfare loss due to RI, ΔV, converge to 0 when channel capacity κ increases to ∞.

[24] When pinning down the undetermined coefficients, we adopt the consumption rule in the RI model, $C_t = H_0 + H_1 \widehat{M}_t$.

[25] Since both M_{t+1} and $\widehat{M}_{t+1}$ follow random walk that do not have unconditional mean, we will use the conditional mean of M_{t+1} and $\widehat{M}_{t+1}$ at t around which welfare is evaluated.

2.2.1. Welfare Loss due to RI

As shown in the preceding subsection, the actual consumption path will deviate from the first-best immediately adjusted consumption path due to information processing constraints. Thus, we are interested in how much utility the consumer will lose in a RI economy. Specifically, we will measure the loss by computing the direct reduction of welfare in steady state or a money metric which is equivalent with the welfare loss.

First, following Cochrane (1989), Pischke (1995), Gabaix and Laibson (2001), and others, we use a money metric to measure welfare cost of deviating from the first-best full information rational expectation solution. As we just showed, we can easily derive the level of expected lifetime utility, $V(M_t)$ and $\widehat{V}(\widehat{M_t})$, by following different decision rules, and then we get the loss of time t expected welfare

$$\Delta V = V(M_t) - \widehat{V}(\widehat{M_t})$$

To convert it to dollars per quarter, following Cochrane (1989) and Pischke (1995), we divide ΔV by the marginal utility of a dollar at time t and then convert it to quarterly rates by multiplying $\frac{R-1}{R}$,

$$\begin{aligned}
\$ \text{ Loss/quarter} &= \frac{R-1}{R}\frac{E[\Delta V]}{u'(\overline{Y})} = \frac{R-1}{R}\frac{E[\Delta V]}{\overline{C}-\overline{Y}} \\
&= \frac{R-1}{R}\sigma\frac{\frac{1}{2}(\frac{R^2-1}{R})\sigma_M^2 + A_2(Var[M_t] - Var[\widehat{M_t}])}{\overline{Y}} \qquad (2.39)
\end{aligned}$$

where $\overline{C}$ is the bliss point of consumption (it is normalized to 1 in our setup), $\overline{Y}$ is the mean income, σ is the local coefficient of relative risk aversion and equal to $\frac{\overline{Y}}{\overline{C}-\overline{Y}}$ for the utility function $u(\overline{Y})$. Following the quarterly data 1954-1990 used in Pischke (pp. 830, 1995), we set the parameters as follows: $\overline{Y} = \$6,929$, the standard deviation of individual income $\omega_\zeta = \$2,470$, $R = 1.01$, and $\sigma = 4$. Using the expression (2.39), table 2 below reported utility costs for several values of the local CRRA σ and channel capacity κ.

Table 2: Utility Losses (\$/quarter)

	$\sigma = 1$	$\sigma = 4$	$\sigma = 10$
$\kappa = 0.3$ bits	0.174	0.69	1.74
$\kappa = 0.6$ bits	0.067	0.27	0.67
$\kappa = 1$ bits	0.028	0.12	0.29
$\kappa = 2$ bits	0.006	0.02	0.06
$\kappa = 3$ bits	0.001	0.01	0.01

It is clear from the table that the welfare losses due to RI are trivial. For example, for $\sigma = 4$ and $\kappa = 1$ bit, the loss only amounts to 12 cents per quarter. This result is similar to the findings by Pischke (1995) who calculated utility losses in the no-aggregate-information model[26]. Furthermore, even for a high value of the CRRA, $\sigma = 10$, and a low channel capacity, $\kappa = 0.3$ bits, the welfare loss is still minor: only around \$1.74.

We can also adopt another simple way to measure the welfare loss due to RI. Suppose that the welfare

[26] He found that in most cases the utility losses due to no information about aggregate shock to income are less than \$1 per quarter.

loss due to RI is equivalent with a reduction in steady state wealth, $(1-\varphi)\overline{M}_t$, where φ is the maximum fraction of steady state wealth that the consumer would like to pay to avoid his welfare loss due to RI. Put it mathematically, in steady state, $\overline{M}_t = E_t[M_{t+1}] = E_t[\widehat{M}_{t+1}]$, we have

$$V((1-\varphi)E_t(M_{t+1})) = \widehat{V}(E_t(\widehat{M}_{t+1}))$$

Solving this quadratic equation for φ yields

$$\varphi = \frac{2 + A_1/(A_2\overline{M}_t) + \sqrt{[2 + A_1/(A_2\overline{M}_t)]^2 - 4(A_0 - B_0)/(A_2^2\overline{M}_t)}}{2}$$

where we ignore the negative eigenvalue since it implies that RI generates positive welfare effect. We consider the following numerical example to evaluate welfare effect of RI. We assume that $R = 1.01, \omega_\zeta^2 = 0.001$, $\kappa = 1$, and Y_t^T follows a uniform distribution over $[0,1]$ with mean 0.5. Note that in this case the steady state permanent income $\overline{M}_t$ should be less than $(1-H_0)/H_1$[27]. Figures 4 and 5 show that the welfare loss in this case is tiny. For example, when $\overline{M}_t = 5, \kappa = 1$, and $\omega_\zeta^2 = 0.001$, welfare loss φ is only around $4 \cdot 10^{-6}$. The figures also show that the welfare loss due to RI increases with income volatility and decreases with channel capacity. The intuition is that since consumers are risk averse, they dislike uncertainty and risk. Hence, since both high income volatility (which has direct effect on welfare) and low capacity (which has indirect effect on welfare) induce high steady state conditional variance of the state, consumers are willing to pay more to avoid this induced uncertainty.

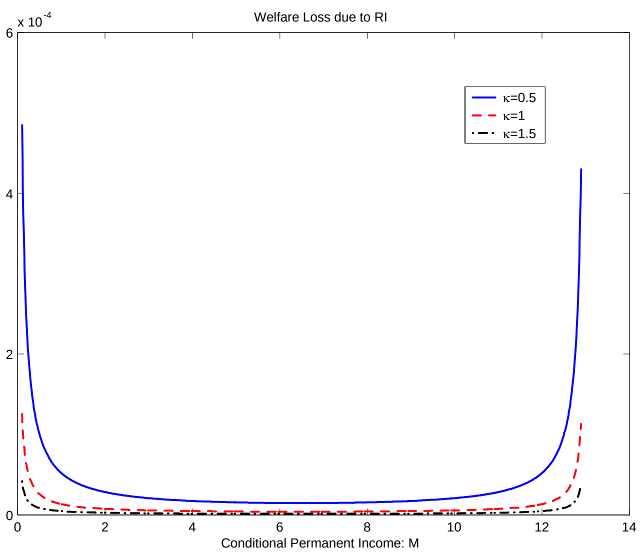

Figure 4

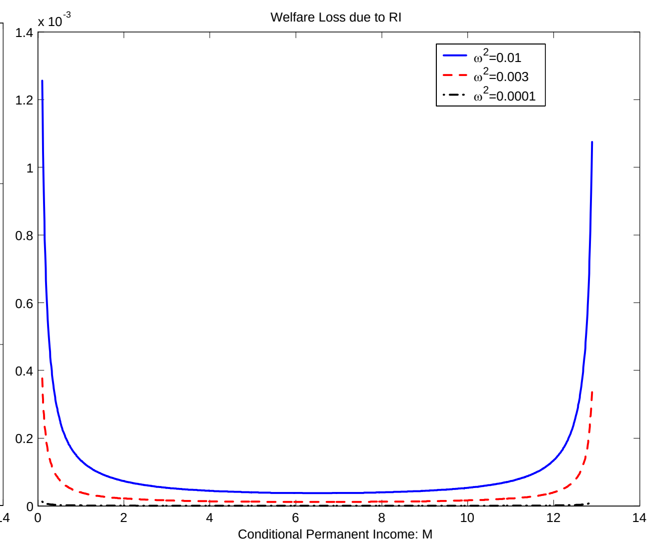

Figure 5

Based on the above calculations, we can conclude that for plausible assumptions, the welfare losses due to RI are negligible[28]. This provides some evidence that it is quite reasonable for consumers to devote low channel capacity in observing and processing information, as the welfare improvement from adopting

[27] That is because in steady state $\overline{C}_t = H_1\overline{M}_t + H_0 \leq 1$ (1 is the bliss point).

[28] This conclusion is also consistent with that in Akerlof and Yellen (1985), Cochrane (1989), Pischke (1995), and Gabaix and Laibson (2001).

the first-best decision rule is trivial. In other words, although consumers can devote much more capacity in processing information and improve their consumption decision, it is also rational for them not to do so since the welfare improvement is so tiny. As a result, this rational inertial behavior can general quite different individual and aggregate consumption and savings dynamics and make the model economy fit the data better.

2.2.2. Welfare Costs of Aggregate Fluctuations under RI

Alternatively, we may evaluate the welfare consequences of RI by examining the welfare costs of aggregate fluctuations in the RI model economy. Lucas (1987) argued that the welfare cost of business cycles may be very small. His estimates of the welfare cost of consumption fluctuations is no more than 0.0001% of aggregate consumption given logarithm preference. Here we will revisit this issue in the RI economy and examine if introducing RI can give us a different conclusion about the welfare cost of aggregate fluctuations. To concentrate on the effects of aggregate fluctuation in income/output[29] on welfare, here we start with the Representative-Agent model where there is no idiosyncratic shocks to income. For simplicity, the notations of income processes used here are same as the ones used in section 2.1, and aggregate income process is also composed by permanent shock Y_t^P and idiosyncratic shock Y_t^T. Hence, the consumption decision rule of the representative agent is $C_t = \frac{R-1}{R}\widehat{M}_t + \frac{1}{R}\overline{Y}^T$ where $\widehat{M}_t = E_t[M_t]$.

The traditional way to think about the welfare costs of business cycles is offering the risk averse consumer two possible consumption streams, one of which is constant and the other has the same mean but fluctuates around the mean with some volatility. The risk averse consumer would always prefer the constant consumption stream and thus require some additional consumption[30] to be indifferent between the two streams. Since RI does introduce additional uncertainty due to information capacity constraints, intuitively, RI would increase the welfare costs of business cycles. We will calculate the exact amounts of the welfare costs in both the full-information rational expectation (henceforth, RE) economy and the RI economy and examine the difference in welfare costs between the two economies.

Following Lucas (1987), the procedure is as follows. First, we denote $W^{RE} = V(\overline{M}_t)$ and $W^{RI} = \widehat{V}(\overline{M}_t)$ as the lifetime welfare under RE and RI[31], respectively, where $\overline{M}_t = E_t[M_{t+1}] = E_t[\widehat{M}_{t+1}]$. Note that both of them are affected by the uncertainty in the economy, which can be reflected in the constant terms in the value function. In the economy without uncertainty, the lifetime welfare can be written as

$$W^C = \frac{1}{1-\beta}[C - \frac{1}{2}C^2]$$

where C is steady state consumption and is equal to $H_0 + H_1\overline{M}_t$.

Second, we let ΔC^{RE} denote the welfare cost due to uncertainty in terms of consumption in the RE economy, and thus the following equality holds

$$\frac{1}{1-\beta}[(C - \Delta C^{RE}) - \frac{1}{2}(C - \Delta C^{RE})^2] = A_0 + A_1\overline{M}_t + A_2^2\overline{M}_t^2, \tag{2.40}$$

[29] We do not model aggregate shocks to technology, government spendings, and monetary policy explicitly in this paper, but assume that this aggregate shocks to income reflect business cycle fluctuations.

[30] This fluctuation effect of uncertainty is always deteremental to welfare since the consumer is risk averse.

[31] Since both follow random walk, there is no traditional steady state here, and we use conditional expectation of M_{t+1} and $\widehat{M}_{t+1}$, $\overline{M}_t$ to represent steady state in both model economies.

which implies a quadratic equation in terms of ΔC^{RE}

$$(\Delta C^{RE})^2 + [1 - (H_0 + H_1 \overline{M}_t)]\Delta C^{RE} + \beta A_2 \omega_\zeta^2 = 0. \tag{2.41}$$

The positive eigenvalue of this equation[32] is

$$\Delta C^{RE} = \frac{-[1 - (H_0 + H_1 \overline{M}_t)] + \sqrt{[1 - (H_0 + H_1 \overline{M}_t)]^2 - 4\beta A_2 \omega_\zeta^2}}{2}. \tag{2.42}$$

Thus, the higher the volatility of fundamental shock, ω_ζ^2, the higher is the welfare costs of business cycles.

Similarly, in the RI economy, we denote ΔC^{RI} as the welfare cost due to both fundamental uncertainty and induced uncertainty due to RI

$$\frac{1}{1-\beta}[(C - \Delta C^{RI}) - \frac{1}{2}(C - \Delta C^{RI})^2] = B_0 + B_1 \overline{M}_t + B_2^2 \overline{M}_t^2 \tag{2.43}$$

Solving it yields

$$\Delta C^{RI} = \frac{-[1 - (H_0 + H_1 \overline{M}_t)] + \sqrt{[1 - (H_0 + H_1 \overline{M}_t)]^2 - 4\beta B_2((R^2 - 1)\sigma_M^2 + \omega_\zeta^2)}}{2}, \tag{2.44}$$

which means that the higher either the volatility of fundamental uncertainty or inattentiveness, the larger is the welfare costs of business cycle fluctuations. The intuition behind this result is that if the consumer has low capacity in observing the state of the economy, he is more possible to be hurt by a sequence of bad aggregate shocks in that he can not adjust his economic decisions instantaneously. As a result, compared with the full-information RE case, here he would prefer to pay more to avoid total uncertainty. Mathematically,

$$\Delta C^{RI} \geq \Delta C^{RE} \tag{2.45}$$

where the equality holds when channel capacity converges to infinity.

Furthermore, it is straightforward to show that ΔC^{RI} should be decreasing with channel capacity κ since higher κ implies lower σ_M^2 given the variance of fundamental shocks. (2.42) and (2.44) imply the upper bounds[33] for the welfare costs in the two cases are $\Delta \overline{C}^{RE} = \sqrt{-4\beta A_2 \omega_\zeta^2}/2$ and $\Delta \overline{C}^{RI} = \sqrt{-4\beta B_2((R^2 - 1)\sigma_M^2 + \omega_\zeta^2)}/2$, respectively. Hence, even if the welfare cost computed from the full-information RE model is very small, the welfare cost implied by the RI model with the same steady state values may be much larger if channel capacity is low enough. Consequently, stabilization macroeconomic policies may be necessary to eliminate some fundamental risks since it implies a large benefit for the consumers. We use $\iota = \Delta \overline{C}^{RI} / \Delta \overline{C}^{RE} - 1$ to measure the additional welfare costs in the RI model. It is straightforward to derive that $\iota = \sqrt{(R^2 - 1)/(\exp(2\kappa) - R^2) + 1} - 1$. It is obvious that if channel capacity is low enough, ι may be high, and it approaches to 0 when κ approaches to ∞.

As we argued in the preceding part, although this upper bound for the welfare costs of business cycles could be large, under reasonable assumptions about income process and conditional mean of permanent

[32] Here we ignore negative eigenvalue in that the risk averse consumer dislikes uncertainty.

[33] Note that when consumption reaches the bliss point in the steady state, the welfare costs reach these upper bounds.

income, the welfare loss is negligible too.

The above story can be extended to heterogeneous-agent case directly. In this case, we assume that every consumer may have different channel capacity in observing the state of the economy, and thus the consumers with low capacity suffer more from business cycle fluctuations than the consumers with high capacity. Furthermore, in this case the average welfare costs in the economy depends on the distribution of consumers over their channel capacity. Hence, how information disperses across the population and changes consumers' attention allocation would have significant welfare effects.

