Rational Inattention, Portfolio Choice, and the Equity Premium*

Yulei Luo

Job Market Paper

January, 2006

Abstract

This paper explores how the introduction of Rational Inattention (RI) affects optimal consumption and portfolio rules and asset pricing in the consumption-based CAPM framework. I first solve an otherwise standard portfolio choice and asset pricing model with RI explicitly and show that RI can generate smooth consumption process and low contemporaneous correlation between consumption growth and asset returns. Second, it is shown that in the RI economy asset returns are determined by the *ultimate* consumption risk rather than the *contemporaneous* risk. As a result, RI has a potential to reduce the demand for the risky asset and could endogenize "limited stock market participation" hypothesis. Third, I show that RI can disentangle the coefficient of relative risk aversion with the elasticity of intertemporal substitution endogeneously by increasing the *effective* CRRA. RI can therefore be an alternative explanation for the equity premium puzzle and the risk free rate puzzle. Fourth, I compare RI with recursive preference, robustness, and habit formation. Fifth, I investigate the implications of RI for optimal consumption and portfolio choice when investment opportunities are stochastic and labor income is modelled explicitly. Finally, I propose a general equilibrium asset pricing framework to examine the implications of RI for the equity premium, the mean ratio of price to dividend, and the equity volatility in equilibrium.

Keywords: *Rational Inattention, Consumption Risk, Portfolio Choice, the Equity Premium*

JEL Classification Codes: *E44, G11, G12*

*I am indebted to Christopher A. Sims for his excellent advice and support during my Ph.D. at Princeton University. I would also like to thank Per Krusell, Jonathan Parker, Ricardo Reis, Charles Whiteman, Wei Xiong, Eric Young, Chi-wa Yuen, and seminar participants at Princeton, the University of Hong Kong, the 2005 Midwest Macro Meeting, the Society for Economic Dynamics 2005 Meeting, and the 2005 Econometric Society World Congress for many helpful suggestions and comments. Email address: yluo@princeton.edu. Web page: http://www.princeton.edu/~yluo.

1. Introduction

Optimal asset allocation is a classic problem in financial economics and macroeconomics. In a single-period setting or a multiperiod setting where investment opportunities are constant, optimal portfolio weights are functions of the first and second moments of asset returns. Specifically, in a portfolio with one riskless asset and one risky asset, the optimal share invested in the risky asset is proportional to the expected excess return and inversely proportional to the volatility of the risky asset. However, given the observed mean and volatility of the risky asset and plausible risk aversion, this kind of models generates a counterfactually high stock market participation rate[1].

Another important topic that is closely related to optimal portfolio choice is asset pricing. According to canonical consumption-based capital asset pricing theory (CCAPM), the expected excess return on any risky portfolio over the riskless interest rate is determined by the quantity of risk times the price of risk, where the quantity of risk is measured by the covariance of the excess return with contemporaneous consumption growth and the price of risk is the coefficient of relative risk aversion (CRRA) of the representative agent. Given i.i.d. stock returns, the optimal portfolio share in equities of an investor with power utility is proportional to the risk premium and the reciprocal of the CRRA and the volatility of asset returns. Hence, the canonical CCAPM theory predicts that equities are not very risky because of the low covariance between equity returns and contemporaneous consumption growth[2]. Consequently, to generate the observed high equity premium measured by the difference between the average real stock return and the average short-term real interest rate, the CRRA must be very high. Mehra and Prescott (1985) examined this issue in the Lucas-type general equilibrium asset pricing framework and first called it "the equity premium puzzle"[3]. Hansen and Jagannathan (1991) and Cochrane and Hansen (1992) related this puzzle to the volatility of the stochastic discount factor (SDF). They argued that this puzzle is such that an extremely volatile SDF is required to match the Sharpe ratio[4]. Kandel and Stambaugh (1991) responded to this puzzle by arguing that the CRRA is indeed much higher than the values traditionally thought. However, as argued in Weil (1989), although a high value of the CRRA can help resolve this puzzle, it brings another puzzle, the risk free rate puzzle[5].

[1] In a country with a well-developed equity culture like the U.S., the direct ownership of publicly traded stocks was 21.3% in the 2001 Survey of Consumer Finance.

[2] Given the property of equity returns, the low covariance is determined by the smoothness of consumption and the low correlation between equtiy return and consumption.

[3] For the history of this puzzle, see Lucas (1978), Grossman and Shiller (1981), Hansen and Singeton (1983), Mehra and Prescott (1985), Hansen and Jagannathan (1991), and Cochrane and Hansen (1992). Kocherlakota (1996), Campbell (1999, 2003), and Cochrane (2005) gave useful reviews on this topic.

[4] It is defined as the ratio of the equity premium to the standard deviation of stock returns.

[5] Weil (1989) argued that if one accepts high risk aversion, the corresponding equation for the risk less rate implies that the risk free interest rate is extremely high. To generate a 5% interest rate, a negative 15% discount rate is

Numerous economic arguments have been proposed to explain this puzzle. A partial list of these explanations in the CCAPM framework includes: habit formation in consumption (Contantinides, 1982; Campbell and Cochrane, 1999), recursive utility (Epstein and Zin,1989; 1991), limited stock market participation (Vissing-Jorgensen, 2002), delayed adjustment (Grossman and Laroque, 1990; Lynch, 1996; Marshall and Parekh, 1999; and Gabaix and Laibon, 2001), and the preference for robustness (Maenhout, 2004).

An implicit but key assumption in the bulk of this work is that individuals have unlimited information-processing capacity and thus can observe the relevant state(s) without errors and react instantaneously and completely to any innovations to equity returns. However, as argued in Sims (1998, 2003, 2005), this assumption is not consistent with the inborn ability of human beings, that is, ordinary people only have limited information-processing capacity. Consequently, people cannot observe the state(s) perfectly and thus have to react to the innovations to the state gradually and with delay. In Sims (2003), this kind of information-processing constraints is first called "Rational Inattention" (henceforth, RI). Recently, RI has been incorporated into a variety of economic models to address some interesting issues, e.g., the observed inertial behavior, the persistence problem, and so on. I will give a brief review later.

For this reason, it is desirable to take RI into account when studying optimal consumption and portfolio rules and asset returns. In this paper, I examine the implications of RI on optimal portfolio choice and asset returns in both the partial equilibrium and general equilibrium CCAPM frameworks. RI was first introduced into economics by Sims (2003). In his RI framework, a concept in information theory, *entropy*, is used to measure the uncertainty of a random variable, and the reduction in the entropy is used to measure information flow. For finite Shannon capacity (a kind of cognitive limitation), the reduction in entropy is bounded above by the limited capacity. As a result, individuals cannot eliminate all uncertainty about the state(s) when new information arrives. He then developed a tractable framework for solving individuals' LQG (Linear-Quadratic-Gaussian) optimization problems when they have finite Shannon capacity and showed that the optimal reactions of individuals with respect to fundamental shocks are delayed; hence RI can be an alternative candidate to explain the observed inertial behavior in the US economy. In contrast to the stochastic optimal control problems with exogenous noises in the stochastic control literature[6], the RI model implies that 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.

required.

[6]See Whittle (1982, 1996) for details.