2.3. Endogenous Channel Capacity

So far, we have not considered costly attention and just set channel capacity κ exogenously. Since attention is a scarce economic resource and there are many competing demands, it is natural to ask how to determine optimal channel capacity endogenously. In this subsection, we modify our baseline model by assuming that there is a cost function associated with attention when individuals make consumption decisions every period. Specifically, we assume that this cost function is an increasing function of capacity and take the form $q\kappa^{\alpha}$. Furthermore, we suppose that the consumers choose optimal capacity to minimize the average difference between the value function in the RI model and the value function in the standard full information model. The optimization problem can then be characterized by

$$\kappa^* = \arg[\frac{1}{2}(\frac{R^2-1}{R})\frac{\omega_\zeta^2}{\exp(2\kappa)-R^2} + \frac{q\kappa^\alpha}{1-\beta}] \tag{2.46}$$

To obtain a closed-form solution, we set $\alpha = 1$ and the FOC is

$$(\frac{R^2-1}{R})\omega_\zeta^2\frac{1}{[\exp(\kappa)-R^2\exp(-\kappa)]^2} = \frac{q}{1-\beta}$$

It can be simplified to

$$(\frac{R^2-1}{R})\omega_\zeta^2\frac{1}{\sinh^2\kappa} \simeq \frac{q}{1-\beta},$$

that is,

$$\kappa^* = \arcsin\mathrm{h}[(1-\beta)(\frac{R^2-1}{R})\omega_\zeta^2/\sqrt{q}] \tag{2.47}$$

Proposition 2. *In this LQG PIH model with RI and costly attention, the optimal capacity is increasing with the persistence and volatility of income shock and the interest rate, and is decreasing with the marginal cost of capacity q.*

The intuition behind the above results is simple: the higher income uncertainty, the larger is the cost of RI due to the exposure to risk, and thus the more attention he uses to monitor his wealth evolution. Similarly, the larger the interest rate, the larger the effects of RI on his future wealth level and then on his welfare. As a result, he chooses higher optimal capacity used in digesting his economic situation. Finally, the higher the marginal cost of capacity, that is, the higher the price of attention, the less amount of attention would be used in his economic decisions. Consider a numerical example. Given the following parameter values: $R = 1.01, \beta = 1/R, \omega_\zeta = 0.06$, and $q = 0.0002$, we find that $\kappa^* = 1.1$ bits.

2.4. Optimal Allocation of Information Capacity

In section 2.1, we show that we can solve the multivariate state RI model analytically by redefining a state variable which is a linear combination of all elements in the original state vector. However, in many cases this kind of linear combinations may not exist. Hence, in these cases, we need to use numerical method to solve steady state conditional variance-covariance matrix Σ which is optimal according to some criterion and satisfies the information processing constraints. The main objective of this subsection is to propose an optimization procedure to solve the optimal capacity allocation problem, that is, the optimal steady state Σ.

We suppose that in steady state Σ_t is a *constant* matrix. As a result, we need not include this second-order moment in the state vector and it affects welfare via other terms in the value function. We propose a two-stage procedure here: in the first stage, we take the optimal steady state Σ as given and then derive the optimal consumption rule, and then substitute the consumption rule into the Bellman equation to pin down the undetermined coefficients in the value function. In the second stage, we choose the optimal steady state Σ to minimize or maximize the proposed criterion under two information processing constraints.

Instead of minimizing the criterion $E_t V(S_t) - \widehat{V}(\widehat{S}_t)$ subject to information constraints as proposed in Sims (2003), here we propose another procedure to solve a multivariate-state model (see Appendix for detailed derivations for optimal Σ):

$$\max_{\Sigma} E[V(S_t) - \widehat{V}(\widehat{S}_t)] \tag{2.48}$$

$$s.t.$$

$$\log_2 |\Psi| - \log_2 |\Sigma| \leq 2\kappa; \;\; \Psi \succeq \Sigma$$

where

$$E[V(S_t) - \widehat{V}(\widehat{S}_t)] = -\frac{1}{1-\beta} tr[(G_1 F_2 G_1' - F_2)\Sigma] + tr[F_2(Var(S_t) - Var(\widehat{S}_t))] \tag{2.49}$$

2.5. Comparison with the PIH Model with Habit Formation

In this subsection, we compare the RI model with the internal habit-formation model. Habit formation in consumption is widely used in economics to study consumption and wealth accumulation behavior, asset pricing, and optimal monetary policy. For example, see Constantinides (1990), Deaton (1992), Boldrin, Christiano, and Fisher (2001), and Amato and Laubach (2004). Intuitively, the RI model should have some similarity with the habit-formation model since the habit formation models also imply that slow adjustment in consumption is optimal because consumers not only smooth consumption level, but also the growth of consumption, while the RI model predicts that slow and delayed adjustment consumption is optimal because capacity constraints make consumers take more time to observe and process information. So even the two kinds of models have very different mechanisms for consumption adjustment, they may have similar impacts on consumption and savings dynamics. We will explore this idea in this subsection.

First, consider a simple model with internal and additive[34] habit formation, in which we assume that higher consumption in last period creates a habit that lowers utility in this period. Under this assumption,

[34] "Internal" means that individuals care about how their consumption in the current period compares to their own consumption in the past.

we have the following preference:

$$V(W_0) = \max_{C_t} E_0 \sum_{t=0}^{\infty} \beta^t u(C_t - \gamma C_{t-1})$$

where $u(C_t - \gamma C_{t-1})$ is also a quadratic function in terms of $\widetilde{C}_t = C_t - \gamma C_{t-1}$, that is, $u(\widetilde{C}_t) = \widetilde{C}_t - \frac{1}{2}\widetilde{C}_t^2$. In this model, in making their consumption decisions, consumers take into account the effects of their current consumption on their future reference stock. Following Alessie and Lusardi (1997), we have the following recursive solution for consumption

$$C_t = \frac{\gamma}{R} C_{t-1} + (1 - \frac{\gamma}{R})(H_1 M_t + H_0) \tag{2.50}$$

where consumption is also a weighted average of past consumption and permanent income; the stronger the habit, the more weight will be assigned to the past consumption. Apparently, equation (2.17) and (2.50) deliver very similar consumption dynamics in the PIH model except that in the RI case there is an endogenous i.i.d. noise term with mean zero that affects consumption. When the two key parameters $\theta = 1 - \frac{1}{\exp(2\kappa)}$ (governs the ability of information processing) and γ (governs the degree of habit formation) satisfy the following equality,

$$1 - \theta = \frac{\gamma}{R} \tag{2.51}$$

the dynamic behaviors of consumption in these two different models are very similar.

As usual, we can rewrite the PIH model in terms of savings rather than consumption. Following Campbell (1987) and Deaton (1992), we define the expression for savings D_t as follows[35]

$$D_t = \frac{R-1}{R} W_t + \frac{1}{R} Y_t - C_t = \frac{R-1}{R} M_t + \frac{1}{R} Y_t^T - C_t. \tag{2.52}$$

Hence, in the PIH model with RI, we have the following recursive saving function

$$D_t = (1-\theta) R D_{t-1} + (1-\theta)\Delta Y_t^T + \frac{1}{R}\theta(Y_t^T - \overline{Y}^T) - \frac{R-1}{R}\theta\xi_t \tag{2.53}$$

where saving depends on past saving and on a linear combination of current income change ΔY_t and the discounted present value of future income changes (note that in the i.i.d. case, this is just $\frac{1}{R}(Y_t^T - \overline{Y}^T)$. The weights in the linear combination depend on the consumer's inattention θ. The stronger the inattention, the less importance of future income changes and the more important of past savings. Furthermore, saving in the model with rational inattention is also affected by the endogenous noise term $-\frac{R-1}{R}\theta\xi_t$.

In the habit formation model, the saving function can be characterized by the following recursive form:

$$D_t = \gamma R D_{t-1} + \gamma \Delta Y_t^T + \frac{1}{R}(1-\gamma)(Y_t^T - \overline{Y}^T) \tag{2.54}$$

Hence, in the habit formation model, saving depends on past saving and on a linear combination of current transitory income change ΔY_t^T and the discounted present value of future income changes. Again, when

[35] In Campbell (1987), he defines that saving $D_t = \frac{r}{1+r} A_t + Y_t - C_t$, where A_t is a single asset carried over from $t-1$. Here we just define $W_t = A_t + Y_t$ as state variable, hence we have the following expression for savings.

$1-\theta=\frac{\gamma}{R}$, the two saving functions, (2.53) and (2.54), are similar too. This conclusion is not surprising since the same condition holds in terms of the consumption functions.

The intuition behind the equality (2.51) is that since the parameter γ in the habit formation model also governs the speed with which reference stock adjusts, the larger γ is, the higher dependence of current consumption on past consumption, while in the RI model, the more the inattentiveness, (i.e., less θ), the less attention the consumer will pay to "news" and thus wealth level in the current period and then depend larger on past *effective* wealth level which can be expressed by past consumption level. Thus, the effects of RI and habit formation on consumption dynamics and thus savings behaviors are very similar except that the recursive consumption and saving equations in the RI model are affected by endogenous noise term ξ_t.

Finally, both models also generate similar wealth accumulation dynamics. In the RI model, combining (2.13) and (2.14) in section 2.1 gives the dynamics of permanent income M_t,

$$M_{t+1}-M_t=\zeta_{t+1}+(R-1)\frac{(1-\theta)\zeta_t-\theta\xi_t}{1-(1-\theta)R\cdot L} \tag{2.55}$$

where $\zeta_{t+1}=\frac{R}{R-1}\varepsilon_{t+1}+\eta_{t+1}$ is a linear combination of permanent and transitory shocks to income and ξ_t is noise term. Since $M_t=W_t+\frac{1}{R-1}Y_t^P$ and Y_t^P is random walk, the above equation of permanent income can be easily transformed to an equation for wealth dynamics, that is,

$$W_{t+1}-W_t=(\varepsilon_{t+1}+\eta_{t+1})+(R-1)\frac{(1-\theta)\zeta_t-\theta\xi_t}{1-(1-\theta)R\cdot L}. \tag{2.56}$$

Hence, in the RI model, wealth also follows a random walk.

By contrast, in the above habit-formation model, combining the flow budget constraint (2.13) and (2.50) yields[36]

$$W_{t+1}-W_t=(\varepsilon_{t+1}+\eta_{t+1})+(R-1)\frac{\gamma}{R}\frac{\zeta_t}{1-\gamma\cdot L}. \tag{2.57}$$

Hence, once $1-\theta=\frac{\gamma}{R}$, individual wealth dynamics in both model are very similar except that in the RI model the individual dynamics is driven by noise terms. However, the impacts of these terms may be largely reduced when aggregating over all individual. As a result, we expect that these two models can generate more similar dynamics in aggregate level.

Finally, although we do not compare the welfare costs of fluctuations in both models here, intuitively, they would have different implications for welfare since in the habit formation model, the preference is changed and then has first-order impact on welfare loss, while in the RI model the preference is unchanged, consumers change their decision rule due to information processing constraints, which has second-order impact on welfare.

3. Consumption Risk and Asset Returns

The standard consumption-CAPM model predicts that the quantity of stock market risk is determined by the contemporaneous covariance of consumption growth with equity returns. Consequently, the low contemporaneous covariance of consumption growth with the excess equity return in the US data implies that equities should not be very risky. However, the mean equity premium over the riskless rate is very high

[36] Note here we assume that $\gamma R<1$.

in the US data[37]. This issue is known as the equity premium puzzle in the asset pricing literature. Mehra and Prescott (1985) first identified it as a puzzle. Numerous economic arguments have been proposed to explain this puzzle. A partial list of these explanations include: habit formation in consumption, non-expected utility function, limited participation, and prospect theory. See Kocherlakota (1996) and Campbell (1999) for a recent survey. Another important potential explanation considers delayed adjustment or adjustment costs, for example, Lynch (1996), Marshall and Parekh (1999), and Gabaix and Laibson (2001). In this section, we will explore a CCAPM model with RI and reconcile three apparent anomalies in the full information CCAPM model simultaneously: a) the excess smoothness of aggregate consumption, b) the low contemporaneous covariance between consumption growth and asset returns, and c) the high equity premium.

In the previous section, we have shown that the LQ PIH model with RI generates slow and smooth responses of consumption with respect to additive income shocks. As a result, aggregate consumption exhibits excess smoothness. As will be shown below, the consumption-CAPM model with RI can also generate the slow adjustment of consumption with respect to changes in wealth[38]. Consequently, the contemporaneous covariance between consumption and wealth understates the risk of equity, thus a long-term consumption provides a correct measure. Hence, incorporating RI into the CCAPM model could be a potential explanation for the slow adjustment of consumption in the data and rationalize the key assumption adopted in Parker (2001, 2003) and Parker and Julliard (2005) that the ultimate consumption risk is a better measure of the risk of equity than the contemporaneous risk.

The main contribution here is that we solve the full-fledged CCAPM model with RI analytically and then evaluate to what extent incorporating RI can better explain the data in some important aspects. Furthermore, when we take the risk premium as given, we can also examine in this model how RI can affect optimal asset allocation. The basic idea is that since the long-term consumption risk is a correct measure for the risk of equity in the RI model and is significantly larger than the contemporaneous consumption risk, the fraction of wealth invested in the risky portfolio should be smaller. In other words, if the consumer can not allocate enough channel capacity in monitoring his financial wealth evolution, it is not rational to invest a large fraction of his wealth in risky portfolio because the innovations to his financial wealth can generate large consumption risk if the capacity is low.

3.1. A Consumption CAPM Model with Rational Inattention

In this subsection, we first incorporate RI into a standard consumption CAPM model and solve it analytically. We then evaluate the effects of RI on consumption dynamics, the covariance of consumption growth and equity returns, and the equity risk premium. Specifically, the procedure is composed by three steps: (1) Following Campbell (1993, 1999), we solve the full-information standard CCAPM model and derive the standard consumption rule. (2) Since the CRRA specification could be fitted into the Gaussian-error framework approximately, introducing RI would replace the true state variable in the standard consumption rule with the perceived state. Furthermore, we can express the change in consumption as a MA(∞) process. (3)

[37] As documented in Campbell (1999), in quarterly US data from 1947.2 to 1996.4, the average real stock return has been 7.6% at an annual rate, while the average riskless real interest rate is low and in the data set the average real return on 3-month Treasury bills have been around 0.8% per year.

[38] This RI model predicts similar aggregate dynamics generated from the delayed adjustment model (e.g., the $6D$ bias model proposed by Gabaix and Laibson) and the habit formation model since all of them predict that aggregate consumption should react to lagged changes in wealth. Furthermore, although the Gaibax-Laibson model and the RI model can generate similar aggregate consumption dynamics, their individual consumption behaviors are different in that the consumers in the Gaibax-Laibson economy update their consumption every D periods, while consumers in the RI economy adjust consumption every period, but are subject to information capacity constraints when they process information about the state.

Given RI, we use S-period pricing function to pricing equity portfolio. In sum, in this section we combine three literatures: the standard log-linearized CCAPM model (see Campbell (1993, 1999)), the RI framework (see Sims (2003)), and the pricing function based on the ultimate consumption risk (see Parker (2001, 2003) and Parker and Julliard (2005)).

Consider the following simple CCAPM model, the identical consumers maximize the following intertemporal welfare[39] by choosing consumption,

$$E_0 \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\gamma} - 1}{1-\gamma}$$

where β is the discount factor, γ is the coefficient of relative risk aversion, and denote $\sigma = 1/\gamma$ the elasticity of intertemporal substitution.

To keep the analysis simple, we assume that there are two tradable financial assets: asset e is risky, with one-period log (continuously compounded) return $r_{t+1}^e (= \log R_{t+1}^e)$, while the other asset f is riskless, with constant log return given by $r^f (= \log R^f)$. We refer to asset e as a portfolio of equities, and to asset f as savings or checking accounts. Furthermore, we assume that r_{t+1}^e has expected return $r^f + \pi$[40] and unexpected component u_{t+1} with $Var[u_{t+1}] = \omega_u^2$.

The flow budget constraint for consumers can be written as

$$W_{t+1} = R_{t+1}^m (W_t - C_t) \tag{3.1}$$

where W_{t+1} is financial wealth which is defined as the value of financial assets carried over from period t at the beginning of period $t+1$, $W_t - C_t$ is savings[41], and R_{t+1}^m is the one-period return on savings and given by

$$R_{t+1}^m = \chi (R_{t+1}^e - R^f) + R^f \tag{3.2}$$

where χ is the proportion of savings invested in the risky asset[42]. Following Campbell and Viceira (2002), we can derive an approximate expression for the log return on wealth as follows

$$r_{t+1}^m = \chi (r_{t+1}^e - r^f) + r^f + \frac{1}{2} \chi (1-\chi) \omega_u^2. \tag{3.3}$$

Since the term $\frac{1}{2}\chi(1-\chi)\omega_u^2 \leq 0.125\omega_u^2 = 0.0028$ is two order of magnitude smaller than the mean values

[39] We may locally approximate the CRRA utility function by a Log-LQ utility function. Formally, we approximate the CRRA function with a second-order Taylor expansion around $\overline{C}_t = E_t[C_{t+1}]$:

$$\begin{aligned} \frac{C_{t+1}^{\gamma}}{\gamma} &\simeq \frac{\overline{C}_t^{\gamma}}{\gamma} + \overline{C}_t^{\gamma} (\log C_{t+1} - \log \overline{C}_t) + \frac{1}{2}(\gamma - 1)\overline{C}_t^{\gamma} (\log C_{t+1} - \log \overline{C}_t)^2 \\ &= \frac{\overline{C}_t^{\gamma}}{\gamma} + \overline{C}_t^{\gamma} c_{t+1} + \frac{1}{2}(\gamma - 1)\overline{C}_t^{\gamma} c_{t+1}^2 \end{aligned}$$

where the coefficient of relative risk aversion is equal to $1-\gamma$. Hence, for plausible values of γ ($\gamma < 0$), the CRRA function can be approximated by the Log-LQ function and then we can apply the same Gaussian-error framework as in the LQ model in the CCAPM model.

[40] where r_f is the risk free rate and satisfies $\exp(r^f)\beta = R^f\beta = 1$, and π is risk premium on this risky portfolio.

[41] For simplicity, we do not model income process explicitly by assuming that all the income flows including labor income can be capitalized into marketable wealth.

[42] In this subsection, we set this proportion exogeneously and only discuss how RI affects asset pricing. We will discuss how RI affects asset allocation given equity premium in the next subsection.

of the returns on equity and the riskless asset given $\omega_u^2 = 0.15^2$, for simplicity, we may assume that $r_{t+1}^m \simeq \chi(r_{t+1}^e - r^f) + r^f$.

Since the return on the portfolio is not constant, the simple discrete-time model can not be solved analytically. Of course, it can be solved by using numerical methods adopted widely in the modern consumption literature and the infinite horizon models of portfolio choice with uninsurable labor income, but here we follow Campbell and Mankiw (1989), Campbell (1993), and Viceira (2001) and use the log-linearization method to solve the model[43]. First, we divide equation (3.1) by W_t and then log-linearize it around steady state $c - w = E(c_t - w_t)$:

$$\Delta w_{t+1} = r_{t+1}^m + \alpha_2 + (1 - 1/\alpha_1)(c_t - w_t) \tag{3.4}$$

where $\alpha_1 = 1 - \exp(c - w)$[44], $\alpha_2 = \log \alpha_1 - (1 - 1/\alpha_1)\log(1 - \alpha_1)$, and lowercase letters denote logs. Next, we can also log-linearize the Euler equation

$$E_t[\beta R_{t+1}^i (C_{t+1}/C_t)^{-1/\sigma}] = 1$$

around[45] $E_t[\log \beta - \frac{1}{\sigma}(c_{t+1} - c_t) + r_{t+1}^i]$ where $i = e$, f, and m, and obtain the following familiar form[46]

$$0 \simeq \log \beta - \frac{1}{\sigma} E_t[c_{t+1} - c_t] + E_t[r_{t+1}^i] + \frac{1}{2} Var_t[r_{t+1}^i - \frac{1}{\sigma}(c_{t+1} - c_t)] \tag{3.5}$$

Furthermore, we guess that the optimal log consumption rule take the following form $c_t = H_0 + H_1 w_t$, and thus

$$\Delta c_{t+1} = H_1 \Delta w_{t+1} \tag{3.6}$$

Combining equation (3.4), (3.5), and (3.6) can pin down the undetermined coefficients in the consumption rule:

$$H_1 = 1 \text{ and } H_0 = \log\{1 - \exp[(\sigma - 1)(r_f + \pi) + \sigma \log \beta + \frac{1}{2}\sigma(1 - \sigma)^2 \omega_u^2\},$$

To apply certainty equivalence principle in the RI framework, here we need to eliminate the precautionary savings terms in the decision rule. The simplest way to do that is to assume that $\sigma = 1$, and thus H_0 reduces to $\log(1 - \beta)$.

Given that the joint log normal distribution of the return on the portfolio and consumption growth, the expected return on the portfolio can be written as

$$\begin{aligned} E_t[r_{t+1}^e] - r^f &\approx \gamma Cov_t[c_{t+1} - c_t, r_{t+1}^e] \\ &= \gamma \chi \omega_u^2 \end{aligned} \tag{3.7}$$

[43] This method proceeds as follows. First, both the flow budget constraint and the consumption Euler equations are log-linearized around steady state, in particular, the Euler equations are log-linearized by a second-order Taylor expansion so that the second-moment effects such as precautionary savings effects are accounted. Second, guess optimal consumption and portfolio choices that verify these log-linearized equations. Finally, it pins down the coefficients of the optimal decision rules by using the method of undetermined coefficients.

[44] Given (3.1), in steady state, $\alpha_1 = 1/R_m$ where R_m is the steady state return on the market portfolio.

[45] Note that the Euler equation

$$E_t[\beta R_{t+1}^i (C_{t+1}/C_t)^{-1/\sigma}] = 1$$

can be written as

$$E_t[\exp(\log \beta - \frac{1}{\sigma}(c_{t+1} - c_t) + r_{t+1}^i)] = 1.$$

[46] Note that this log Euler equation holds exactly given consumption growth and returns are jointly conditionally lognormal.

where we use the formula that $\Delta c_{t+1} = \Delta w_{t+1} = r^m_{t+1} + \alpha_2 + (1 - 1/\alpha_1)H_0$.

Hence, the expected return on the risky asset is determined by the contemporaneous covariance of its return and consumption growth as well as risk aversion. Given γ is 1 and χ is 0.75, the equity premium is close to $0.75\omega_u^2$ which is around 2.4%[47]. Hence, this simple model may generate very reasonable risk premium given this low value for the coefficient of relative risk aversion, however, as we can see from equation (3.7), this result is generated from the unrealistic contemporaneous covariance between consumption growth and asset return which is $225 \cdot 10^{-4}$ annually[48] given $\omega_u = 15\%$. In Campbell's US dataset, the standard deviation of real consumption growth is around 0.01 annually, and the covariance of consumption growth with risky returns is around $3.2 \cdot 10^{-4}$. Therefore, one puzzle for this simple consumption CAPM model is that the model predicts too high contemporaneous covariance between consumption growth and asset returns as well as too high volatility in consumption. From the LQ PIH model developed in section 2, we have shown that introducing RI can not only reduce the variability of consumption growth, but also reduce the contemporaneous correlation between consumption growth and income innovation, thus it would be worth examining if the RI hypothesis can help solve the puzzles in this simple consumption CAPM.

Since the CRRA specification could be approximated by log-LQ framework, we can apply Gaussian-error framework approximately. Hence, adding RI in the above model yields the following modified consumption rule

$$c_t = H_0 + \widehat{w}_t \tag{3.8}$$

and the information state $\widehat{w}_t$ can be characterized by the following Kalman equation

$$\widehat{w}_{t+1} = (1 - \theta)\widehat{w}_t + \theta(w_{t+1} + \xi_{t+1}) \tag{3.9}$$

where θ and ξ_{t+1} have the same definitions as in section 2. Hence, consumption growth can now be written as

$$\Delta c_{t+1} = \Delta \widehat{w}_{t+1} = [\theta\chi u_{t+1} + \theta\chi \frac{((1-\theta)/\alpha_1)u_t}{1 - ((1-\theta)/\alpha_1) \cdot L}] + [\theta\xi_{t+1} - \frac{\theta\xi_t}{1 - ((1-\theta)/\alpha_1) \cdot L}] + \Omega \tag{3.10}$$

where L is the lag operator and Ω is the irrelevant constant term. Furthermore, for simplicity, we assume that the endogenous noise terms in the second bracket of the above expression will be cancelled out when aggregating over all consumers[49]. Consequently, we have the following proposition

Proposition 3. *Aggregate consumption growth in the CCAPM model with RI can be written as*[50]

$$\Delta c_{t+1} = \theta\chi u_{t+1} + \theta\chi \frac{((1-\theta)/\alpha_1)u_t}{1 - ((1-\theta)/\alpha_1) \cdot L}, \tag{3.11}$$

[47] Campbell (1999) reported a series of main moments from his dataset including 11 main industried countries. For the US stock market, their estimate of the standard deviation of unexpected log excess return ω_u is around 18% per year when the sample period is from $1891 - 1994$ and around 15% when the sample period is from $1947 - 1996$.

[48] We define the covariance of consumption growth and asset returns as $Cov(\Delta c_{t+1}, r^e_{t+1}) = \rho(\Delta c_{t+1}, r^e_{t+1}) \cdot \sigma(\Delta c_{t+1}) \cdot \sigma(r^e_{t+1})$. where $\rho(\Delta c_{t+1}, r^e_{t+1})$ is the correlation between equity return and consumption growth and $\sigma(\Delta c_{t+1}) = \chi\sigma(r^e_{t+1})$ is the standard deviation of consumption growth.

[49] As we discussed in section 2, this is an extreme case, and actually the individual noise terms may include a common component.

[50] *Since we focus on aggregate behavior and to avoid the notation confusion, in the following equation, we still use c to represent aggregate consumption.*

which implies that 1) the covariance between aggregate consumption growth and asset returns is[51]

$$Cov(\Delta c_{t+1}, r^e_{t+1}) = \theta\chi\omega_u^2, \tag{3.12}$$

2) the standard deviation of consumption growth is

$$\sigma(\Delta c_{t+1}) = \phi\chi\omega_u \tag{3.13}$$

where $\phi = \theta\sqrt{1 + \frac{((1-\theta)/\alpha_1)^2}{1-((1-\theta)/\alpha_1)^2}}$, *3) the correlation between consumption growth and equity return is then*

$$\rho(\Delta c_{t+1}, r^e_{t+1}) = 1/\sqrt{1 + \frac{((1-\theta)/\alpha_1)^2}{1-((1-\theta)/\alpha_1)^2}} \tag{3.14}$$

and 4) the autocorreltion of consumption growth is then

$$\rho_{\Delta c}(j) = Corr(\Delta c_t, \Delta c_{t+j}) = \frac{(\theta\chi)^2}{(\theta\chi)^2 + (\theta\chi(1-\theta)/\alpha_1)^2/[1-((1-\theta)/\alpha_1)^2]}[(1-\theta)/\alpha_1]^j \tag{3.15}$$

where $j \geq 1$.

Proof. (see Appendix B). ¥

Equation (3.11) means that aggregation consumption adjusts gradually to the shock to asset returns, and thus the contemporaneous covariance between consumption growth and asset returns becomes[52] $\theta\chi\omega_u^2$ rather than $\chi\omega_u^2$, that is, the measured contemporaneous covariance between consumption growth and risky returns will be lowered by θ. In the above simple CCAPM model without RI, this contemporaneous covariance is around $12.3 \cdot 10^{-4}$ at the quarterly frequency, and this figure is well above its US empirical counterpart which is around $0.8 \cdot 10^{-4}$. When $\theta = 0.1$, the theoretical covariance value becomes $1.23 \cdot 10^{-4}$ which is much closer to the empirical value.

Equation (3.13) means that RI can largely reduce the standard deviation of consumption growth since ϕ is less than 1 . Figure 6 below shows the relationship between the standard deviation of consumption growth and channel capacity. In the quarterly US data, the standard deviation of consumption growth is around $0.54 \cdot 10^{-2}$, which is well below $1.64 \cdot 10^{-2}$, the value predicted by the standard CAPM model without RI. However, in our RI model with $\theta = 0.3$, the theoretical value of $\sigma(\Delta c_{t+1})$ becomes $0.5 \cdot 10^{-2}$ which is also close to the empirical one.

[51] *Note that if we assume that there is also limited stock market participation in our RI economy and the fraction of wealth shares for stockholders is λ_s which is around 28% based on the US data, the comtemporaneous covariance between aggregate consumption growth with asset returns becomes $Cov(\Delta c_{t+1}, r^e_{t+1}) = \theta\chi\lambda_s\omega_u^2$.*

[52] To match the empirical evidences, here we may also consider the limited stock market participation effect, which is measured by $\lambda_s < 1$. Consequently, the contemporaneous covariance becomes $\theta\chi\lambda_s\omega_u^2$.

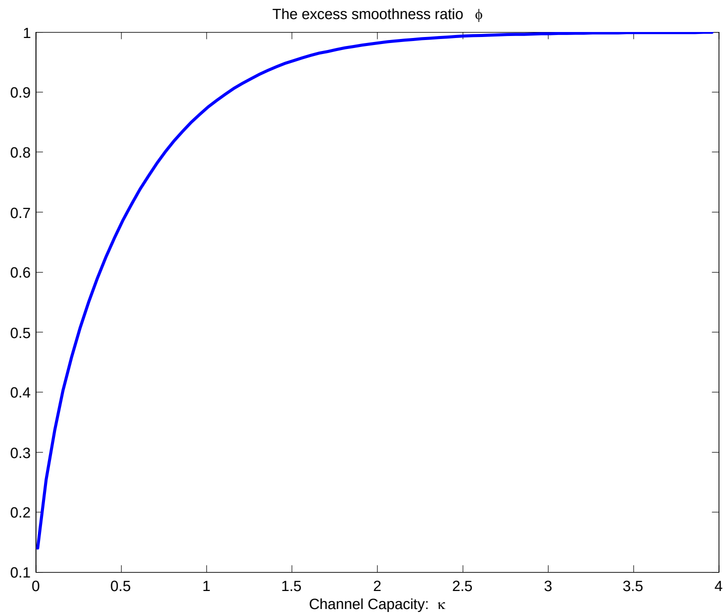

Figure 6

For illustrative purpose, we can consider the following question: what will an economist equipped with the consumption CAPM model find if he observes quarterly data from our RI economy, but thinks he is observing data from the standard model? This question can be answer after calculating

$$\widetilde{\gamma} \simeq \frac{\pi + r^f}{Cov_t[c_{t+1} - c_t, r^m_{t+1}]} = \frac{1}{\theta}\gamma$$

In other words, the estimate of the coefficient of relative risk aversion will be biased up by factor $\frac{1}{\theta}$. For example, if the true γ is 3 and $\theta = 0.1$, he will find that the estimated value $\widetilde{\gamma}$ will be 30. If considering other frictions like limited stock market participation and time aggregation, the estimated γ is further misleading.

3.2. Risk Premium and the Long-term Consumption Risk

Parker (2001, 2003) and Parker and Julliard (2005) argued that the long-term risk is a better measure of the ture risk of the stock market if consumption reacts with a delay to changes in wealth because the contemporaneous covariance of consumption and wealth understates the risk of equity. The hypothesis of RI here provides a potential explanation for the slow adjustment of consumption.

In this subsection, we attempt to reconcile two phenomena: the low covariance between consumption growth and the equity return and high equity premium in our RI model. We have shown above that introducing RI does reduce the contemporaneous covariance between Δc_{t+1} and r^e_{t+1}, and next, we show why this low covariance does not imply low equity premium. Following Parker (2003), we define the ultimate consumption risk as the covariance of asset returns and consumption growth over the period of the return and many following periods. Since our RI model predicts that consumption reacts to the innovations to asset returns gradually and slowly, it can rationalize the assumption used in Parker (2003) that consumption risk should be long term instead of contemporaneous. Furthermore, given the analytical solution for consumption growth in our RI model, it is straightforward to calculate the ultimate consumption risk. Specifically, when consumers behave optimally, we have the two consumption Euler equations for the market portfolio and the risk free asset

$$E_t[R^e_{t+1}(C_{t+1+S}/C_t)^{-\gamma}] = E_t[R^f_{t+1}(C_{t+1+S}/C_t)^{-\gamma}] \tag{3.16}$$

Log-linearizing equation (3.16) yields[53]

$$E_t[r^e_{t+1}] - r^f + \frac{1}{2}Var_t[r^e_{t+1}] = \gamma Cov_t[c_{t+1+S} - c_t, r^e_{t+1}] \tag{3.17}$$

Hence, we can write the mean asset return as follows

$$\begin{aligned} E_t[r^e_{t+1}] - r^f &\approx \gamma Cov_t[c_{t+1+S} - c_t, r^e_{t+1}] \\ &= \gamma \sum_{s=0}^{S} Cov_t[c_{t+1+s} - c_{t+s}, r^e_{t+1}] \end{aligned}$$

where $c_{t+1+S} - c_t = \sum_{s=0}^{S} \Delta c_{t+1+s}$ and $\Delta c_{t+1+s} = \theta\chi u_{t+1+s} + \theta\chi \frac{((1-\theta)/\alpha_1)u_{t+s}}{1-((1-\theta)/\alpha_1)\cdot L}$.

Hence, the ultimate risk of asset to consumption is larger than the contemporaneous risk to consumption since

$$\lim_{S\to\infty} \sum_{s=0}^{S} Cov_t[c_{t+1+s} - c_{t+s}, r^e_{t+1}] = \frac{\theta}{1-(1-\theta)/\alpha_1}\chi\omega_u^2 > \chi\omega_u^2, \tag{3.18}$$

that is, the impacts of the risk on consumption can last infinite following quarters. Since aggregate consumption adjusts gradually to the shocks to asset returns as predicted by our RI model, this ultimate consumption risk should be the best measure of the risk of an asset.