In this paper, I introduce RI into the CCAPM frameworks and explore its implications for optimal consumption and portfolio choice and asset returns. As a first contribution, I solve a partial equilibrium CCAPM model[7] with RI explicitly and all results are then obtained in closed form. The explicit expressions of optimal consumption and portfolio rule in terms of primitive parameters and the parameter governing the degree of RI (κ) give us clear insights about the impacts of RI on portfolio choice and asset returns.

Second, after solving the model, I show that RI can substantially affect the intertemporal allocation of consumption and thus generate a smooth consumption process and low contemporaneous covariance between consumption growth and asset returns[8]. In addition, the model implies that lagged equity returns could be used to predict future consumption growth, which is consistent with the data to some extent[9]. The intuition is simple. Because the individuals devote their finite information capacity to observing the state of their financial wealth, and thus the state can not be perfectly observed, optimal consumption and portfolio decisions are made relative to a noisy signal of the true state, and information about changes in the true state are not entirely incorporated into forecasts. In other words, they have to take some time to digest new information about the state. Hence, past information about equity returns is helpful to predict future consumption growth.

Third, I show that just as argued in Parker (2001; 2003) and Parker and Jullard (2005), because consumption takes many periods to adjust with respect to the innovations to risky assets in the RI economy, the quantity of the risk of the risky portfolio should be determined by its *ultimate* consumption risk instead of the *contemporaneous* one. Specifically, the usual Euler equation does not hold because consumption cannot react instantaneously and completely to the innovations due to limited capacity in processing information. Instead, a long-term Euler equation holds in the RI world because consumption reacts slowly with respect to the changes in wealth and takes many periods to adjust. Hence, the equity premium should be measured by the ultimate consumption risk. Because the ultimate consumption risk is larger than the contemporaneous one, the individuals will face larger risk if measuring the risk correctly and will then require higher equity premium for compensating. In sum, incorporating RI can reconcile three apparent anomalies in the full-information CCAPM model simultaneously: a) the excess smoothness of aggregate consumption,

[7] Actually, it is a log-linearized version of the standard Merton and Samuelson model.

[8] This RI model predicts similar aggregate dynamics generated from the delayed adjustment model (e.g., the $6D$ bias model proposed by Gabaix and Laibson (2001)) because both of them predict that aggregate consumption should react to lagged changes in wealth. However, 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.

[9] When we solve the benchmark CCAPM model with RI, we assume that the equity return is i.i.d.; hence we cannot address the question: can consumption growth be used to predict equity returns? But when we allow time-varying investment opportunities, it will be clear that current consumption growth can predict future equity returns.

b) the low contemporaneous covariance between consumption growth and asset returns, and c) the high equity premium.

Fourth, it is straightforward to show that incorporating RI into this expected utility maximization framework can also disentangle the CRRA from the elasticity of intertemporal substitution (EIS) by increasing the *effective* CRRA and reducing the EIS[10]. Hence, if individuals are highly inattentive, the model can generate high equity premium by just increasing the *effective* CRRA that is an increasing function of the degree of inattention and leaving the *true* CRRA unchanged. RI can therefore be an alternative explanation for both the equity premium puzzle (by increasing the effective CRRA) and the risk free rate puzzle (by remaining the true CRRA unchanged). In other words, RI can reconcile low estimates of risk aversion obtained from experimental evidence or introspection, with high estimated values of risk aversion based on asset pricing data. Further, given i.i.d. equity returns, it is easily shown that RI reduces the demand for the risky asset. The intuition is that if agents can not allocate enough channel capacity in monitoring their financial wealth evolution, it is not rational for them to invest a large fraction of their wealth in the risky portfolio because the innovations to their financial wealth can generate large consumption risk in the long run if the capacity is low. As a result, RI provides a possibility for "the limited stock market participation" observed in the data.

Fifth, I explore the different implications of RI, recursive preference, and the preference for robustness on asset holdings and asset returns. In the Epstein-Zin (1989) and Weil (1989)'s recursive utility framework, CRRA and EIS are exogenously separated in the preference. Hence, increasing relative risk aversion can help resolve the equity premium puzzle and will not cause the risk free rate puzzle because the CRRA and the IES are unrelated in the recursive utility framework. In the robustness asset pricing literature, Maenhout (2004) used similar methodology as developed by Hansen and Sargent (2002) and Anderson et al. (2002), and showed that the preference for robustness dramatically reduces the optimal share of the portfolio allocated in equities because investors are very conservative or pessimistic when forming portfolios due to robustness (model uncertainty)[11].

Sixth, in the expected utility framework the CRRA and the EIS are closely related. This link between the two distinct parameters is broken in the recursive utility framework. Hence, incorporating RI into the recursive utility framework allows us to explore the interactions of RI, the CRRA, and the IES when determining portfolio choice and asset returns. It is shown that

[10] RI reduces the EIS in consumption because individuals are *constrained* to substitute consumption over time due to finite capacity. This is an objective source of reluctance instead of a subjective one from changing the preference directly.

[11] In fact, we can also regard RI as a source of pessimism or conservatism.

incorporating RI in this framework then goes further in explaining the equity premium puzzle and the risk free rate puzzle.

Seventh, in the data it seems that expected asset returns vary through time so that investment opportunities are not constant, and the evidence for predictable variation in the equity premium is particularly strong. Several papers have been developed to solve for optimal portfolios in models with realistic predictability of returns. See Campbell and Viceira (1999, 2002) for detailed discussions. Hence, it seems natural to analyze the optimal asset demands including both myopic demand and hedging demand of the consumers in the RI economy. It is shown that RI reduces the speed of adjusting the portfolio with respect to the innovations to equity returns.

Finally, the high equity premium is, literally speaking, no puzzle in the above partial equilibrium CCAPM framework because asset prices are specified exogenously. The endogenous variables in this model are consumption processes of individual consumers. Aggregating over all consumers yields an aggregate consumption process. It is therefore interesting to examine the implications of RI on the equity premium puzzle in the general equilibrium CCAPM framework. Specifically, I show that in a Lucas-type equilibrium asset pricing model, RI reduces the mean ratio of price to dividend, increases the equity premium, and provides an explanation for the equity volatility puzzle.

Recently, there have been some papers that incorporate explicit information processing constraints into a variety of theoretical models and explore how they affect the optimal decision rules of consumers or firms, as well as their 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. Peng (2004) explored the effects of information constraints on the equilibrium asset price dynamics and consumption behavior under the continuous-time CARA framework. Moscarini (2003) derived optimal time-dependent adjustment rules from the information constraints in a continuous-time framework. Kasa (2005) compared RI with the preference for robustness and showed that they are observationally equivalent in the sense that a higher filter gain can be interpreted as either a strong preference for robustness or a strong ability to process information. Luo (2005) examined the implications of RI for consumption dynamics in the permanent income hypothesis model and showed that RI can be an alternative explanation for the excess smoothness puzzle and the excess sensitivity puzzle. Luo and Young (2005) examined the effects of RI on the amplification and propagation of aggregate shocks in a stochastic growth model. Maćkowiak and Wiederholt (2005) explored the implications of RI on optimal sticky prices. Nieuwerburgh and Veldkamp (2005a;

2005b) study information acquisition, portfolio under-diversification, and the home bias puzzle. 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 showed 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[12].