In the following numerical example, we will show that to what extent our RI model can better fit the US and International data in the following three dimensions: low consumption variability, low contemporaneous covariance between consumption growth and equity returns, and high risk premium. We assume that in the annual frequency $R^e = 1.02$, $R^f = 1.0025$, and $\chi \in [0.25, 0.5, 0.75]$, and then $R^m \in [1.0069, 1.0113, 1.0156]$. Thus, the effect of RI on the ultimate consumption risk can be measured by the term: $\frac{\theta}{1-(1-\theta)R^m}$. The following figure plots the relationship between this term and channel capacity κ.

[53]Note that the unconditional moments also hold if we assume that consumption growth and asset returns are joint unconditional normal distribution.

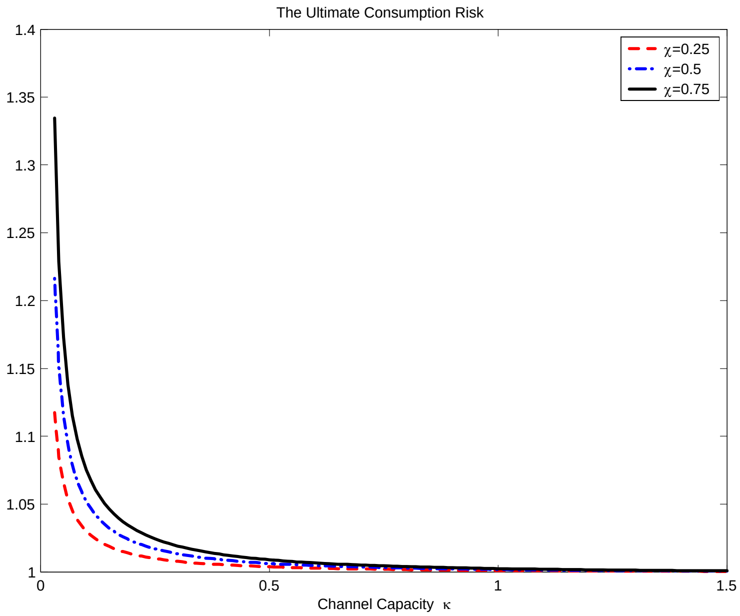

Figure 7

This figure shows that the effects of RI on the ultimate consumption risk is decreasing with channel capacity, that is, the larger the inattention of the consumer used in monitoring the dynamics of his financial wealth, the larger effect of RI on the ultimate consumption risk of the risky asset for the consumer. The intuition is same as before: the less information capacity the consumer used in his economic decisions, the larger the long-term consumption risk, and then he would hold less risky asset.

3.3. Implications of RI for Optimal Asset Allocation

After examining the impact of RI on the equity premium, it is natural to ask such a question: what is the impact of RI on optimal portfolio choice? To examine this issue, we need to fix the equity premium exogenously. Based on the pricing equation above, since the ultimate consumption risk is the best measure in the RI economy, we have

$$\begin{aligned} \pi &\approx \gamma Cov_t[\lim_{S\to\infty}(c_{t+1+S}-c_t), r^e_{t+1}] \\ &= \gamma\frac{\theta}{1-(1-\theta)/\alpha_1}\chi\omega_u^2 \end{aligned} \tag{3.19}$$

where π is the equity premium and χ is the fraction of wealth invested in stock market. We therefore have the following proposition about the optimal asset allocation in the risky asset.

Proposition 4. *The optimal asset allocation in the RI economy can be expressed by*

$$\chi = \vartheta_1\frac{\pi}{\gamma\omega_u^2} \tag{3.20}$$

where $\vartheta_1 = \frac{1-(1-\theta)/\alpha_1}{\theta} < 1$ is inversely proportional to the ultimate consumption risk of risky asset.

As predicted by the standard CCAPM model, the optimal fraction of savings invested in the risky asset is proportional to the risk premium (π) and the reciprocal of both the coefficient of relative risk aversion

(γ) and the variance of unexpected component in the risky asset (ω_u^2). What's new in the RI model is that the optimal allocation to the risky asset also depends on the degree of inattention. The larger the inattention, the higher is the ultimate consumption risk. As a result, individuals with low attention would invest lower share in risky asset. The intuition for this result is simple. In the RI economy, a one percent negative shock in individuals' financial wealth would affect their consumption more than that predicted by the full-information model. For this reason, rational inattentive individuals are willing to invest less in the risky asset. Note that we can rewrite the expression (3.20) as

$$\chi = \frac{\pi}{\widetilde{\gamma}\omega_u^2} \tag{3.21}$$

where the *effective* coefficient of relative risk aversion $\widetilde{\gamma} = \gamma/\vartheta > \gamma$. Hence, the asset allocation of the inattentive consumers is similar with the allocation of the more risk averse individuals[54]: both groups hold less risky portfolio.

3.4. Review of Related Empirical Evidence

A number of existing survey evidences support that 1) investors do not have enough knowledge about the evolution of their financial wealth and consequently can not adjust their consumption fully in response to the innovations to the returns, and 2) the innovations to their financial assets can be used to predict their future change in consumption. For example, for 1), Dynan and Maki (2000) analyzed the responses to the Consumer Expenditure Survey (CEX) from 1996 1. to 1999 1. and found that around one-third of stockholders reported no change in the value of their assets while the US stock markets rose over 15% per year during this sample period[55], and for 2), in the same paper, they also reported that for stockholders with over $\$10,000$ in securities, a 1% increase in the value of securities holding would cause lasting impacts on consumption growth and eventually consumption would increase by 1.03%, one third of which increases during the first 9 months, another third of which occurs from 10th month to 18th month, another quarter of which occurs form the 19th month to the 27th month, and the left occurs from the 28th month to the 36th month. This evidence may be largely captured by our RI model since equation (3.11) implies that

$$\begin{aligned} c_{t+1+S} - c_t &= \sum_{s=0}^{S} \Delta c_{t+1+s} \\ &= \theta\chi[1 + (1-\theta)/\alpha_1 + \cdots + ((1-\theta)/\alpha_1)^S]u_{t+1} \end{aligned}$$

where $\lim_{S\to\infty}[c_{t+1+S} - c_t] = \frac{\theta\chi}{1-(1-\theta)/\alpha_1}u_{t+1}$. Consider a numerical example (the time unit here is 3 quarters) in which $R^f = 1.005$, $\alpha_1 \simeq 1/R^f$, and $\theta = 0.2$ such that $\lim_{S\to\infty}[c_{t+1+S} - c_t] = 1.03$ as estimated from the data. When $S = 0$, $c_{t+1+S} - c_t = 0.36$, when $S = 1$, $c_{t+1+S} - c_t = 0.58$, and when $S = 2$, $c_{t+1+S} - c_t = 0.74$. Thus, our numerical example can generate very similar results as those estimated from the US data. Furthermore, using the estimation results from Dynan and Maki (2000), we plot figure 8 to illustrate to what extent our RI model can match the survey results. In the left figure, we define stockholders (henceforth, "SH") as households with securities $> \$1,000$, while in the right figure, we define SH as households with securities

[54] According to the mutual-fund separation theorem, more risk-averse individuals should hold more of their wealth in the riskless asset.

[55] Kennickell, et. al (2000) and Starr-McCluer (2000) also reported similar results based on alternative survery sources.

> \$10,000. And when we plot the the profile generated from our model, we calibrate the observation weight θ such that the initial jump of consumption to the shock to asset returns can match the data exactly, and then check if the responses to past shocks during the following 27 quarters (3 time units) can also fit the the dynamic responses reflected in the data. The left figure below shows that the RI model with $\kappa = 0.14$ can fit the empirical results very well: the responses of consumption to the innovations is muted initially and then increases gradually over time. The right figure also shows the similar pattern of the responses, though the fit is not as good as the left one.

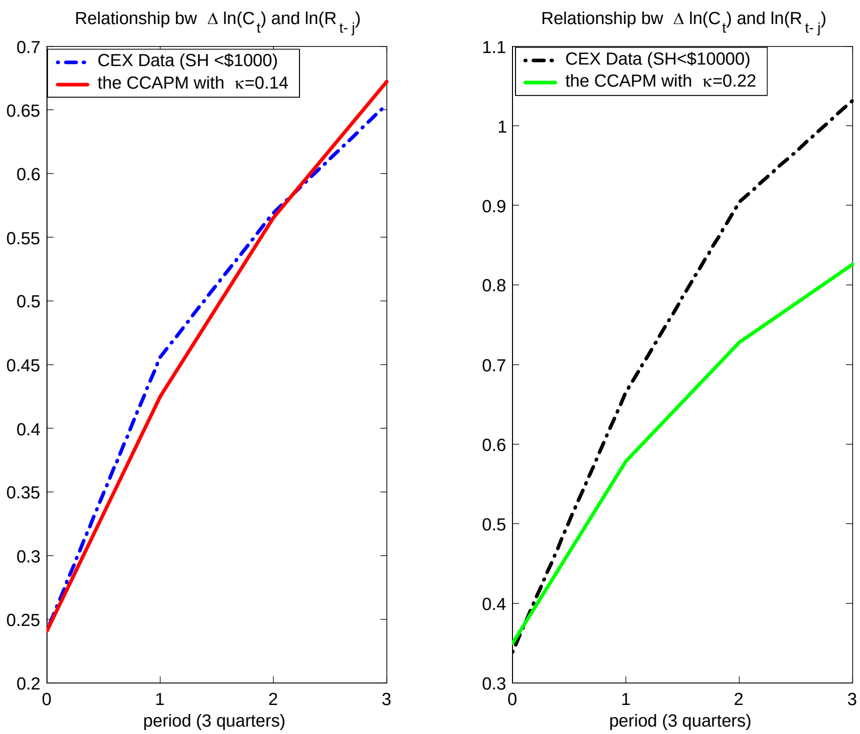

Figure 8

As in Gabaix and Laibson (2001), we are also interested in whether the values of $Cov[c_{t+1+S} - c_t, r^e_{t+1}]$ generated from our RI model can match the empirical counterparts. We use the cross-country panel dataset created by Campbell (1999), and plot the empirical covariances $Cov[c_{t+1+S} - c_t, r^e_{t+1}]$ in US and the average covariance across countries with large stock markets[56] in the following figure. The figure show a main feature in the data: the empirical covariances gradually increases with the horizon, s. Note that in the full information CCAPM model, the covariance should initially jump to a plateau and stay there. The following figure also show the covariance profiles generated by our RI model with different values of channel capaicty. It is obvious that the RI model can capture this apparent empirical feature successfully: the covariance slowly rise over time. The intuition here is same as before: if a large number of consumers/investors in the economy can not digest the innovations to their financial wealth and monitor their wealth evolution due to limited information capacity constraints, aggregate consumption should react to the shock to asset return with delay and be sensitive to lagged shocks. Actually, it is not unreasonable. For example, as argued in Thaler (1990), some consumers may put their retirement wealth in one of their mental accounts[57] and ignore

[56] Following the same criterion (ordered the countries in the dataset by the ratio of stock market capitalization to GDP) used in Gabaix and Laibson (2001), they are Switzerland (0.87), the United Kingdom (0.8), the United States (0.72), the Netherlands (0.46), Australia (0.42), and Japan (0.4).

[57] This mental account can be regarded as "asset account" and the MPC from this account is less than the MPC from the

the accumulating financial wealth until their retirement age 65.

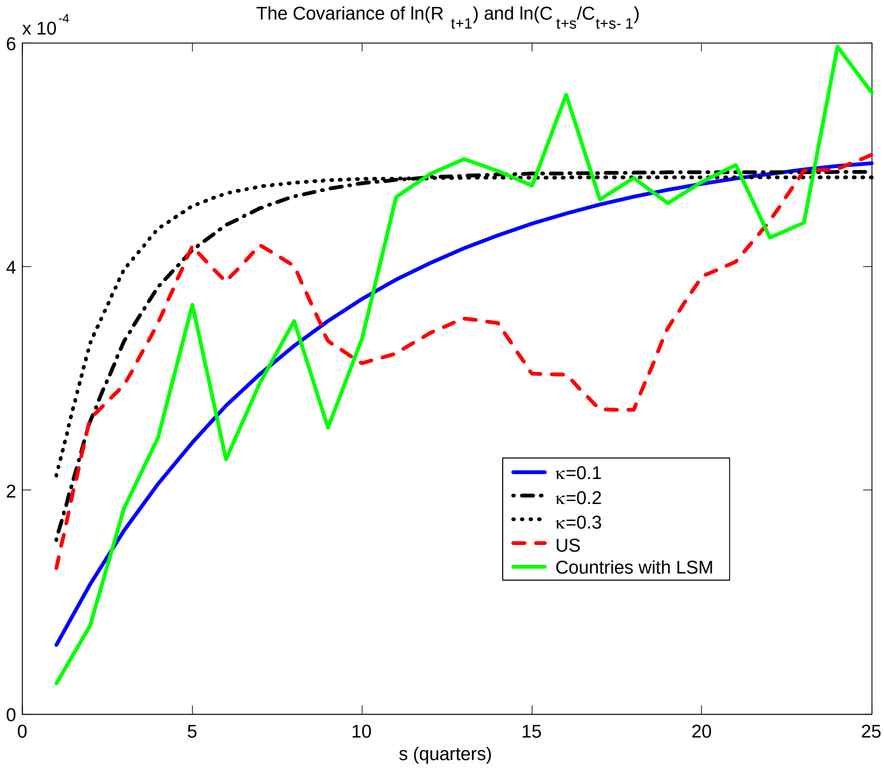

Figure 9

Furthermore, Parker (2001) used data from the CEX of the Bureau of Labor Statistics and calculate the covariance and risk aversion using the impulse responses to returns in a vector autoregression (VAR). This method provided a clear picture of consumption dynamics following an innovation in excess returns. Specifically, he estimated a three-variable VAR in excess returns, the logarithm of consumption, and the dividend-to-price ratio, each with four lags. Figure 1 reported in Parker (2001) ploted the responses of flow consumption to an innovation in excess returns and clearly showed that flow consumption adjusts gradually in response to innovation in excess returns and the adjustment lasts many periods, as the RI model predicts.

3.5. Incorporating Labor Income Risk

So far we have assumed that labor income can be capitalized into marketable wealth. But in reality human wealth[58] is nontradable in the market since it is difficult to sell claims against future labor income. Consequently, investors would adjust their financial asset holdings to take account of their implicit holdings of human wealth. For most investors, human wealth may tilt financial portfolios towards higher holdings of risky assets. Hence, it is interesting to examine how labor income risk affects both risk premium and asset allocations in the RI economy. In this subsection, we adopt the same preference specification as the preceding subsection, while model the flow budget constraint for consumers differently:

$$W_{t+1} = R^m_{t+1}(W_t + Y_t - C_t) \tag{3.22}$$

"current income" account.

[58] An individual's labor income can be seen as a dividend on the individual's implicit human wealth.

where W_{t+1} is financial wealth at the beginning of $t+1$ carried over from t, $W_t + Y_t - C_t$ is savings, R^m_{t+1} is the one-period return on savings given by equation (3.2), and Y_{t+1} is labor income. Following Viceira (2001) and Campbell and Viceira (2002), we specify the process for labor income as follows

$$Y_{t+1} = Y_t \exp(v_{t+1} + g) \tag{3.23}$$

where $\nu_{t+1} \sim NIID(0, \omega_v^2)$ and g is the deterministic growth rate. The empirical evidence suggests that individual labor income is composed by both permanent and transitory shocks. Here we ignore transitory shocks to labor income to make the notations simple. Furthermore, we assume that the innovation to labor income may be contemporaneously correlated with the innovation to equity return:

$$Cov_t(u_{t+1}, v_{t+1}) = \omega_{uv}. \tag{3.24}$$

Note that $\omega_{uv} = 0$ if labor income risk is idiosyncratic.

Following the same procedure used in the preceding subsection to derive optimal consumption and portfolio choice rule, we first log-linearize the flow budget constraint (3.22) around $c - y = E[c_t - y_t]$ and $w - y = E[w_t - y_t]$ as follows

$$w_{t+1} - y_{t+1} \simeq \eta + \eta_w (w_t - y_t) - \eta_c (c_t - y_t) - \Delta y_{t+1} + r^m_{t+1} \tag{3.25}$$

where lowercase letters denote variables in logs and η, η_w, and η_c are log-linearization constants that are given in Appendix B. The log consumption function takes the form

$$c_t = H_0 + H_1 m_t \tag{3.26}$$

where m_t is a new state variable which is defined as $w_t + \frac{1-\eta_w+\eta_c}{\eta_w - 1} y_t$ and

$$H_1 = \frac{\eta_w - 1}{\eta_c} \text{ and } H_0 = \frac{1}{\eta_c}\{\eta - g - \frac{\sigma}{H_1}\log\beta + (1 - \frac{\sigma}{H_1})E[r^m_{t+1}] - \frac{1}{2}\Xi\} \tag{3.27}$$

where Ξ is the precautionary savings term (See Appendix B for details).

As shown in the Appendix B, for the CCAPM model with labor income, Ξ and then H_0 explicitly depend on both the variance of the unexpected log equity return and the variance of labor income growth. As a result, we cannot apply the Gaussian-error framework for RI in this case directly since the certainty equivalence principle does not hold generally. But we may impose some condition to eliminate precautionary savings and then fit it into the RI framework as we characterize for the standard LQ case. It is clear from the expression for Ξ :

$$\Xi = (1 - \frac{1}{\sigma}H_1)^2 Var[r^m_{t+1}] + (\frac{1}{\sigma}H_1)^2 \lambda Var[v_{t+1}] - 2\frac{1}{\sigma}H_1\lambda(1 - \frac{1}{\sigma}H_1)Cov[r^m_{t+1}, v_{t+1}] \tag{3.28}$$

where $\lambda = \frac{1-\eta_w+\eta_c}{\eta_w-1}$. When the elasticity of substitution σ is set to $H_1 \in (0,1)$, $\Xi = \frac{1-\eta_w+\eta_c}{\eta_w-1} Var[v_{t+1}] = \frac{1}{\exp(c-y)-1} Var[v_{t+1}] > 0$. Hence, when $w - y$ is large[59] and $Var[v_{t+1}]$ is small enough, the precautionary

[59] Note that $c - y$ is a linear function of $w - y$.

savings term Ξ should be close to 0. As documented in Viceira (2001), Campbell and Viceira (2002), and others, $Var[v_{t+1}]$ is around 10% per year, but here we set may it to a very low value since we are just interested in accessing how RI affect risk premium and asset allocation in the presence of labor income risk rather than calibrating to the real economy.

Hence, adding RI in this model yields the following modified consumption rule

$$c_t = H_0 + H_1 \widehat{m}_t \tag{3.29}$$

and the information state $\widehat{m}_t$ can be characterized by the following Kalman equation

$$\widehat{m}_{t+1} = (1-\theta)\widehat{m}_t + \theta(m_{t+1} + \xi_{t+1}) \tag{3.30}$$

where θ and ξ_{t+1} have the same definitions as in section 2. Hence, consumption growth can now be written as

$$\Delta c_{t+1} = H_1\{[\theta\zeta_{t+1} + \theta\eta_w \frac{(1-\theta)\zeta_t}{1-(1-\theta)\eta_w \cdot L}] + [\theta\xi_{t+1} - \frac{\theta\xi_t}{1-(1-\theta)\eta_w \cdot L}] + \Omega\}$$

where $\zeta_{t+1} = \chi r^e_{t+1} + \lambda v_{t+1}$, L is the lag operator and Ω is constant term. Furthermore, for simplicity, we assume that the endogenous noise terms in the second bracket of the above expression will be cancelled out when aggregating over all consumers[60]. Consequently, we have the following proposition

Proposition 5. *Aggregate consumption growth in the CCAPM model with RI can be written as*[61]

$$\Delta c_{t+1} = H_1[\theta\zeta_{t+1} + \theta\eta_w \frac{(1-\theta)\zeta_t}{1-(1-\theta)\eta_w \cdot L} + \Omega], \tag{3.31}$$

which implies that 1) the contemporaneous covariance between aggregate consumption growth and asset returns is

$$Cov(\Delta c_{t+1}, r^e_{t+1}) = H_1\theta\chi\omega_u^2 + H_1\theta\lambda\omega_{uv}, \tag{3.32}$$

2) the ultimate covariance between consumption growth and asset return is

$$Cov[\lim_{S\to\infty}(c_{t+1+S} - c_t), r^e_{t+1}] = H_1 \frac{\theta}{1-(1-\theta)\eta_w}\chi\omega_u^2 + H_1\theta\lambda\omega_{uv}, \tag{3.33}$$

and 3) the optimal asset allocation is

$$\chi = \vartheta_2[\frac{\pi}{\gamma H_1 \omega_u^2} - \frac{\theta\lambda\omega_{uv}}{\gamma\omega_u^2}] \tag{3.34}$$

where $\vartheta_2 = \frac{1-(1-\theta)\eta_w}{\theta} < 1$.

Proof. (see Appendix B). ¥

[60] As we discussed in section 2, this is an extreme case, and actually the individual noise terms may include a common component.

[61] *Since we focus on aggregate behavior and to avoid the notation confusion, in the following equation, we still use c to represent aggregate consumption.*

Here $\vartheta_2 = \frac{1-(1-\theta)/\alpha_4}{\theta} < 1$ is inversely proportional to the ultimate consumption risk of risky asset. The term in the bracket represents the optimal allocation to the risky asset that has two components. The first characterizes the optimal allocation when labor income risk is uncorrelated with the risky asset, while the second component is an income hedging demand component. The desirability of the risky asset depends not only on its expected excess return relative to its variance, but also on its ability to hedge consumption against bad realizations of labor income[62].

In Viceira (2001) and Campbell and Viceira (2003), they analyzed optimal portfolio choices of long-horizon investors with undiversifiable labor income risk and showed that positive correlation between labor income innovations and unexpected asset returns reduces the investor's willingness to hold the risky asset because the risky asset provides poor hedge against unexpected declines in labor income. Since $\vartheta_2 < 1$ and $\theta < 1$, we can see from above that if we measure consumption risk correctly, RI reduces not only the standard optimal allocation in the risky asset but also the hedging component. One percent negative shock in both individuals' financial wealth and human wealth would affect their consumption more than that predicted by the full-information model. Hence, the hedging demand is reduced.

4. Extension to the Risk-sensitive LQG PIH Model

Many consumption and saving empirical studies found that consumers are risk averse and precautionary savings motive is one of the basic motives for savings. Hence, it is worthwhile to extend and generalize the earlier LQG PIH model by introducing the preference for risk sensitivity that allows us to address the interactions among RI, risk aversion and labor income uncertainty. Specifically, in this section, we introduce risk-sensitivity into the LQG PIH model with RI via an additional single risk sensitivity parameter γ. Furthermore, this feature allows us disentangle risk aversion from intertemporal substitution so we can solely evaluate the effects of risk aversion on consumption and savings[63]. In particular, we can discuss the effects of RI on the impacts of risk aversion and income uncertainty on precautionary savings and the MPC out of income and wealth. One feature of the LEQG model is that the traditional certainty equivalence (CE) no longer holds because the variance of fundamental shocks affects the optimal decision, while a *modified* certainty equivalence principle still holds. See Whittle (1981, 1982, 1990, 1996) for details. This feature will play an important role in our analysis.

In this model, the utility index is an exponential transformation of a LQG time separable felicity function[64]. The main advantage of the LEQG framework is that it not only remains the appealing properties of LQG framework like the *modified* certainty equivalence principle, but also introduces precautionary savings motive by the exponential transformation. The LEQG problems are widely studied in automatic control and engineering literatures. This extension to the LEQG framework was started by Jacobson (1973, 1977) and extended by Whittle (1981, 1982, 1990, 1996) and others. Ploeg (1993) adopted this framework to seek closed-form solutions in a model with precautionary savings. Recently, it was also applied in macroeconomics

[62] Note that if this covariance is negative, then the risky asset offers a good hedge against negative income shock and then increases the demand for the risky asset.

[63] Tallarini (2000) adopted a risk-sensitive RBC model and showed that increasing risk aversion and holding the intertemporal elasticity of substitution constant works to raise the market price of risk and lower the risk-free rate simultanenously.

[64] There are another ways to introduce precautionary savings motive in the LQG PIH models. For example, Caballero (1990) and Weil (1993) study precautionary savings in the CARA models; and Carroll (1997), Gourinchas and Parker (1999, 2001), and so on adopt CRRA preference to address precautionary savings. However, in those models with CARA or CRRA preferences and information capacity constraints, minimizing the objective functions subject to these constraints does not imply that the conditional distribution of W_t given information available at time t is $N(\widehat{W}_t, \Sigma_t)$.

by Hansen, Sargent, and their coauthors. See Hansen and Sargent (1995), Hansen, Sargent, and Tallarini (1999), and Tallarini (2000) for details. This paper, in contrast, pursues to solve a LEQG PIH model with RI. The closed-form solutions derived from our model can show that incorporating RI will induce state-estimation risk and then affect precautionary savings, the MPC out of income and wealth, and welfare losses due to RI.

4.1. An Analytical Solution for The LEQG PIH Model with Rational Inattention

In this section, we solve the LEQG PIH model with RI analytically[65]. We consider an economy in which consumers have risk-sensitive preference. In particular, each consumer has the preference of the form

$$\widehat{V}_t(\widehat{S}_t) = \max_{C_t}\{\ln[E_t \exp u(C_t, S_t)] - \beta\frac{1}{\gamma}\ln E_t[\exp(-\gamma\widehat{V}_{t+1}(\widehat{S}_{t+1}))]\} \tag{4.1}$$

where C_t and S_t are controls and states, respectively, $\widehat{S}_s = E_t(S_t)$, $\Sigma_s = Var_t(S_t)$, γ is the parameter for the risk-sensitive preference, and β is the discount factor. Particularly, in most preference-based models, the state variable does not enter the utility function. Given that C_t is known at time t, the above Bellman equation reduces to

$$\widehat{V}_t(\widehat{S}_t) = \max_{C_t}\{u(C_t) - \beta\frac{1}{\gamma}\ln E_t[\exp(-\gamma\widehat{V}_{t+1}(\widehat{S}_{t+1}))]\} \tag{4.2}$$

In this case, risk aversion is measured by the parameter γ. When $\gamma = 0$, the preference collapses into the familiar expected utility case, and when $\gamma > 0$ the consumer is more risk averse relative to the expected utility case. Furthermore, given the LQ felicity function, $u(C_t) = C_t - \frac{1}{2}C_t^2$, the elasticity of intertemporal substitution is $\frac{1-C_t}{C_t}$. Note that this representation is similar to the value function derived in Hansen et. al (1999) except that here the expectation is formed under information processing constraints.

Note that in this LEQG framework, the loss function is defined as

$$\min_{q(Y|X)} E[\exp(Y - X)^2]$$

subject to the information constraints. Y is our action and X is a random variable that must be observed through a finite channel capacity. $q(Y|X)$ is the optimal conditional distribution of $Y|X$, subject to the information constraint? See Sims (section 3, 2003) for details about optimization with information-flow constraints under the LQ framework. The main difference between that model with ours is that now the objective function (loss function) is an exponential transform of the LQ function. We will show that under this specification, the optimal conditional distribution is still Gaussian.

Define a new variable $Z = (Y - X)^2$, and $\overline{Z}$ is the steady state value. Suppose that Z fluctuates around its steady state such that local linearization around steady state is valid, we can approximate the above

[65] For the details about solving the general LEQG model with RI, see the Technical Appendix for this paper which is available from the author's homepage: http://www.princeton.edu/~yluo.

optimization problem as follows

$$\begin{aligned}\min E[\exp(Y-X)^2] &\simeq \min E[\exp(\overline{Z})+\exp(\overline{Z})(Z-\overline{Z})] \\ &= \min \exp(\overline{Z})E[(Y-X)^2] + \text{Constants} \\ &= \min E[(Y-X)^2] \text{ when } \overline{Z}=0\end{aligned}$$

Hence, with an exponential-quadratic loss function, if a local approximation is valid, minimizing the loss function under the information processing constraint means that the $N(\widehat{S}_t, \Sigma_t)$ is an approximation of the true condition distribution of the state given information at time t.

Using the risk-sensitive LQG control methods used in Whittle (1982, 1992) and Hansen and Sargent (1995), in the separate Technical Appendix, we show that the consumption decision rule[66] in the PIH model is[67]

$$C_t = K \cdot \overline{M}_t + k \tag{4.3}$$

where

$$\begin{aligned} K &= \frac{R^2-1/\beta}{R^2-\gamma\omega^2},\ k = (1+\frac{R^2-1/\beta}{1/\beta-\gamma\omega^2})^{-1}[1+\frac{R^2-1/\beta}{1/\beta-\gamma\omega^2}(\frac{\overline{Y}^T}{R-1}-\frac{1}{R-1})],\ \text{and} \\ \overline{M}_t &= (\gamma\Sigma\Pi+I)^{-1}\widehat{M}_t + (\gamma\Sigma\Pi+I)^{-1}\gamma\Sigma\Pi M^* \end{aligned}$$

where Π and M^* are the steady state matrices in the future stress $V_s(M_s) = -\frac{1}{2}(M_s - M_s^*)'\Pi_s(M_s - M_s^*)$ (see Technical Appendix for details). For simplicity, we assume that $R\beta = 1$,thus the above expressions can be simplified as follows

$$K = \frac{R(R-1)}{R^2-\gamma\omega^2} \text{ and } k \simeq (1+\frac{R(R-1)}{R-\gamma\omega^2})^{-1}\overline{Y}^T \tag{4.4}$$

To compare this consumption function with the one from the LQG PIH model, we rewrite (4.3) as follows

$$C_t = \widetilde{K} \cdot \widehat{M}_t + \widetilde{k} \tag{4.5}$$

where

$$\widetilde{K} = \frac{R(R-1)}{R^2-\gamma\omega^2}\frac{1}{1+\gamma\sigma^2R(R-1)/(R^2-\gamma\omega^2)} \text{ and} \tag{4.6}$$

$$\widetilde{k} = \frac{R(R-1)}{R^2-\gamma\omega^2}\frac{1}{1+1/[\gamma\sigma^2R(R-1)/(R^2-\gamma\omega^2)]} + (1+\frac{R(R-1)}{R-\gamma\omega^2})^{-1}Y^T \tag{4.7}$$

where σ^2, $\widehat{M}_t$, and M_t have been defined in section 2.

Given the above expressions for $\widetilde{K}$ and $\widetilde{k}$, it is obvious that RI (measured by σ^2) affects precautionary savings via the first term in the expression of $\widetilde{k}$ and the whole expression of $\widetilde{K}$. Note that when $\gamma = 0$ (the expected utility case), the consumption function reduces to $C_t = \frac{R-1}{R}\widehat{M}_t + \frac{1}{R}Y^T$ that is just the consumption rule derived in section 2, when $\kappa = \infty$ (no information processing constraint), it reduces to

[66] Here we consider the same model setup as described in Section 2.1, except the presence of the preference for risk sensitivity. The notations used here are the same ones as those we used in section 2.

[67] In steady state, Σ_t converges to constant Σ.

that derived from the standard LEQG model with perfect observation, and when $\gamma = 0$ and $\kappa = \infty$, it reduces to the standard consumption function $C_t = \frac{R-1}{R}(W_t + \frac{1}{R-1}Y_t^P) + \frac{1}{R}Y^T$ as the standard LQ PIH model predicts. In the LEQG model without RI, income uncertainty can increase precautionary savings premium by reducing k and increase the MPC out of wealth by increasing K[68], while in the LEQG model with RI, RI does reduce the MPC out of *perceived* state $\widehat{M}_t$ by increasing the steady state conditional variance σ^2. Intuitively, the reason that channel capacity can raise the slope of the consumption function is that an extra bit used in observing state will loosen the consumer's information processing constraint, as a result, reduce the conditional variance σ^2 and thus encourage extra consumption. With higher inattention, consumption may be less sensitive to the short-term wealth fluctuation from income shock. Gabaix and Laibson (2001) discussed how fixed cost of monitoring wealth can reduce the reaction of consumption with respect to the short-run wealth fluctuation, and in our framework, it is also the case: the consumers with low channel capacity means that their monitoring and observing costs might be higher than those with high channel capacity in observing.