This paper is organized as follows. Section 2 presents and solves a partial equilibrium CCAPM model with RI, and examines its implications for portfolio choice and the equity premium. Section 3 discusses the implications of RI in the recursive utility framework. Section 4 considers stochastic investment opportunities to analyze the implication of RI on asset demands in some special cases. Section 5 analyzes the implications of labor income risk in the CCAPM model with RI. Section 6 studies the implications of RI on asset prices in the Lucas-type general equilibrium framework. Section 7 presents some empirical evidence. Section 8 concludes. Appendices contain the proofs and derivations that are omitted from the main text.

2. Consumption-based CAPM Models with Rational Inattention

In this section, I first present a simple standard CCAPM model, and then discuss how to incorporate RI into this framework and solve the model in close form approximately. The model proposed in this section is based on three literatures. First, I follow Campbell (1993; 1999) and Campbell and Viceira (1999) and solve the CCAPM by log-linearizing the budget constraint and the Euler equation and by using the method of undetermined coefficients to find policy functions[13]. Second, since the original CRRA specification can be approximated by a log-LQ framework under some conditions, it can be fitted into the Gaussian-error RI framework developed in Sims (2003) approximately as in the pure LQ problem. Third, in the RI economy, we need to use the ultimate consumption risk to price risky assets. The logic is based on the work by Parker (2001; 2003) and Parker and Julliard (2005).

[12] 'RI' modeled in Sims (2003) and others is based on Shannon channel capacity, whereas 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.

[13] Actually, it is in the vein of Merton (69) and Samuelson (69). Here I adopt Campbell's discrete-time log-linearized version because it has an approximate linear-quadratic-Gaussian framework and thus RI can be easily introduced.

2.1. Specification and Solution of the Standard CCAPM Model

Before setting up and solving the CCAPM model with RI, it is helpful to present the standard CCAPM model first and then discuss how to introduce RI in this framework. Here I consider a simple partial equilibrium CCAPM model[14]; the identical consumers maximize the following intertemporal welfare by choosing consumption,

$$V = \mathbf{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\gamma} - 1}{1-\gamma} \tag{2.1}$$

where C_t represents individual's consumption at time t, β is the discount factor, γ is the coefficient of relative risk aversion (CRRA), and $\sigma = 1/\gamma$ is the elasticity of intertemporal substitution (EIS). When $\gamma = 1$, the utility function becomes logarithmic form, $\log C_t$.

To keep the analysis simple, I 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 $r^f (= \log R^f)$. I refer to asset e as a portfolio of equities, and to asset f as savings or checking accounts. Furthermore, I assume that r_{t+1}^e has expected return μ, where $\mu - r^f$ is the equity premium, and an unexpected component u_{t+1} with $\mathbf{var}(u_{t+1}) = \omega_u^2$.

The flow budget constraint for consumers can be written as

$$W_{t+1} = R_{t+1}^p (W_t - C_t) \tag{2.2}$$

where W_{t+1} is an individual's 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[15], and R_{t+1}^p is the one-period return on savings given by

$$R_{t+1}^p = \chi_t (R_{t+1}^e - R^f) + R^f \tag{2.3}$$

where $\chi_t = \chi$ is the proportion of savings invested in the risky asset[16]. As in Campbell (1993), we can easily derive an approximate expression for the log return on wealth as follows,

$$r_{t+1}^p = \chi (r_{t+1}^e - r^f) + r^f + \frac{1}{2}\chi(1-\chi)\omega_u^2. \tag{2.4}$$

Because the term $\frac{1}{2}\chi(1-\chi)\omega_u^2$ is one order of magnitude smaller than the mean values of the

[14] The model is based on Campbell (1993) and widely adopted in the macroeconomics and finance literature.

[15] 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.

[16] Given i.i.d. equity returns and power utility function, the share invested in equities, χ_t, is constant over time.

returns on equity and the riskless asset given the realistic value of ω_u^2 observed in the data, for simplicity, we may assume that $r_{t+1}^p \simeq \chi(r_{t+1}^e - r^f) + r^f$.

As 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 I follow Campbell (1993), Campbell and Viceira (1999), and Viceira (2001) and use the log-linearization method to solve the model[17].

First, I divide equation (2.2) by W_t and then log-linearize it around steady state $c - w = \mathbf{E}(c_t - w_t)$:

$$\Delta w_{t+1} = r_{t+1}^p + \psi + (1 - 1/\phi)(c_t - w_t) \tag{2.5}$$

where $\phi = 1 - \exp(c - w)$, $\psi = \log\phi - (1 - 1/\phi)\log(1 - \phi)$, and lowercase letters denote logs. Next, I log-linearize the Euler equation

$$\mathbf{E}_t[\beta R_{t+1}^k (C_{t+1}/C_t)^{-1/\sigma}] = 1$$

around[18] $\mathbf{E}_t[\log\beta - \frac{1}{\sigma}(c_{t+1} - c_t) + r_{t+1}^k]$ where $k = e,\ f$, and p, and obtain the following familiar form[19]

$$0 \simeq \log\beta - \frac{1}{\sigma}\mathbf{E}_t[c_{t+1} - c_t] + \mathbf{E}_t[r_{t+1}^k] + \frac{1}{2}\mathbf{var}_t[r_{t+1}^k - \frac{1}{\sigma}(c_{t+1} - c_t)] \tag{2.6}$$

Furthermore, guess that the optimal log consumption rule takes the following form

$$c_t = b_0 + b_1 w_t, \tag{2.7}$$

and thus we can easily get the expression for the change in consumption, $\Delta c_{t+1} = b_1 \Delta w_{t+1}$.

Combining equations (2.5), (2.6), and (2.7) gives the undetermined coefficients in the consump-

[17] 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 the 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.

[18] Note that the Euler equation

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

can be written as

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

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

tion rule[20]:

$$b_1 = 1 \text{ and } b_0 = \log\{1 - \exp[(\frac{1}{\gamma} - 1)\mathbf{E}_t[r^p_{t+1}] + \frac{1}{\gamma}\log\beta + \frac{1}{2\gamma}(1-\gamma)^2\mathbf{var}_t[r^p_{t+1}]\}, \tag{2.8}$$

and the optimal portfolio rule

$$\chi = \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\gamma b_1 \omega_u^2} \tag{2.9}$$

Hence, the optimal portfolio has mean $\mathbf{E}_t[r^p_{t+1}] = r^f + \frac{(\mu - r^f)(\mu - r^f + \frac{1}{2}\omega_u^2)}{\gamma\omega_u^2}$ and variance $\mathbf{var}_t[r^p_{t+1}] = \frac{(\mu - r^f + \frac{1}{2}\omega_u^2)^2}{\gamma^2\omega_u^2}$. It is also worth noting that for the isoelastic utility, the portfolio rule is independent of the consumption rule, and when $\gamma = 1$, the consumption rule is also independent of the stochastic properties of financial assets.