Next, as usual, we can rewrite the PIH model in terms of savings rather than consumption. In the LEQG PIH model with RI, we also define the expression for savings as $D_t = \frac{R-1}{R}W_t + \frac{1}{R}Y_t - C_t = \frac{R-1}{R}M_t + \frac{1}{R}Y_t^T - C_t$. Substituting the consumption function $C_t = \widetilde{K} \cdot \widehat{M}_t + \widetilde{k}$ into this saving equation yields

$$
\begin{aligned}
D_t = \{ & \frac{R-1}{R}(M_t - \widehat{M}_t) + [\frac{R-1}{R} - \frac{R(R-1)}{R^2 - \gamma\omega^2}\frac{1}{1 + \gamma\sigma^2 R(R-1)/(R^2 - \gamma\omega^2)}]\widehat{M}_t\} + \frac{1}{R}Y_t^T \\
& - \{\frac{R(R-1)}{R^2 - \gamma\omega^2}\frac{1}{1 + 1/[\gamma\sigma^2 R(R-1)/(R^2 - \gamma\omega^2)]} + (1 + \frac{R(R-1)}{R - \gamma\omega^2})^{-1}Y^T\}
\end{aligned} \tag{4.8}
$$

In contrast to the standard CARA model, it is obvious that here precautionary saving depends not only on the constant term but also on the MPC out of *perceived* state. Given the saving function (4.8), there is no clear conclusion about if RI would increase or decrease precautionary savings. For simplicity, here we first focus on the effects of RI and ignore the term $\gamma\omega^2$ since it is small relative to R^2 and thus only has marginal contribution to the saving function. Hence, the above saving function becomes

$$
D_t = \frac{R-1}{R}(M_t - \widehat{M}_t) + \frac{1}{R}(Y_t^T - Y^T) + \frac{R-1}{R}\frac{1}{1 + 1/[\gamma\sigma^2(R-1)/R]}(\widehat{M}_t - 1) \tag{4.9}
$$

Since savings in the LQG PIH model with RI is $D_t = \frac{R-1}{R}(M_t - \widehat{M}_t) + \frac{1}{R}(Y_t^T - Y^T)$, it is clear that RI has two opposite effects on savings: one is from reducing the constant term by $\frac{R-1}{R}\frac{1}{1+1/[\gamma\sigma^2(R-1)/R]}$ in the saving function and the other is from adding an extra term $\frac{R-1}{R}\frac{1}{1+1/[\gamma\sigma^2(R-1)/R]}\widehat{M}_t$. The second effect is via the MPC out of perceived state, which does not emerge in the standard CARA model. The intuition behind this is simple: RI reduces the MPC out of perceived wealth because of the delay and smoothness of consumers' reaction to "news", consequently, for given current total wealth, saving at current period will increase. The intuition is that the less attention the consumer used in observing the state, the larger is his exposure to labor income risk, in that he may not react to income shocks immediately when they occur. This larger risk induces higher precautionary savings because the consumer need to save more to self-insure himself against a sequence of bad income shocks[69]. As for the net impact of RI on savings, we have the

[68] This conclusions is consistent with what Ploeg (1993) found, though it is not what the standard CARA models predict in that they predict that the MPC out of wealth is independent of the variance of labor income and risk aversion parameter. See Caballero (1990) and Weil (1993) for details.

[69] This mechanism of the effects of RI on savings is similar to the one discussed in Reis (2003) where it is shown that if the

following proposition.

Proposition 6. *Whether RI increases precautionary savings or not depends on if* $\widehat{M}_t > 1$ *or not.*

The proof is straightforward. Note that since $M_t = W_t + \frac{1}{R-1}Y_t^P$ and $\widehat{M}_{t+1} = \widehat{M}_t + \theta R\frac{(1-\theta)\zeta_t - \theta\xi_t}{1-(1-\theta)R\cdot L} + \theta(\zeta_{t+1} + \xi_{M,t+1})$, $\widehat{M}_t$ is a random walk. Conditional on some initial value $\widehat{M}_0$, the conditional mean of $\widehat{M}_t$ is $E_0[\widehat{M}_t] = \widehat{M}_0$.

Next, we will do some comparative statics analysis to explore the effects of RI, risk aversion, and labor shock on consumption and savings dynamics. We found that in the presence of RI, the effects of labor income on both MPC and precautionary savings are altered. The following proposition summarizes the results.

Proposition 7. *With rational inattention, the effects of labor income uncertainty* ω_ζ^2 *and risk aversion parameter* γ *on the MPC out of perceived wealth and savings are reduced, that is, for a given income shock, the consumer with low attention in monitoring and observing his wealth evolution would have lower MPC and higher savings if certain condition is satisfied.*

Proof: The proof is straightforward. Note that the parameter of risk aversion γ and labor uncertainty ω^2 affect the consumption function in the same way. The effects of RI are reflected by a factor $\frac{1}{1+\gamma\sigma^2 R(R-1)/(R^2-\gamma\omega^2)}$ in the $\widetilde{K}$ term and a factor $\frac{1}{1+1/[\gamma\sigma^2 R(R-1)/(R^2-\gamma\omega^2)]}$ in the $\widetilde{k}$ term. Since $\sigma^2 = \frac{\omega^2}{\exp(2\kappa)-R^2}$, lower κ that represent less attention reduces the first factor and increases the second factor.

Furthermore, it is also interesting to see under what conditions for information processing channel capacity, the impacts of risk aversion and labor income uncertainty on the consumption function will be altered. The following proposition summarize the results.

Proposition 8. *RI can affect the magnitude of the impacts of risk aversion and income uncertainty on both precautionary savings and the MPC out of permanent income and wealth but cannot reverse the effects for a plausible value of channel capacity.*

Proof: From the consumption function, it is straightforward to obtain

$$\frac{d(\widetilde{K})}{d\omega^2} = \begin{cases} > 0 \text{ if } \kappa > \frac{3}{2}(R-1) \\ = 0 \text{ if } \kappa = \frac{3}{2}(R-1) \\ < 0 \text{ if } \kappa < \frac{3}{2}(R-1) \end{cases}, \frac{d(\widetilde{K})}{d\gamma} = \begin{cases} > 0 \text{ if } \kappa > \frac{3}{2}(R-1) \\ = 0 \text{ if } \kappa = \frac{3}{2}(R-1) \\ < 0 \text{ if } \kappa < \frac{3}{2}(R-1) \end{cases}$$

Since the interest rate $R-1$ is a small value, RI will reduce the impacts of risk aversion and income shock on the MPC out of income and wealth only if consumers use very small portion of his channel capacity in monitoring and observing his wealth dynamics. Since this value is quite below the usual capacity people use in monitoring their financial affairs, it is reasonable to believe that RI does not reverse the impacts of γ and ω^2 on the MPC and just smooth the reactions of consumption.

4.1.1. The Impulse Responses Function of Consumption

In this subsection, we will analyze how risk aversion affect the dynamic behavior of consumption and wealth in the presence of information constraints. To do that, we need to assemble the consumption decision rule,

consumer faces higher planning costs and remains inattentive longer, he will save more.

the dynamic budget constraint, and the Kalman equation used in characterizing the evolution of $\widehat{W}_t$ to form a three-dimension system of difference equations:

$$C_t = \widetilde{K} \cdot \widehat{M}_t + \widetilde{k} \tag{4.10}$$

$$M_{t+1} = R(M_t - C_t) + \zeta_{t+1} + \overline{Y}^T \tag{4.11}$$

$$\widehat{M}_{t+1} = \widehat{M}_t + \theta(M_{t+1} + \xi_{t+1} - \widehat{W}_t) \tag{4.12}$$

where $\zeta_{t+1} = \frac{R}{R-1}\varepsilon_{t+1} + \eta_{t+1}$ and $M_t = W_t + \frac{1}{R-1}Y_t^P$. Figure 11 below plots the responses of consumption with respect to two components in income and one error shock, with channel capacity 1. The two dotted horizontal lines accompanied with the the responses to both permanent and transitory shocks represent the levels of the flat responses of consumption in the absence of information capacity constraints. We can see that all properties of IRFs found in the LQG PIH model remain unchanged here: consumption responds slowly and gradually to income shocks and responds immediately to the information processing error.

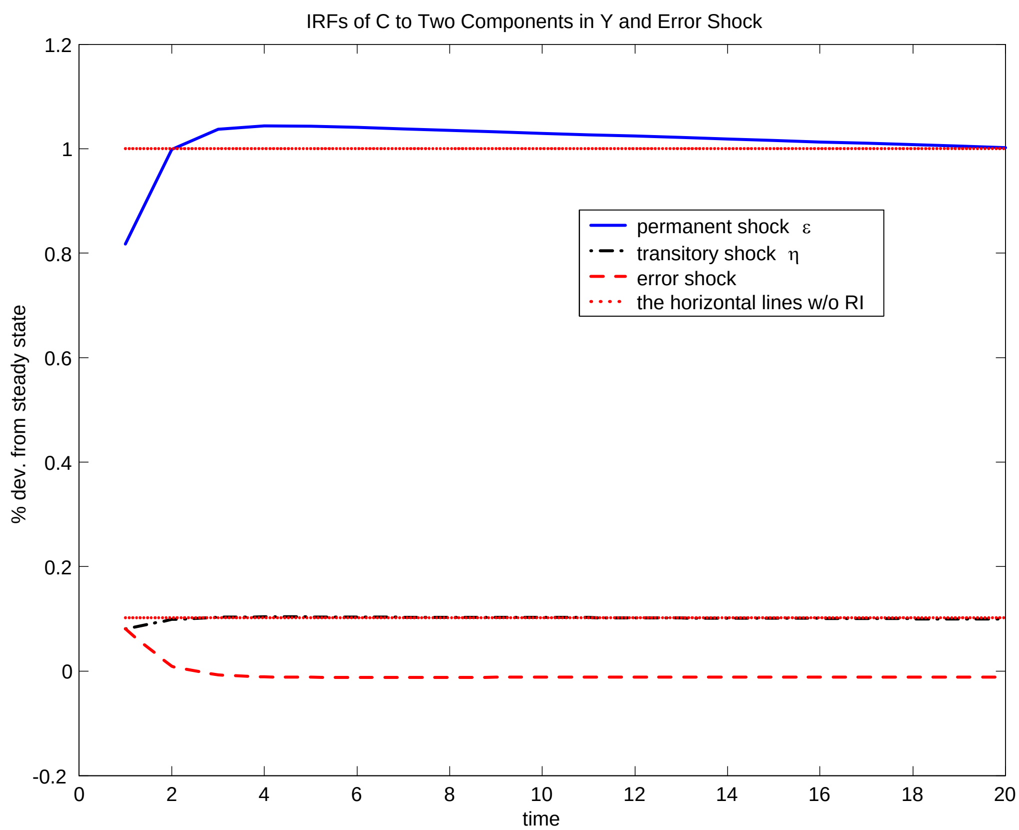

Figure 11

4.1.2. The Excess Smoothness and Excess Sensitivity Puzzles Revisited

In this subsection, we will revisit two puzzles in the consumption literature: the excess smoothness puzzle and the excess sensitivity puzzle. Before we examine how RI affects the change in consumption in this model, it is helpful to see how risk-sensitivity and income uncertainty affect consumption growth in the absence of

RI first. Using the consumption function derived from the LEQG PIH model, we have

$$\begin{aligned}\Delta C_t &= K\Delta M_t = K[(R-1)M_{t-1} - RC_{t-1}] + K\zeta_t + K\overline{Y}^T \\ &= [(R-1) - RK]C_{t-1} + K\zeta_t + K\overline{Y}^T - KRk\end{aligned} \tag{4.13}$$

where $K = \frac{R(R-1)}{R^2-\gamma\omega^2}$ and $k = (1 + \frac{R(R-1)}{R-\gamma\omega^2})^{-1}$. As in Ploeg (1993), the growth of consumption depends on both the "news" about income, ζ_t, and past consumption C_{t-1}. Furthermore, given the initial value of state variable M_0, we can express past consumption C_{t-1} as follows

$$C_{t-1} = K\sum_{j=t-1}^{1} \zeta_{t-j} + K \cdot M_0 + k \tag{4.14}$$

where we need to use the fact that $M_t = M_{t-1} + \zeta_t$.

Thus, in this case, consumption no longer follows a random walk. Note that when $\gamma = 0$, this impact disappears. Consumption is also sensitive to past income innovations, which is obvious from the expression of C_{t-1}. Hence, although consumption reacts to both current and past shocks in this LEQG model, the mechanism and the magnitudes of responses to income shocks are quite different from the RI models[70]. Since $K > \frac{R-1}{R}$, it follows that the MPCs out of both the permanent shock and the i.i.d. shock are larger than the ones obtained by the standard LQG PIH model. In other words, risk sensitivity increases the sensitivity of changes in consumption to unanticipated changes in income. This sensitivity of consumption is usually called "making hay while the sun shines" and is consistent with the empirical evidence reported in Flavin (1981). However, Campbell and Deaton (1989) argued that there is excess sensitivity of consumption to anticipated changes in income and too little sensitivity to unanticipated changes in income. They then suggested that this might be one of the reasons why consumption is too smooth. Hence, the presence of "making hay while the sun shines" makes it more difficult to explain the puzzle of why consumption is so smooth. Intuitively, introducing RI can reduce the sensitivity of consumption to unanticipated permanent and transitory income shocks and make consumption more smooth. Hence, incorporating RI into the standard LEQG model could also make the model be more reconcile with the empirical evidence. To see it clearly, we can express change in consumption in the RI model as follows

$$\Delta C_t = \widetilde{K} \cdot \Delta\widehat{M}_t = (R - 1 - R\widetilde{K})C_{t-1} + \theta R\widetilde{K}\frac{(1-\theta)\zeta_{t-1} - \theta\xi_{t-1}}{1-(1-\theta)R \cdot L} + \theta\widetilde{K}(\zeta_t + \xi_t) + \text{constants} \tag{4.15}$$

where we use the formula $M_t - \widehat{M}_t = \frac{(1-\theta)\zeta_t - \theta\xi_t}{1-(1-\theta)R\cdot L}$. Hence, this expression shows that RI reduces the response of consumption to contemporaneous income shock by a factor θ but increases the responses to past income shocks. Following the same procedure used in section 2, it is straightforward to show that introducing RI in the LEQG model could better explain the empirical evidence about consumption smoothness.

[70] In this model, the coefficients attached to past shocks are same for all past periods, while the coeffiecients obtained from the LQG PIH model with RI are decreasing over time. Furthermore, the magnitude in the LEQG model is $[(R-1) - RK]K$, which is close to zero.

4.1.3. Welfare Analysis

In this subsection, we evaluate how income uncertainty, risk aversion, and RI affect consumers' intertemporal welfare. We begin with the following Bellman equation characterizing the LEQG PIH model with RI:

$$\widehat{V}(\widehat{M}_t) = \max_{C_t}\{u(C_t) - \frac{1}{\gamma}\ln E_t[\exp(-\gamma\beta\widehat{V}(\widehat{M}_{t+1}))]\} \tag{4.16}$$

which can be rewritten in a more familiar form:

$$\widehat{V}(\widehat{M}_t) = \max_{C_t}\{u(C_t) + \beta[-\frac{1}{\widetilde{\gamma}}\ln[E_t(\exp(-\widetilde{\gamma}\widehat{V}(\widehat{M}_{t+1})))]]\} \tag{4.17}$$

where $\widetilde{\gamma} = \gamma\beta$ and $-\frac{1}{\widetilde{\gamma}}\ln[E_t(\exp(-\widetilde{\gamma}\widehat{V}(\widehat{M}_{t+1})))]$ is the risk-adjusted next period's value function. Note that since

$$-\frac{1}{\widetilde{\gamma}}\ln[E_t(\exp(-\widetilde{\gamma}\widehat{V}(\widehat{M}_{t+1})))] \simeq E_t\widehat{V}(\widehat{M}_{t+1}) - \frac{1}{2}\widetilde{\gamma}Var_t\widehat{V}(\widehat{M}_{t+1}), \tag{4.18}$$

the above Bellman equation reduces to[71]

$$\widehat{V}(\widehat{M}_t) = \max_{C_t}\{u(C_t) + \beta[E_t(\widehat{V}(\widehat{M}_{t+1})) - \frac{1}{2}\widetilde{\gamma}Var_t(\widehat{V}(\widehat{M}_{t+1}))]\} \tag{4.19}$$

As usual, we guess that $\widehat{V}(\widehat{M}_t) = E_0 + E_1\widehat{M}_t + E_2\widehat{M}_t^2$and after substituting the optimal consumption rule: $C_t = \widetilde{k} + \widetilde{K}\cdot\widehat{M}_t$ where $\widetilde{k}$ and $\widetilde{K}$ are defined above and are both dependent on the properties of income uncertainty and endogenous noise as well as risk aversion. Substituting them into the Bellman equation, we can pin down the undetermined coefficients by matching as follows:

$$E_1 = \frac{\widetilde{K}(1-\widetilde{k})}{1-\beta}, E_2 = -\frac{\frac{1}{2}\widetilde{K}}{1-\beta}, \text{ and } E_0 = \frac{\widetilde{k} - \frac{1}{2}\widetilde{k}^2 + \beta(E_2 - \frac{1}{4}\widetilde{\gamma}E_1^2)[(R^2-1)\sigma^2 + \omega^2]}{1-\beta}$$

where we assume that $Var_t(\widehat{M}_{t+1}^2) = 0$ since this term is a fourth moment of variable $\widehat{M}_{t+1}$.

We can now analyze the effect of RI on the welfare of consumers starting from the initial period 0. The constant term E_0 implies the impacts of RI on the welfare if given the initial estimated state $\widehat{M}_0$. Starting from period 0, the value function is

$$\widehat{V}(\widehat{M}_0) = E_0(\omega^2, \kappa, \gamma) + E_1\widehat{M}_0 + E_2\widehat{M}_0^2$$

which measures the level of intertemporal welfare of the consumer when the optimal path under information processing constraints is followed. For simplicity, we assume that the initial information state $\widehat{M}_0$ is given exogenously. Since the E_0 is a complicated function of the key parameters we are interested in, we need to use a numerical example to illustrate the impacts of risk aversion, income uncertainty, and RI on the welfare. In this example, we also set $R = 1.01$, $\beta = 1/R$, and assume that Y_t follows a uniform distribution over $[0, 2]$ with mean 1[72], and then explore the impacts of risk aversion and income uncertainty on the consumers with limited channel capacity[73].

[71] Hence, the LEQG assumption corresponds locally to the mean-variance approach.
[72] Note that in this case, the steady state wealth $W \leq 1$ since $C \leq C_{bliss} = 1$.
[73] In this example, the pattern of the effects of risk aversion and income uncertainty on the consumers with RI does not change

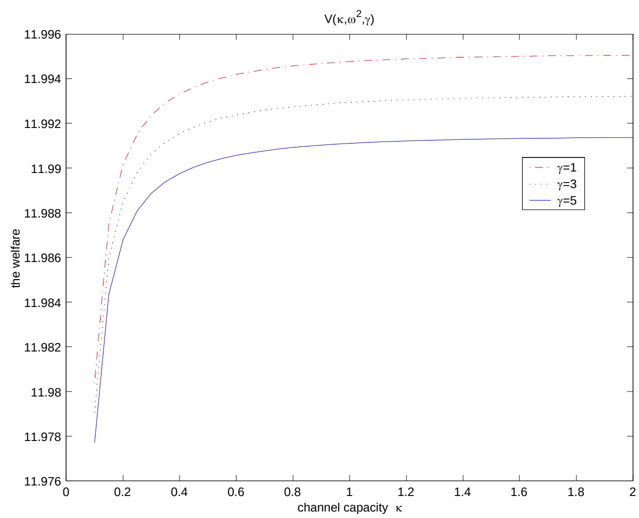

Figure 12

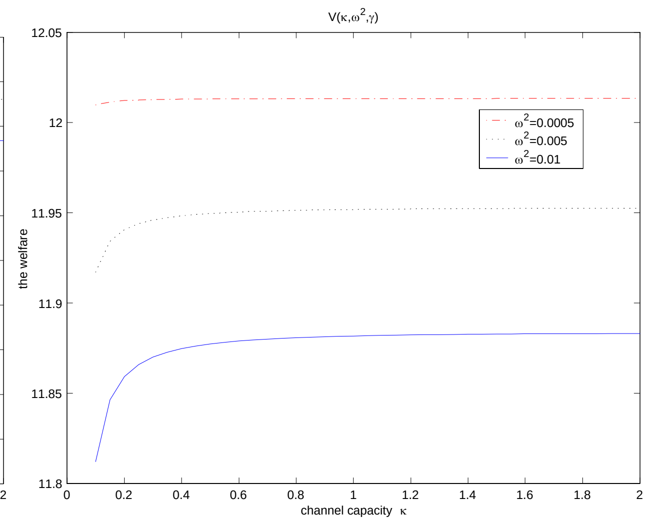

Figure 13

Figure 12 shows that for any variance of income shocks, intertemporal welfare of the risk averse consumer increases with channel capacity because low capacity means high conditional variance σ^2. Furthermore, it is obvious from the figure that for given κ the welfare of the consumers with low risk aversion is higher than the welfare of the consumers with high risk aversion. The intuition is simple: the consumers with higher risk aversion dislike uncertainty induced by both the fundamental shocks to income and the conditional variance of the state that cannot be observed perfectly due to RI more than the consumers with lower risk aversion. Figure 13 shows that for any value of γ, intertemporal welfare of the risk averse consumer increases with channel capacity because low capacity means high conditional variance σ^2. Furthermore, for the given level of κ, higher labor income uncertainty reduces welfare since the consumer is risk averse, which is also predicted by the standard PIH models.

5. Conclusions

In this paper, we examined the effects of information processing constraint on the intertemporal allocation of consumption, aggregate dynamics, asset pricing and asset allocation, and welfare in the PIH model and the CCAPM model. It is shown that incorporating RI into these standard model can generate different individual and aggregate dynamics and make the models better explain the data in several aspects. Specifically, in the context of the LQ PIH framework, introducing RI can generate (1) the gradual responses of consumption to various shocks to wealth, (2) the excess sensitivity of consumption to lagged shocks, and (3) the excess smoothness of aggregate consumption. Furthermore, we compared this model with the habit-formation model and found that both model can generate similar consumption, savings, and wealth dynamics. We then analyzed welfare consequences due to RI and found that the welfare losses due to deviating from the first-best instantly adjusted consumption path is very tiny. Finally, we discussed endogenizing channel capacity. In the context of the CCAPM model, we found that introducing RI can generate consumption smoothness, the low contemporaneous correlation between consumption growth and asset returns, as well as

with the reasonable steady state wealth level.

high equity risk premium measured by long-term consumption risk. Furthermore, it is shown that consumers with low attention would be invested less in risky assets. In the final substantive section of this paper, we solved a risk-sensitive LQ PIH model with RI and discussed the interaction among risk-sensitivity, labor income risk, and RI.

Given these findings, it seems promising to incorporating RI into other frameworks, e.g., the monetary model or sophisticated general equilibrium model. Specifically, (1) it is well-known that in the baseline New Keynesian monetary models, the dynamics of output and inflation that are characterized by the "IS" equation and the "New Keynesian" Phillips curve do not depend on any lagged variables but entirely depend on expectations of future variables, monetary policy and exogenous force processes, that is, they are completely forward-looking. However, as have been documented in many empirical studies, these forward looking models can not match some important features of the data. For example, they are unable to replicate the high serial correlations found in both output and inflation. Hence, since we have demonstrated in this paper that RI can generate the persistence of wealth shocks endogeneously in the PIH model and the CCAPM model, it may be a potential explanation for the persistence problem in the monetary models. (2) Since RI largely affects the optimal intertemporal allocation of consumption, it is also worthwhile examining if incorporating RI into a Heterogeneous-agent Dynamic Stochastic General Equilibrium (DSGE) model can help us answer some important questions in the literature. For example, can RI play a role in shaping the skewed wealth distribution in the U.S.? or can RI increase aggregate savings in equilibrium? We know that the simple heterogeneous-agent DSGE models cannot account for the very skewed wealth distribution in the U.S. (e.g., Aiyagari1994, Krusell and Smith, 1998). We leave these interesting topics for future research.

6. Appendix

6.1. Appendix A: Derivation for Optimal Allocation of Capacity

Instead of minimizing the criterion $E_t V(S_t) - \widehat{V}(\widehat{S}_t)$ subject to information constraints as proposed in Sims (2003), here we propose another procedure to solve a multivariate-state model. The procedure is as follow. First, we guess the value function from the RI model as follows:

$$\widehat{V}(\widehat{S}_t) = F_0 + F_1 \widehat{S}_t + \widehat{S}_t' F_2 \widehat{S}_t$$

and the state evolution equation is

$$S_{t+1} = G_0 + G_1 S_t + G_2 C_t + \varepsilon_{t+1}$$

where S_t is state variable and C_t is control variable. In the presence of RI, we have the following Bellman equatin:

$$\widehat{V}(\widehat{S}_t) = E_t u(C_t, S_t) + \beta E_t \widehat{V}(\widehat{S}_{t+1})$$

Substituting the guessed value function and the consumption function into this Bellman equation yields

$$F_0 + F_1\widehat{S}_t + \widehat{S}_t' F_2 \widehat{S}_t = (H_0 + H_1\widehat{S}_t) - \frac{1}{2}(H_0 + H_1\widehat{S}_t)'(H_0 + H_1\widehat{S}_t)$$
$$+ \beta[F_0 + F_1(G_0 + G_2H_0 + (G_1 + G_2H_1)\widehat{S}_t) + E_t(\widehat{S}_{t+1}' F_2 \widehat{S}_{t+1})]$$

where
$$E_t(\widehat{S}_{t+1}' F_2 \widehat{S}_{t+1}) = (G_0 + G_1\widehat{S}_t + G_2C_t)' F_2 (G_0 + G_1\widehat{S}_t + G_2C_t) + tr(F_2(\Psi_t - \Sigma_{t+1}))$$

Hence, for $\widehat{S}_t'(\cdot)\widehat{S}_t$ term:

$$F_2 = -\frac{1}{2}H_1'H_1 + \beta(G_1 + G_2H_1)'F_2(G_1 + G_2H_1) \tag{6.1}$$

This is a standard Lyaponov equation and can be solved numerically by the usual way.

For $\widehat{S}_t$ term, we have

$$F_1 = H_1 - H_0'H_1 + \beta[F_1(G_1 + G_2H_1) + \frac{1}{2}(G_0 + G_2H_0)'F_2(G_1 + G_2H_1)] \tag{6.2}$$

Note that F_1 is $(1 \times ns)$ vector and given F_2, it can be solve from above.

For the constant term, we have

$$F_0 = H_0 - \frac{1}{2}H_0'H_0 + \beta[F_0 + F_1(G_0 + G_2H_0) + (G_0 + G_2H_0)'F_2(G_0 + G_2H_0) + tr(F_2\Omega) + tr((G_1F_2G_1' - F_2)\Sigma)] \tag{6.3}$$

Next, the procedure to derive the optimal steady state Σ is as follows:

$$\max_{\Sigma} E[V(S_t) - \widehat{V}(\widehat{S}_t)] \tag{6.4}$$

$$s.t.$$

$$\log_2|\Psi| - \log_2|\Sigma| \leq 2\kappa;\ \ \Psi \succeq \Sigma$$

where $E[V(S_t) - \widehat{V}(\widehat{S}_t)] = -\frac{1}{1-\beta}tr[(G_1F_2G_1' - F_2)\Sigma] + tr[F_2(Var(S_t) - Var(\widehat{S}_t))]$.

We can also solve $Var(S_t)$ and $Var(\widehat{S}_t)$ by the following procedure. The dynamic system for given optimal choice of Σ is as follows

$$S_t = G_0 + G_2H_0 + G_1S_{t-1} + G_2H_1\widehat{S}_{t-1} + \varepsilon_t \tag{6.5}$$
$$\widehat{S}_t = (I - \Theta)(G_1 + G_2H_1)\widehat{S}_{t-1} + \Theta(S_t + \xi_t) \tag{6.6}$$

Since $\widehat{S}_t = E(S_t|\digamma_t)$, we have $E(\widehat{S}_t) = E[E(S_t|\digamma_t)] = E(S_t)$. Next, we need to calculate the unconditional covariance matrix of $Var(\widehat{S}_t)$

$$\Gamma_0 \begin{bmatrix} S_t \\ \widehat{S}_t \end{bmatrix} = \Gamma_1 \begin{bmatrix} S_{t-1} \\ \widehat{S}_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \Theta\xi_t \end{bmatrix} + \text{Constant} \tag{6}$$

where $\Gamma_0 = \begin{bmatrix} I & 0 \\ -\Theta & I \end{bmatrix}$ and $\Gamma_1 = \begin{bmatrix} G_1 & G_2 H_1 \\ 0 & (I-\Theta)G_3 \end{bmatrix}$. Since Γ_0 is nonsingular, we can obtain the following reduced form:

$$\begin{bmatrix} S_t \\ \widehat{S}_t \end{bmatrix} = \Gamma_0^{-1}\Gamma_1 \begin{bmatrix} S_{t-1} \\ \widehat{S}_{t-1} \end{bmatrix} + \Gamma_0^{-1} \begin{bmatrix} \varepsilon_t \\ \Theta\xi_t \end{bmatrix} + \Gamma_0^{-1} \cdot \text{Constant} \tag{6.7}$$

Taking unconditional variance on both sides yields:

$$Var(\begin{bmatrix} S \\ \widehat{S} \end{bmatrix}) = (\Gamma_0^{-1}\Gamma_1) Var(\begin{bmatrix} S \\ \widehat{S} \end{bmatrix})(\Gamma_0^{-1}\Gamma_1)' + Var(\Gamma_0^{-1} \begin{bmatrix} \varepsilon_t \\ \Theta\xi_t \end{bmatrix}) \tag{8}$$

Thus, we can calculate the unconditional covariance matrix $Var(S_t)$ and $Var(\widehat{S}_t)$ by solving the above Lyapunov equation.

6.2. Appendix B

6.2.1. Deriving the Consumption Rule

Adding RI in the above model yields the following modified consumption rule

$$c_t = H_0 + \widehat{w}_t \tag{6.8}$$

and the information state $\widehat{w}_t$ can be characterized by the following Kalman equation

$$\widehat{w}_{t+1} = (1-\theta)\widehat{w}_t + \theta(w_{t+1} + \xi_{t+1}) \tag{6.9}$$

where θ and ξ_{t+1} have the same definitions as in section 2. Combining these two equations with the log-linearized flow budget constraint

$$\Delta w_{t+1} = r_{t+1}^m + \alpha_2 + (1 - 1/\alpha_1)(c_t - w_t), \tag{6.10}$$

we obtain

$$\begin{aligned} \Delta\widehat{w}_{t+1} &= \theta(w_{t+1} - \widehat{w}_t) + \theta\xi_{t+1} \\ &= (\theta/\alpha_1)(w_t - \widehat{w}_t) + (\theta\chi u_{t+1} + +\theta\xi_{t+1}) + \theta[(1-\chi)r^f + \alpha_2 + (1 - 1/\alpha_1)H_0] \end{aligned}$$

where $w_t - \widehat{w}_t = \frac{(1-\theta)\chi u_t - \theta\xi_t}{1-((1-\theta)/\alpha_1)\cdot L} + \Omega$ because

$$w_{t+1} - \widehat{w}_{t+1} = ((1-\theta)/\alpha_1)(w_t - \widehat{w}_t) + (1-\theta)\chi u_{t+1} - \theta\xi_{t+1} + (1-\theta)[(1-\chi)r^f + \alpha_2 + (1-1/\alpha_1)H_0]$$

where L is the lag operator and Ω is an irrelevant constant term.

Hence, consumption growth can now be written as

$$\Delta c_{t+1} = \Delta \widehat{w}_{t+1} = [\theta\chi u_{t+1} + \theta\chi \frac{((1-\theta)/\alpha_1)u_t}{1-((1-\theta)/\alpha_1)\cdot L}] + [\theta\xi_{t+1} - \frac{\theta\xi_t}{1-((1-\theta)/\alpha_1)\cdot L}] + \Omega \tag{6.11}$$

where L is the lag operator and Ω is the constant term. If we assume that the endogenous noise terms in the second bracket of the above expression will be cancelled out when aggregating over all consumers. Consequently, we can have proposition 3 directly.