Hence, (2.7) and (2.9) characterize the solution to the standard CCAPM completely. Further, given the CRRA utility and i.i.d. equity returns, the value function of the consumer in terms of log wealth can be then written as follows

$$V(w_t) = \frac{\exp(-\gamma b_0 + (1-\gamma)w_t)}{1-\gamma} \tag{2.10}$$

2.2. Implications for the Implied Equity Premium

Strictly speaking, there is no equity premium puzzle or riskfree rate puzzle in this partial equilibrium framework because asset returns are given exogenously and we just solve for optimal consumption and portfolio rules given the returns. I will discuss the implications of RI for the equity premium puzzle in a general equilibrium framework in section 6. Here I only briefly discuss how the equity premium puzzle appears in this partial equilibrium model.

First, we need to use the implied relationship between consumption growth and asset returns. Based on the results of the above CCAPM case, we know that the joint distribution of the returns on the equity and consumption growth is log normal. Hence, the expected excess return can be written as[21]

$$\begin{aligned} \mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2 &= \gamma\mathbf{cov}_t[c_{t+1} - c_t, r^e_{t+1}] \\ &= \gamma\chi\omega_u^2 \end{aligned} \tag{2.11}$$

where we use the formula that $\Delta c_{t+1} = \Delta w_{t+1} = r^p_{t+1} + \psi + (1 - 1/\phi)b_0$.

[20] For the detailed derivations, see Appendix B in Viceira (2001) and Campbell and Viceira (2002).

[21] Given i.i.d. asset returns, this expression for the conditional expection of excess returns is equivalent to the corresponding unconditional expression.

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 3 and χ is 0.25[22], the equity premium is close to $0.75\omega_u^2$ which is around 2.4%[23]. 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 (2.11), this result is generated from the unrealistic contemporaneous covariance between consumption growth and asset return which is $5.6 \cdot 10^{-3}$ annually[24] 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 equity returns is around $32 \cdot 10^{-}5$. Therefore, one puzzle for this simple consumption CAPM model is that the model predicts too much contemporaneous covariance between consumption growth and asset returns due to large consumption volatility and high contemporaneous correlation between consumption growth and equity returns. As shown in Campbell (1999, 2003), once we substitute the empirical values of the exess returns and the covariance between equity returns and consumption growth into (2.11), the implied relative risk aversion, γ, is extremely high for both US data and international data. See Campbell (2003) for a detailed discussion about the equity premium puzzle based on this equation. Luo (2005) showed that introducing RI into a standard LQ PIH model can reduce the variability of consumption growth when labor income is nonstationary, and then have a potential to explain the excess smoothness puzzle (the Deaton's puzzle) and the excess sensitivity puzzle in the consumption literature. Hence, it seems promising to incorporate RI into this canonical CCAPM framework to examine if it would be an alternative explanation for the equity premium puzzle and the risk free rate puzzle.

2.3. Incorporating RI into the Standard CCAPM

In the standard LQ problem with imperfect observations on the state, we just need to apply the separation principle and replace the true state with its estimated one. To apply this principle here, we need to show that the volatility of stock returns has no impact on the decision rule, that is, there is no precautionary saving here. From the decision rule (2.7) with the coefficients, b_0 and b_1, defined in (2.8), we can see that the precautionary saving term is zero only if the utility is logarithmic,

[22] Here setting the value of χ to be 0.25 is based on a calibration exercise in Gabaix and Laibson (2001) where they assumed that all capital is identical to stock market capital and calibrate the equity share of total wealth is around 0.22.

[23] Campbell (1999) reported a series of main moments from his dataset including 11 main industrialized 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 $1947-1996$.

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

that is, $\gamma = \sigma = 1$. However, in this case the utility function as well as the value function is linear if we reformulate the problem with the log variables, c_t and w_t. As a result, the loss function due to RI is also a linear function; therefore, minimizing this loss function subject to the information constraints does not imply that the true state follows a normal distribution given information at t. See Sims (2005) for a discussion about the RI model with linear utility function. However, when γ is very close to 1 from above [25] such that $(\gamma - 1)c_t$ is close to 0, the CRRA utility function can be approximated by a log-LQ function. Specifically, since $(\gamma - 1)c_t$ is close to 0, we can approximate the utility function around $(\gamma - 1)c_t = 0$ as follows:

$$\frac{C_t^{1-\gamma} - 1}{1-\gamma} = \frac{(\exp(c_t))^{1-\gamma} - 1}{1-\gamma} \simeq c_t + \frac{1}{2}(1-\gamma)c_t^2 + \frac{1}{3}(1-\gamma)^2 c_t^3 \tag{2.12}$$

where the third term, $\frac{1}{3}(1-\gamma)^2 c_t^3$, measures prudence. Note that since the ratio of the third term to the second term in the above approximation, $\frac{2}{3}(\gamma - 1)c_t$, is close to 0 when $(\gamma - 1)c_t$ is close to 0, the original CRRA function can be replaced by a log-LQ function, $c_t + \frac{1}{2}(1-\gamma)c_t^2$.

I next discuss how to incorporate RI in this log-linearized CCAPM model. First, I redefine c_t as the new control variable and w_t as the new state variable. Then, following Sims (2003), Luo (2005), and Luo and Young (2005), I present the CCAPM model with RI as follows[26]. I assume that the individuals maximize their lifetime utility subject to both the usual flow budget constraint and the information-processing constraint that will be specified later. The dynamic optimization problem is

$$\widehat{V}(\widehat{w}_t) \simeq \max_{c_t, \mathcal{D}_t} \mathbf{E}_0 \sum_{t=0}^{\infty} \beta^t [c_t + \frac{1}{2}(1-\gamma)c_t^2] \tag{2.13}$$

subject to

$$\Delta w_{t+1} = r_{t+1}^p + \psi + (1 - 1/\phi)(c_t - w_t)$$

$$w_{t+1} | \mathcal{I}_{t+1} \sim \mathcal{D}_{t+1} \tag{2.14}$$

$$w_t | \mathcal{I}_t \sim \mathcal{D}_t \tag{2.15}$$

$$\text{given } w_0 | \mathcal{I}_0 \sim N(\widehat{w}_0, \Sigma_0) \tag{2.16}$$

[25] Calibrated macroeconomic models designed to match growth and business cycle facts typically require that the CRRA be close to 1 (slightly greater than 1) and the IES be also close to 1 (slightly less than 1). For example, see Weil (1989).

[26] As suggested by Chris Sims, a more ideal way to solve this model with RI is to solve for the nonlinear Euler equation for this nonlinear model with RI first and then approximate the nonlinear Euler equation and the nonlinear flow budget constraint. However, this approach is still not feasible now. I leave it for future research.

and the requirement that the rate of information flow at $t+1$ implicit in the specification of the distributions, $\mathcal{D}_t$ and $\mathcal{D}_{t+1}$ be less than channel capacity. $\widehat{w}_t$ is the estimated state variable, and $\mathcal{I}_t$ is the information available at time t. The expectation is formed under the assumption that $\{c_t\}_0^\infty$ are chosen under the information processing constraints.

In the presence of RI, the individuals in the economy cannot observe the state(s) perfectly because observing the state involves information transfer at a limited channel capacity κ, so that they have to choose the nature of the endogenous error optimally. For simplicity, here I assume that all individuals in the model economy have the same channel capacity; hence the average capacity in the economy is equal to individual capacity[27]. In this case the effective state variable is not the traditional state variable (e.g., the wealth level w_t in this model), but the so-called *information state*: the distribution of the state variable w_t conditional on the information set available at time t, $\mathcal{I}_t$. In other words, it expands the state space to the space of distributions on w_t. We therefore have a "curse of dimensionality" problem.[28] Fortunately, the above problem can be approximated by a Linear-Quadratic-Gaussian framework in which the conditional distributions are Gaussian; the first two moments, the conditional mean $\widehat{w}_t$ and the conditional covariance matrix Σ_t, are therefore sufficient to characterize the effective state. Hence, $\widehat{V}(\widehat{w}_t)$ is the value function under RI, and $V(w_t)$ is the value function from the standard model, where the consumers are assumed to have unlimited channel capacity and thus can observe the state perfectly. Finally, I define the loss function at t due to imperfect information as the difference between these two value functions, that is, $\Delta V = V(w_t) - \widehat{V}(\widehat{w}_t)$.

Here I 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.[29] With finite capacity, the agents will choose a signal that reduces the uncertainty of the state. Formally, this idea can be described by the following information constraint

$$\mathcal{H}(w_{t+1}|\mathcal{I}_t) - \mathcal{H}(w_{t+1}|\mathcal{I}_{t+1}) \leq \kappa \tag{2.17}$$

where κ is the consumer's information channel capacity, $\mathcal{H}(w_{t+1}|\mathcal{I}_t)$ denotes the entropy of the state prior to observing the new signal at $t+1$, and $\mathcal{H}(w_{t+1}|\mathcal{I}_{t+1})$ is the entropy after observing the new signal. κ imposes an upper bound on the amount of information – that is, the change in

[27] Assuming that channel capacity follows some distribution complicates the problem when aggregating, but does not change the main findings.

[28] Another type of models which also have the problem of the curse of dimensionality are the heterogeneous-agent models with both idiosyncratic shock and aggregate shock. In those models, the measure is a state variable. Sims (2005) contains a discussion of some of the complications that such a model would present.

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

the entropy – that can be transmitted in any given period[30].

So far I have not shown that $\mathcal{D}_t$ is a normal distribution. As in Sims (2003; 2005) and Luo (2005), I propose the following procedure to deduce its property. First, I guess that the loss function the agent used to deduce the distribution of the actual state w_t is quadratic in terms of $w_t - \widehat{w}_t$. Minimizing this guessed loss function subject to information constraints implies that $w_t|\mathcal{I}_t \sim N(\widehat{w}_t, \Sigma_t)$[31]. Second, using this key property, I can derive the optimal consumption function and the value function in the RI case. Third, given the derived value functions in both the full-information case and the RI case, I show that the loss function is indeed a quadratic function in terms of $w_t - \widehat{w}_t$[32]. Hence, I verify the guess that $\mathcal{D}_t$ is indeed a normal distribution $N(\widehat{w}_t, \Sigma_t)$[33].

Therefore, (2.17) can be rewritten as

$$\frac{1}{2}[\log \Psi_t - \log \Sigma_{t+1}] \leq \kappa \tag{2.18}$$

where $\Sigma_{t+1} = \mathbf{var}_{t+1}(w_{t+1})$ and $\Psi_t = \mathbf{var}_t(w_{t+1})$ are the posterior and the prior variance-covariance matrices of the state vector. This means that given a finite capacity κ per time unit, the optimizing consumer would choose a signal that reduces the conditional variance by[34]

$$\kappa = \frac{1}{2}[\log \Psi_t - \log \Sigma_{t+1}] \tag{2.19}$$

Note that here I use the fact that the entropy of a Gaussian random variable is equal to half of its logarithm variance plus some constant term. In the univariate state case this information constraint completes the characterization of the optimization problem and everything can be solved analytically, whereas for the multivariate state case, we need another information constraint, that is, $\Psi_t \succeq \Sigma_{t+1}$. This constraint embodies the restriction that precision in the estimates of the state cannot be improved by forgetting some components (making the change in their entropy negative)

[30] 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.

[31] See Appendix B for a detailed derivation.

[32] Hence, minimizing this loss function under the IPC also implies that $S_t|\mathcal{I}_t \sim N(\widehat{S}_t, \Sigma_t)$. See Appendix B for a detailed derivation.

[33] We can treat the above problem as a two-step optimization procedure. First, given the observed signal about the state, we could derive the optimal decision rule in the case with imperfect information. Second, given the optimal decision rule, the individual minimizes the expected loss function by choosing the information system, that is, the property of the signal.

[34] Note that given Σ_t, choosing Σ_{t+1} is equivalent with choosing the noise $\mathbf{var}(\xi_t)$ since the usual updating formula for the variance of a Gaussian distribution is

$$\Sigma_{t+1} = \Psi_t - \Psi_t(\Psi_t + \mathbf{var}(\xi_t))^{-1}\Psi_t$$

where Ψ_t is a function of Σ_t.

and using that extra capacity to reduce other components by more than κ.

Then flow budget constraint (2.5) implies that

$$\mathbf{E}_t[w_{t+1}] = \mathbf{E}_t[r^p_{t+1}] + \psi + \widehat{w}_t,$$

$$\mathbf{var}_t[w_{t+1}] = \mathbf{var}_t[r^p_{t+1}] + (1/\phi)^2\Sigma_t,$$

(2.19) therefore implies that

$$\kappa = \frac{1}{2}[\log(\mathbf{var}_t[r^p_{t+1}] + (1/\phi)^2\Sigma_t) - \log\Sigma_{t+1}], \tag{2.20}$$

which has a steady state

$$\Sigma = \frac{\mathbf{var}_t[r^p_{t+1}]}{\exp(2\kappa) - (1/\phi)^2} \tag{2.21}$$

where $\mathbf{var}_t[r^p_{t+1}] = \chi^2\omega_u^2$.

Hence, we can apply the separation principle[35] in this log-LQ CCAPM model and then have the following modified consumption rule

$$c_t = b_0 + \widehat{w}_t \tag{2.22}$$

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

$$\widehat{w}_{t+1} = (1-\theta)\widehat{w}_t + \theta w^*_{t+1} \tag{2.23}$$

where $\theta = 1 - \frac{1}{\exp(2\kappa)}$ is the optimal weight on observation, $w^*_{t+1} = w_{t+1} + \xi_{t+1}$ is the observed signal, and ξ_{t+1} are the i.i.d. endogenous noise with $\mathbf{var}(\xi_{t+1}) = \Sigma/\theta$.

Note that consumers with more channel capacity will choose to observe a less noisy signal about the state of the world because

$$\frac{\partial\mathbf{var}(\xi_{t+1})}{\partial\kappa} < 0 \text{ as long as } \kappa > \frac{1}{4}\log((1/\phi)^2).$$

Note that when γ is close to 1, $\phi = \beta$. Hence, for a typical quarterly estimation or calibration (say, $\beta = 0.99$), this above condition requires that $\kappa > 0.005$ nats; and for an annual exercise ($\beta = 0.96$), this requires that $\kappa > 0.02$ nats. Both required capacities seem very low and in the following

[35] This principle says that under the LQ assumption optimal control and state estimation can be decoupled. See Whittle (1982, 1996) for detailed discussions. Of course, this modified decision rule can be derived by solving the stochastic optimal control problem explicitly. The detailed derivations are available from the author by request.

discussions I assume this condition is always satisfied. Further, $\frac{\partial \mathbf{var}(\xi_{t+1})}{\partial \beta} < 0$ implies that more patient consumers choose to observe more carefully because they face higher costs in the future due to having a suboptimal asset holding level.

Hence, individual consumption growth can now be written as[36]

$$\Delta c_{t+1} = \Delta \widehat{w}_{t+1} = [\theta \chi u_{t+1} + \theta \chi \frac{((1-\theta)/\phi) u_t}{1 - ((1-\theta)/\phi) \cdot L}] + [\theta \xi_{t+1} - \theta \frac{(\theta/\phi) \xi_t}{1 - ((1-\theta)/\phi) \cdot L}] \tag{2.24}$$

where L is the lag operator. See Appendix C for the derivations.

Note that in the above expression for the change in consumption, the optimal share invested in stock market is still undetermined. In the next subsection, I will use this expression for consumption growth and the Euler equation to derive the optimal portfolio rule. It will be more clear that how optimal consumption rule and portfolio rule depends on each other in the RI model.

2.4. Long-term Consumption Risk and the Demand for the Risky Asset

Parker (2001; 2003) and Parker and Julliard (2005) argued that the long-term risk is a better measure of the true 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. Hence, we need to use the long-term consumption risk to measure the risk of the equity in the RI model because consumption reacts gradually and with delay to the innovations to the equity.

In this subsection, I first define the long-term consumption risk in the RI model and then derive the optimal portfolio rule. Substituting the optimal portfolio rule into the consumption rule and the changes in consumption gives us a complete solution to this simple optimal consumption and portfolio choice model with RI. In the next subsection, these results will be summarized in a proposition and I will further discuss how introducing RI reconciles two important phenomena in the US economy: the smoothness of consumption, the low contemporaneous covariance between consumption growth and the equity return, and the high equity premium.

Following Parker's work, I define the long-term consumption risk as the covariance of asset returns and consumption growth over the period of the return and many following periods. Because the RI model predicts that consumption reacts to the innovations to asset returns gradually and slowly, it can rationalize the assumption used in Parker's papers that consumption risk should be long term instead of contemporaneous. Given the above analytical solution for consumption growth, it is straightforward to calculate the ultimate consumption risk in the RI model. Specifically, when

[36] Note that this MA(∞) expression requires that $(1-\theta)/\phi < 1$, which is equivalent to $\kappa > \frac{1}{4}\log((1/\phi)^2)$.

consumers behave optimally but only have finite capacity, we have the following equality for the risky asset e and the risk free asset f[37]:

$$\mathbf{E}_t[R^e_{t+1}C^{-\gamma}_{t+1+S}] = \mathbf{E}_t[R^f C^{-\gamma}_{t+1+S}], \tag{2.25}$$

which can be transformed to the following stationary form:

$$\mathbf{E}_t[R^e_{t+1}(C_{t+1+S}/C_t)^{-\gamma}] = \mathbf{E}_t[R^f(C_{t+1+S}/C_t)^{-\gamma}] \tag{2.26}$$

where S is the number of periods in the future[38]. The standard equality $\mathbf{E}_t[R^e_{t+1}C^{-\gamma}_{t+1}] = \mathbf{E}_t[R^f C^{-\gamma}_{t+1}]$ does not hold here because consumption reacts slowly with respect to the innovations to equity returns and thus cannot adjust immediately and completely.

Log-linearizing equation (2.26) yields[39] the following pricing equation

$$\begin{aligned}\mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2 &= \gamma\mathbf{cov}_t[c_{t+1+S} - c_t, r^e_{t+1}] \\ &= \gamma\sum_{s=0}^{S}\mathbf{cov}_t[\Delta c_{t+1+s}, r^e_{t+1}]\end{aligned} \tag{2.27}$$

where I use the fact that $c_{t+1+S} - c_t = \sum_{s=0}^{S}\Delta c_{t+1+s}$.

Hence, substituting the expression for consumption growth, $\Delta c_{t+1+s} = [\theta\chi u_{t+1+s} + \theta\chi\frac{((1-\theta)/\phi)u_{t+s}}{1-((1-\theta)/\phi)\cdot L}] + [\theta\xi_{t+1+s} - \theta\frac{(\theta/\phi)\xi_{t+s}}{1-((1-\theta)/\phi)\cdot L}]$, into the above equation, it is easy to show that when S approaches to infinity, the ultimate consumption risk is larger than the contemporaneous consumption risk because

$$\lim_{S\to\infty}\sum_{s=0}^{S}\mathbf{cov}_t[\Delta c_{t+1+s}, r^e_{t+1}] = \frac{\theta}{1-(1-\theta)/\phi}\chi\omega_u^2 > \chi\omega_u^2, \tag{2.28}$$

that is, the impacts of the risk on consumption can last infinite following quarters. Note that since consumption adjusts gradually to the shocks to asset returns in the RI model, this ultimate consumption risk is the best measure of the riskiness of the equity.

Furthermore, since the return on the equity is i.i.d. with mean μ, I rewrite the above pricing

[37] The equality can be obtained by using $S+1$ period consumption growth to price a multiperiod return formed by investing in equity for one period and then transforming to the risk free asset for the next S periods. Hence, the following multiperiod moment condition holds

$$C_t^{-\gamma} = \mathbf{E}_t[\beta^{S+1}C^{-\gamma}_{t+S+1}R^e_{t+1}(R^f)^S].$$

[38] This measure has some appealing features, see Parker (2003) for discussion.

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

equation as follows

$$\mu - r^f + \frac{1}{2}\omega_u^2 = \gamma \frac{\theta}{1-(1-\theta)/\phi}\chi\omega_u^2 \tag{2.29}$$

where $\mu - r^f$ is the equity premium, χ is the fraction of wealth invested in the equity, $\phi = 1 - \exp(b_0)$[40], and the ultimate risk is represented by $\frac{\theta}{1-(1-\theta)/\phi} > 1$. It is worth emphasizing that b_0 itself also depends on χ if γ is different from 1, and when γ approaches 1, b_0 approaches β. Therefore, when γ is close to 1 from above, we have the following proposition about the optimal asset allocation in the risky asset.

Proposition 1. *The optimal asset allocation of an investor in the RI economy can be expressed by*

$$\chi = \vartheta \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\gamma\omega_u^2} \tag{2.30}$$

where $\vartheta = \frac{1-(1-\theta)/\phi}{\theta} < 1$ is inversely proportional to the ultimate consumption risk of the equity

Note that the ultimate consumption risk, $\frac{\theta}{1-(1-\theta)/\phi}$, is increasing with the degree of inattention. The following figure plots the relationship between the ultimate risk and channel capacity κ.

[Insert Figure 1 about here]

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 amount of risky assets.

As predicted by the standard CCAPM model, the optimal fraction of savings invested in the risky asset is proportional to the risk premium $(\mu - r^f)$ 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 is 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 shares in risky assets. The intuition for this result is simple. In the RI economy, a one percent negative shock in individuals' financial wealth would

[40] Note that b_0 is the coefficient of optimal consumption rule and defined in (2.8).

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 (2.30) as

$$\chi = \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\widetilde{\gamma}\omega_u^2} \tag{2.31}$$

where the *effective* coefficient of relative risk aversion is $\widetilde{\gamma} = \gamma/\vartheta > \gamma$. Hence, the asset allocation of the inattentive consumers is similar to the allocation of the more risk averse individuals[41]: both groups hold less risky assets.

Note that I assume at the beginning all investors have same preference and devote same amount of channel capacity in monitoring their wealth evolution, so they choose to invest the same fraction of their wealth in the equity. Consequently, the fraction of the total wealth invested in stock market in the RI economy is the same as the optimal share in the equity of individual investor. Therefore, we could use some stylized aggregate facts and this formula to calibrate the degree of average inattention in this economy. We have assumed that labor income is marketable and certain in this simple model and thus included in the total wealth W_t, so we can regard human capital as a kind of the riskless asset. Following Gabaix and Laibson (2001), using that fact that the ratio of total nonhuman capital[42] to total wealth is around 0.22, we have that

$$\frac{1-(1-\theta)/\phi}{\theta} \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\gamma\omega_u^2} = 0.22. \tag{2.32}$$

Hence, substituting $\gamma = 1.01$, $\phi = \beta = 0.96$, $\mu - r^f + \frac{1}{2}\omega_u^2 = 0.06$, $\omega_u^2 = 0.18^2$, we can calculate that θ is around 0.05, that is, if on average only 5% of the uncertainty about the wealth can be eliminated after observing and processing information, that the otherwise standard portfolio choice model can generate the observed share of total wealth invested in stock market. If we look at the standard model without RI, to generate the same fraction invested in stock market, we must have a much higher value of γ. For the same process of asset returns, this value is around 9. Note that this value is not extremely high as predicted in the asset pricing literature because it includes another puzzle: consumption is as volatile as wealth.

So far I have assumed that every agent has same channel capacity, however, in reality different agents would have different levels of channel capacity in observing and processing information. Hence, in this heterogeneous-attention case, consumers have different demands for the risky asset even if they have the same preference and face the same asset returns process. Based on the

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

[42] Here I assume that all capital is identical to stock market capital.

formula (2.30), it is straightforward to show that "limited stock market participation"[43] can arise endogeneously in the presence of RI. The intuition is that the demand for the equity is decreasing with the degree of inattention, so some consumers do not have any demand for the equity when the degree of their attention is low enough, that is, when the degree of RI, θ, reaches the critical value $1-\phi$, the demand for the equity is zero. Note that here I assume that there is no any transaction costs in this economy. So it might be very likely that those consumers with low capacity would also choose not to participate in the stock market even if the transaction cost is assumed to be very tiny and the demand for the equity is nonzero[44]. I won't explore this issue in detail here and would leave it for future research. Hence, the model composed of consumers with different degrees of RI can generate "limited stock market participation" endogeneously. Some consumers who know that they cannot devote enough capacity in processing information about their wealth would choose not to invest in stock markets.

2.5. A Complete Characterization of the Model's Dynamic Properties

Combining (2.30) with (2.22) gives us optimal consumption and portfolio rules in the RI model. The following proposition summarizes these results.

Proposition 2. *Given finite channel capacity κ, the optimal share invested in the equity is*

$$\chi^* = \vartheta \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\gamma \omega_u^2} \tag{2.33}$$

where $\vartheta = \frac{1-(1-\theta)/\phi}{\theta} < 1$; Furthermore, optimal consumption rule is then

$$c_t = b_0 + \widehat{w}_t \tag{2.34}$$

and the estimated state $\widehat{w}_t$ is characterized by the following Kalman filtering equation

$$\widehat{w}_t = (1-\theta)\widehat{w}_{t-1} + \theta w_t^* \tag{2.35}$$

where $\theta = 1 - \frac{1}{\exp(2\kappa)}$ is the optimal weight on observation, $w_t^* = w_t + \xi_t$ is the observed signal with ξ_t are the i.i.d. endogenous noise with $\mathbf{var}(\xi_{t+1}) = \frac{\chi^{*2}\omega_u^2}{\theta(\exp(2\kappa)-(1/\phi)^2)}$.

It is clear from this proposition that optimal consumption and portfolio rules are interdependent each other in the RI model. The optimal portfolio rule is similar to the standard Merton solution,

[43] See Vissing-Jørgensen (2002) for a detailed discussion on limited asset market participation.
[44] For example, when $\kappa = 0.02$ nats, $\gamma = 1.01$, $\beta = 0.96$, $\mu - r^f + \frac{1}{2}\omega_u = 0.06$, and $\omega_u = 16\%$, $\chi \simeq 0.01$.

where the usual CRRA is replaced by the effective CRRA. RI amounts therefore to an increase in the effective relative risk aversion.

For the case without RI, the ratio of consumption to wealth is constant, that is, $\frac{C_t}{W_t} = \exp(b_0)$. However, when introducing RI, this ratio is a stochastic process the true wealth cannot be observed perfectly. Formally, we have the following proposition:

Proposition 3. *The ratio of consumption to true wealth in the RI case is a stochastic process rather than a constant in both individual level and aggregate level, and both the expected ratio of individual consumers and the average ratio over all consumers are greater than the ratios implied by the model without RI. Furthermore, these ratios increase with the degree of RI.*

Proof. For any individual, we can rewrite the ratio of his consumption to true wealth level as follows

$$\frac{C_t}{W_t} = \frac{C_t}{\widehat{W}_t}\frac{\widehat{W}_t}{W_t}$$

where $\frac{C_t}{\widehat{W}_t} = \exp(b_0)$, $\frac{\widehat{W}_t}{W_t} = \exp(\widehat{w}_t - w_t) = \exp(\frac{-(1-\theta)\chi u_t + \theta\xi_t}{1-((1-\theta)/\phi)\cdot L} - \Omega_1)$, and $\chi = \vartheta\frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\gamma\omega_u^2}$. Hence, taking unconditional expectation on both sides yields

$$\mathbf{E}[\frac{C_t}{W_t}] = \exp(b_0)\exp[\frac{1}{2}((1-\theta)\chi)^2\frac{\omega_u^2}{1-((1-\theta)/\phi)^2} + \frac{1}{2}\theta^2\frac{\omega_\xi^2}{1-((1-\theta)/\phi)^2}] \tag{2.36}$$

Similarly, after aggregating over all consumers, we can also find that the average ratio of consumption to true wealth is also greater than the standard one because

$$\overline{(\frac{C_t}{W_t})} = \exp(b_0)\exp[\frac{1}{2}((1-\theta)\chi)^2\frac{\omega_u^2}{1-((1-\theta)/\phi)^2}], \tag{2.37}$$

where $\overline{(\frac{C_t}{W_t})}$ represents the average level of the ratio. Note that here all the private noises are cancelled out after aggregation. ■

Here the expected high ratio of consumption to wealth at the optimum is due to imperfect observations. In other words, Imperfect observation generates extra saving due to the fundamental shocks and the endogenous noises, and thus at the optimum increases the ratio of consumption to wealth.

Given these optimal rules, I next provide a general characterization of the dynamic properties of the economy described above. Specifically, I analyze several important properties of the RI economy: excess smoothness of consumption growth, low contemporaneous covariance between consumption growth and asset returns, postive autocorrelation of consumption growth, and non-zero covariance of consumption growth and lagged equity returns.

Substituting (2.30) into (2.24) yields

$$\Delta c_{t+1} = (1-(1-\theta)/\phi)\frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\gamma\omega_u^2}\left[u_{t+1} + \frac{((1-\theta)/\phi)u_t}{1-((1-\theta)/\phi)\cdot L}\right] + \left[\theta\xi_{t+1} - \theta\frac{(\theta/\phi)\xi_t}{1-((1-\theta)/\phi)\cdot L}\right] \tag{2.38}$$

Further, because ξ_t are i.i.d. information-processing induced endogenous idiosyncratic noise (a kind of private information) with a mean of zero, we could assume that in the expression for individual consumption growth all noise terms would be cancelled out when aggregating over all consumers[45]. Consequently, after aggregating[46], we have the following proposition

Proposition 4. *Aggregate consumption growth in the CCAPM model with RI can be written as*[47]

$$\Delta c_{t+1} = \theta\chi^* u_{t+1} + \theta\chi^*\frac{((1-\theta)/\phi)u_t}{1-((1-\theta)/\phi)\cdot L}, \tag{2.39}$$

which implies that 1) the covariance between aggregate consumption growth and asset returns is[48]

$$\mathbf{cov}(\Delta c_{t+1}, r_{t+1}^e) = \theta\chi^*\omega_u^2, \tag{2.40}$$

2) the standard deviation of consumption growth is

$$\sigma(\Delta c_{t+1}) = \lambda\omega_u \tag{2.41}$$

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

$$\rho(\Delta c_{t+1}, r_{t+1}^e) = \sqrt{1-((1-\theta)/\phi)^2}, \tag{2.42}$$

[45] Note that as discussed in Sims (2003), there might have been some aggregate component in the noise terms left after aggregating. If this is the case, we can see below that this additional term would affect the smoothness of aggregate consumption, but have no impact on the covariance between consumption growth and equity returns and thus the equity premium and optimal portfolio choice. The reason is that the noises and the innovations to equity returns are independent and thus have no correlation each other.

[46] For simplicity, we assume that individuals have identical channel capacity, and to avoid confusion, we still use c to represent aggregate consumption.

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

[48] *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* $\mathbf{cov}(\Delta c_{t+1}, r_{t+1}^e) = \theta\chi\lambda_s\omega_u^2$.

4) the autocorreltion of consumption growth is then

$$\rho_{\Delta c}(j) = \mathbf{corr}(\Delta c_t, \Delta c_{t+j}) = [1 - ((1-\theta)/\phi)^2][(1-\theta)/\phi]^j \tag{2.43}$$

where $j \geq 1$*, and 5) the covariance between consumption growth and lagged equity returns is*

$$\mathbf{cov}(\Delta c_{t+1}, r^e_{t+1-j}) = \theta\chi^*((1-\theta)/\phi)^j \omega_u^2 \tag{2.44}$$

where $j \geq 1$.

Proof. Based on the derivations in Appendix C, it is straightforward to obtain the above results. ■

Equation (2.39) means that aggregate consumption adjusts gradually to the shocks of asset returns, and thus, the contemporaneous covariance between consumption growth and asset returns becomes[49] $\theta\chi^*\omega_u^2$ rather than $\chi\omega_u^2$[50], that is, the measured contemporaneous covariance between consumption growth and risky returns will be lowered by a factor $1-(1-\theta)/\phi$. In the above simple CCAPM model without RI, this contemporaneous covariance is around $1.23 \cdot 10^-3$ at the quarterly frequency, and this figure is well above its US empirical counterpart which is around $0.08 \cdot 10^-3$. When $\theta = 0.1$, the theoretical covariance value becomes $1.23 \cdot 10^-4$, which is much closer to the empirical value.

Equation (2.41) means that RI can also reduce the standard deviation of consumption growth since λ is less than 1 . Figure 2 shows the relationship between the smoothness of consumption growth and channel capacity. It is clear from this figure that the smoothness of consumption growth is increasing with the degree of inattention. 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 CCAPM 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.

[Insert Figure 2 about here]

Equation (2.43) shows that the autocorrelation of aggregate consumption growth is greater than 0[51]. This finding is roughly consistent with some empirical evidence[52] that the autocorrelation is

[49] 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 $\theta\chi\lambda_s\omega_u^2$.

[50] Note that $\chi = \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\gamma\omega_u^2}$ is the optimal share invested in stock market in the model without RI.

[51] Note that the standard CCAPM case implies that the autocorrelation of consumption growth is 0, that is, consumption growth is i.i.d.

[52] For example, Piazzesi (2001) reported the autocorrelation of consumption growth at different lags together with 95% confidence bounds and found that they are significant in the first several quarters. See section 1.2 in her paper

significant up to the third quarter, which means that consumption growth is definitely not i.i.d[53].

Equation (2.44) shows that the covariance between consumption growth and lagged equity returns is greater than zero, that is, lagged equity returns can be used to predict future consumption growth. Further, the impacts of lagged returns on future consumption growth are decreasing at the rate $(1-\theta)/\phi$. It is obvious that in the absence of RI, lagged returns have no impact on consumption growth. In the last section, I discuss some existing empirical evidence that shows that consumers may not have enough knowledge about their financial wealth and lagged returns can be used to predict future consumption growth.

From the above proposition, it is obvious that given the empirical evidence, we can use these dynamic properties, for example, (2.42) or (2.43)[54] to calibrate the average degree of inattention in the economy governed by the deep parameter, κ. For example, the observed contemporaneous correlation between aggregate consumption growth and asset returns is around 0.34 in the US quarterly data[55]; hence one can calibrate the value of θ by setting $\sqrt{1-((1-\theta)/\phi)^2} = 0.23$ and using the fact that $\phi \simeq \beta$. After choosing $\beta = 0.96$, we have a calibrated θ equal to 0.1, that is, $\kappa = 0.07$ bits, which means that to match the observed correlation between consumption growth and equity returns, the average degree of RI in the economy would be quite low. Further, using the expression for the autocorrelation $\rho_{\Delta c}(j) = [1-((1-\theta)/\phi)^2][(1-\theta)/\phi]^j$ where $j \geq 1$ and the empirical counterparts estimated from the US data, we may also calibrate κ. Piazzesi (2001) adopted the maximum likelihood method to estimate the US log aggregate consumption growth composed of an AR(3) process combined with a Garch (1,1) process and found that $\rho_{\Delta c}(1) = 0.31$ with T-statistics 4.01[56]. Hence, setting $[1-((1-\theta)/\phi)^2][(1-\theta)/\phi] = 0.31$ yields $\theta = 0.27$, that is, $\kappa = 0.22$ bits, which also implies a low average level of inattention. Note that here I did not use confidence bounds based on standard errors. Hence, I cannot determine what value of κ can satisfy all observed evidence exactly. However, it is obvious that a low value of κ can roughly capture some main properties of the US data about consumption growth and asset returns: the standard deviation of aggregate consumption, the correlation between consumption growth and asset returns, and the autocorrelation of consumption growth.

---

for a detailed description of the estimation.

[53] Of course, the empirical evidence on serial correlation in aggregate consumption growth is mixed. For example, Measurement problems may bias these autocorrelation in either direction. Empirical estimates of discrete-time Markov models by Kandel and Stambaugh (1991) and Mehra and Prescott