6.2.2. Deriving the Consumption Rule in the Presence of Labor Income Risk

We first divide the flow budget constraint (3.22) by Y_{t+1} and log-linearize it around $c - y = E[c_t - y_t]$ and $w - y = E[w_t - y_t]$ as follows

$$w_{t+1} - y_{t+1} \simeq \eta + \eta_w(w_t - y_t) - \eta_c(c_t - y_t) - \Delta y_{t+1} + r^m_{t+1} \tag{6.12}$$

where

$$\begin{aligned} \eta_w &= \frac{\exp(w-y)}{1+\exp(w-y)-\exp(c-y)} > 0,\ \eta_c = \frac{\exp(c-y)}{1+\exp(w-y)-\exp(c-y)} > 0, \\ \text{and } \eta &= -(1-\eta_w+\eta_c)\log(1-\eta_w+\eta_c) - \eta_w\log(\eta_w) + \eta_c\log(\eta_c) \end{aligned}$$

Second, to reduce this multivariate state case to the univariate state case, we need to define a new state variable that is a certain linear combination of w_t and y_t. Following the same procedure in section 2, we rewrite the log-linearized budget constraint as follows:

$$w_{t+1} + \lambda y_{t+1} = \eta + \eta_w(w_t + \lambda y_t) - \eta_c c_t + r^m_{t+1} - g + \lambda v_{t+1}$$

where $\lambda = \frac{1-\eta_w+\eta_c}{\eta_w - 1}$. Define the new state $m_t = w_t + \lambda y_t$, we have

$$m_{t+1} = \eta + \eta_w m_t - \eta_c c_t + r^m_{t+1} - g + \lambda v_{t+1} \tag{6.13}$$

Third, the log-linearized Euler equation is as follows

$$0 = \log\beta - \frac{1}{\sigma}E_t[c_{t+1} - c_t] + E_t[r^m_{t+1}] + \frac{1}{2}Var_t[r^m_{t+1} - \frac{1}{\sigma}(c_{t+1} - c_t)]. \tag{6.14}$$

Furthermore, we guess that the optimal log consumption rule take the following form $c_t = H_0 + H_1 m_t$. Hence, $c_{t+1} - c_t = H_1(m_{t+1} - m_t)$.

Combining it with the log-linearized budget constraint yields

$$\begin{aligned} E_t[c_{t+1} - c_t] &= H_1 E_t[m_{t+1} - m_t] \\ &= H_1[\eta + (\eta_w - 1 - \eta_c H_1)m_t - \eta_c H_0 + E[r^m_{t+1}] - g] \end{aligned} \tag{6.15}$$

The log-linearized Euler equation implies that

$$\begin{aligned}
E_t[c_{t+1}-c_t] &= \sigma\{\log\beta + E[r^m_{t+1}] + \frac{1}{2}\Xi\} \qquad (6.16)\\
\text{where} \qquad \Xi &= Var_t[r^m_{t+1} - \frac{1}{\sigma}H_1(m_{t+1}-m_t)]\\
&= (1-\frac{1}{\sigma}H_1)^2 Var[r^m_{t+1}] + (\frac{1}{\sigma}H_1)^2\lambda Var[v_{t+1}] - 2\frac{1}{\sigma}H_1\lambda(1-\frac{1}{\sigma}H_1)Cov[r^m_{t+1}, v_{t+1}]
\end{aligned}$$

Equalizing the right-hand side of equation (6.15) and (6.16) and identifying coefficients, we obtain two key coefficients in the consumption function:

$$H_1 = \frac{\eta_w - 1}{\eta_c}, H_0 = \frac{1}{\eta_c}\{\eta - g - \frac{\sigma}{H_1}\log\beta + (1-\frac{\sigma}{H_1})E[r^m_{t+1}] - \frac{1}{2}\Xi\}$$

6.2.3. Deriving the Expression of Change in Consumption in the Presence of Labor Income Risk

Adding RI in the above model yields the following modified consumption rule

$$c_t = H_0 + H_1\widehat{m}_t \qquad (6.17)$$

and substituting it into (6.13) yields

$$m_{t+1} = \eta - g - \eta_c H_0 + (1-\chi)r^f + \eta_w m_t - (\eta_w - 1)\widehat{m}_t + \chi r^e_{t+1} + \lambda v_{t+1}. \qquad (6.18)$$

Furthermore, the information state $\widehat{m}_t$ is characterized by the following Kalman equation

$$\widehat{m}_{t+1} = (1-\theta)\widehat{m}_t + \theta(m_{t+1} + \xi_{t+1}) \qquad (6.19)$$

Combining these three equations yields

$$\Delta\widehat{m}_{t+1} = \theta\eta_w(m_t - \widehat{m}_t) + \theta[\chi r^e_{t+1} + \lambda v_{t+1} + +\xi_{t+1}] + \Omega_1$$

where $\Omega_1 = \theta[\eta - g - \eta_c H_0 + (1-\chi)r^f]$, and

$$m_t - \widehat{m}_t = \frac{(1-\theta)(\chi r^e_t + \lambda v_t) - \theta\xi_t}{1-((1-\theta)\eta_w)\cdot L} + \frac{\Omega_2}{1-(1-\theta)\eta_w}$$

since $m_{t+1} - \widehat{m}_{t+1} = ((1-\theta)\eta_w)(m_t - \widehat{m}_t) + (1-\theta)(\chi r^e_{t+1} + \lambda v_{t+1}) - \theta\xi_{t+1} + \Omega_2$, where L is the lag operator and $\Omega_2 = (1-\theta)[\eta - g - \eta_c H_0 + (1-\chi)r^f]$. Hence, consumption growth can now be written as

$$\begin{aligned}
\Delta c_{t+1} &= H_1\{[\theta(\chi r^e_{t+1} + \lambda v_{t+1}) + \theta\eta_w\frac{(1-\theta)(\chi r^e_t + \lambda v_t)}{1-((1-\theta)\eta_w)\cdot L}]\\
&\quad + [\theta\xi_{t+1} - \frac{\theta\xi_t}{1-((1-\theta)/\alpha_1)\cdot L}] + \Omega\}
\end{aligned}$$

where Ω is the constant term.

6.2.4. Proof of Proposition 5

Proof. (1) it is straightforward from the expression (3.31).

(2)

$$\begin{aligned} Cov[\lim_{S\to\infty}(c_{t+1+S}-c_t), r^e_{t+1}] &= Cov[\lim_{S\to\infty}(\sum_{s=0}^{S}\Delta c_{t+1+s}), r^e_{t+1}] \\ &= H_1\theta\lim_{S\to\infty}[1+(1-\theta)\eta_w+\cdots+((1-\theta)\eta_w)^S]\chi\omega_u^2+H_1\theta\lambda\omega_{uv} \\ &= H_1\frac{\theta}{1-(1-\theta)\eta_w}\chi\omega_u^2+H_1\theta\lambda\omega_{uv}. \end{aligned}$$

(3) Since

$$\begin{aligned} \pi &\approx \gamma Cov_t[\lim_{S\to\infty}(c_{t+1+S}-c_t), r^e_{t+1}] \\ &= \gamma[H_1\frac{\theta}{1-(1-\theta)\eta_w}\chi\omega_u^2+H_1\theta\lambda\omega_{uv}], \end{aligned}$$

we can easily obtain

$$\chi = \vartheta_2[\frac{\pi}{\gamma H_1\omega_u^2}-\frac{\theta\lambda\omega_{uv}}{\gamma\omega_u^2}] \tag{6.20}$$

References

[1] Adam, Klaus (2004), “Optimal Monetary Policy in the Presence of Imperfect Common Knowledge,” Manuscript, European Central Bank.

[2] Aiyagari, Rao (1994), “Uninsured Idiosyncratic Risk and Aggregate Saving,” *Quarterly Journal of Economics* 109, 659-684.

[3] Akerlof, George and Janet Yellen (1985), “Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?” *American Economic Review* 1985, 75, 708-720.

[4] Alessie, Rob and Annamaria Lusardi (1997), “Consumption, Saving and Habit Formation,” *Economics Letters* 55, 103-108.

[5] Amato, Jeffery and Thomas Laubach (2004), “Implications of Habit Formation for Optimal Monetary Policy,” *Journal of Monetary Economics* (forthcoming)

[6] Ameriks, John, Andrew Caplin, and John Leahy (2003), “Wealth Accumulation and the Propensity to Plan,” Forthcoming *Quarterly Journal of Economics*

[7] Ash, Robert (1990), “*Information Theory*,” Dover Publications, Inc.

[8] Ball, Lawrence, Gregory Mankiw, and Ricardo Reis (2003), “Monetary Policy for Inattentive Economies,” Forthcoming *Journal of Monetary Economics*.

[9] Boldrin, Michele, Christiano, Lawrence, and Fisher, Jonas (1997), “Habit Persistence and Asset Returns in an Exchange Economy”, *Macroeconomic Dynamics* 1, no. 2.

[10] Boldrin, Michele, Christiano, Lawrence, and Fisher, Jonas (2001), "Habit Persistence, Asset Returns, and the Business cycle," *American Economic Review.* 91, 149-166

[11] Caballero, Richard (1990), "Consumption Puzzles and Precautionary Savings," *Journal of Monetary Economics 25, 136-136.*

[12] Campbell, John (1987), "Does Saving Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis," *Econometrica* 55, 1249- 1273.

[13] Campbell, John and Angus Deaton (1989), "Why is Consumption So Smooth?," *Review of Economic Studies* 56, 357-374.

[14] Campbell, John (1993), "Intertemporal Asset Pricing without Consumption Data," *American Economic Review*, vol. 83, 487-512.

[15] Campbell, John (1999), "Asset Prices, Consumption, and the Business Cycle," in *Handbook of Macroeconomics*, edited by Taylor and Woodford.

[16] Campbell, John and Luis M. Viceira (2002), "Strategic Asset Allocation: Portfolio Choice for Long-Term Investors," Oxford University Press, UK

[17] Carroll, Christopher (2001), "Precautionary Saving and the Marginal Propensity to Consume out of Permanent Income," Manuscript, John Hopkins University.

[18] Cochrane, John (1989), "The Sensitivity of Tests of the Intertemporal Allocation of Consumption to Near-Rational Alternatives," *American Economic Review*, vol. 79, 319-337.

[19] Constantinides, George (1990). 'Habit Formation: A Resolution of the Equity Premium Puzzle'. *Journal of Political Economy*, 98, 519-43.

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

[21] Deaton, Angus (1992), Understanding Consumption, Oxford University Press, Oxford.

[22] Dynan, Karen and Maki, Dean (2000). "Does Stock Market Wealth Matter for Consumption?" Federal Reserve Board Mimeo.

[23] Epstein, Larry, and Stanley Zin (1989). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57:4, 937-969.

[24] Flavin, Marjorie (1981), "The adjustments of consumption to changing expectations about future income," *Journal of Political Economy* 89, 974-1009.

[25] Fuhrer, Jeffrey (2000), "Habit Formation in Consumption and Its Implications for Monetary Policy," *American Economic Review* 90, 367-390.

[26] Gabaix, Xavier and David Laibson (2001), "The 6D bias and the Equity Premium Puzzle," NBER *Macroeconomic Annual*.

[27] Gali, Jordi (1991) "Budget Constraints and Time Series Evidence on Consumption," *American Economic Review*, vol. 81, 1238-1253.

[28] Gumbau-Brisa, Fabià (2003), "Inflation Persistence and Learning in General Equilibrium," Manuscript, Harvard University.

[29] Hall, Robert (1978). "Stochastic Implications of the life cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 91, 249-265.

[30] Hansen, Lars., and Thomas Sargent (1995), "Discounted Linear Exponential Quadratic Gaussian Control," *IEEE Transactions on Auto. Control,* Vol. 40, No.5, 968-971.

[31] Hansen, Lars., Thomas Sargent, and Thomas Tallarini (1999), "Robust Permanent Income and Pricing," *Review of Economic Studies* 66, 873-907.

[32] Hansen, Lars. and Thomas Sargent (2003), Robust Control and Economic Model Uncertainty, Manuscript, Stanford University

[33] Jacobson, David H. (1973). "Optimal Stochastic Linear Systems with Exponential Performance Criteria and Their Relation to Deterministic Differential Games, " *IEEE Transactions on Auto. Control,* Vol. 18, No.2, 124-131.

[34] Jacobson, David H. (1977). Extensions of Linear-Quadratic Control, Optimization and Matrix Theory. Academic Press, New York.

[35] Kreps, David. and E. Porteus (1978), "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica, XLVI,* 185-200.

[36] Krusell, Per and Anthony Smith, Jr. (1996), "Rules of Thumb in Macroeconomic Equilibrium: A Quantitative Analysis," *Journal of Economic Dynamics and Control* 20, 527-558.

[37] Krusell, Per and Anthony Smith, Jr. (1998), "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy* 106, 867-896.

[38] Lucas, Robert E. (1972), "Expectations and the Neutrality of Money," *Journal of Economic Theory* 4, 103-124.

[39] Luo, Yulei (2004), "Rational Inattention and Aggregate Fluctuation," Manuscript, Princeton University.

[40] Lynch, Anthony. (1996), "Decision Frequency and Synchronization across Agents: Implications for Aggregate Consumption and Equity Returns," *Journal of Finance* 51 (4), 1479-1497.

[41] Mankiw, Gregory and R Reis (2002), "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, vol 117 (4).

[42] Marshall, David and Nayan Parekh (1999), "Can Costs of Consumption Adjustment Explain Asset Prcing Puzzles ," *Journal of Finance* 54 (2) 623-654.

[43] Mehra, Rajnish and Prescott, Edward (1985), "The Equity Premium: A Puzzle," *Journal of Monetary Economics* 15, 145-161.

[44] Moffitt, Robert and Peter Gottschalk (1995), "Trends in the Covariance Structure of Earnings in the U.S.: 1969-1987," Manuscript, Johns Hopkins University and Boston College.

[45] Moscarini, Giuspeppe (2003), "Limited Information Capacity as a Source of Inertia," Forthcoming, *Journal of Economic Control and Dynamics.*

[46] Parker, Jonanthan (1999), "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes," *American Economic Review*, Vol 89, No.4, 959-973.

[47] Parker, Jonanthan (2001), "The Consumption Risk of the Stock Market," *Brookings Papers on Economic Activity*, No.2, 279-348.

[48] Parker, Jonanthan (2003), "Consumption Risk and Expected Stock Returns," *NBER* Working Paper 9548.

[49] Parker, Jonanthan and Christian Julliard (2004), "Consumption Risk and the Cross-section of Expected Returns," Forthcoming, *Journal of Political Economy.*

[50] Peng, Lin and Wei Xiong (2001), "Time to Digest and Volatility Dynamics," Manuscript, Baruch College and Princeton University.

[51] Peng, Lin (2004), "Learning with Information Capacity Constraints," Manuscript, Baruch College, City University of New York.

[52] Pischke, Jorn-Steffen (1995), "Individual Income, Incomplete Information, and Aggregate Consumption," *Econometrica* 63 (4), 805-840.

[53] Ploeg, Frederick Van Der (1993), "A Closed-form Solution for a Model of Precautionary Saving,". *Review of Economic Studies*, 60, 385-395.

[54] Reis, Ricardo (2003), "Inattentive Consumer," Manuscript, Harvard University.

[55] Rios-Rull, José Víctor (1999), "Computation of Equilibria in Heterogeneous-agent Models," Manuscript, University of Pennsylvania.

[56] Sargent, Thomas (1993), *Bounded Rationality in Macroeconomics*, Oxford University Press.

[57] Shannon, C.E. (1948), "A Mathematical Theory of Communication," *The Bell System Technical Journal* 27, 623-656.

[58] Sims, Christopher A. (1998) "Stickiness," *Carnegie-Rochester Conference Series on Public Policy*, Vol 49(1), 317-356.

[59] Sims, Christopher A. (2003) "Implications of Rational Inattention," *Journal of Monetary Economics* 50, 3, 665-690.

[60] Souleles, Nicholas S. (1999), "The Responses of Household Consumption to Income Tax Refunds," *American Economic Review*, vol. 89 (4), 947-958.

[61] Storesletten, K., C.I. Telmer, and A. Yaron (2001), "Cyclical Dynamics in Idiosyncratic Labor Market Risk," *Journal of Political Economy* 112(3), 695-717.

[62] Tallarini, Thomas D. (2000), "Risk-sensitive Real Business Cycles," *Journal of Monetary Economics* 45, 507-532.

[63] Thaler, Richard (1990), "Anomalies: Saving, Fungibility, and Mental Accounts," *Journal of Economic Perspectives*, Vol. 4, No. 1. 193-205.

[64] Viceira, Luis M. (2001), "Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income," *The Journal of Finance*, Vol. 56, No. 2: 433-470, April 2001

[65] Weil, Phillippe. (1993). "Precautionary Savings and the Permanent Income Hypothesis," *Review of Economic Studies*, 60, 367-383.

[66] Whittle, Peter (1981). "Risk-sensitive Linear/Quadratic/Gaussian Control," *Advances in Applied Probability*,13, 764-777.

[67] Whittle, Peter (1982) "*Optimization Over Time*," Volume 1, Chichester: John Wiley & Sons.

[68] Whittle, Peter (1990) "*Risk-sensitive Optimal Control*," Wiley, Chichester and New York.

[69] Whittle, Peter (1996), "*Optimal Control: Basics and Beyond*," John Wiley & Sons.

[70] Woodford, Michael (2001) "Imperfect Common Knowledge and the Effects of Monetary Policy," Princeton University, in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds.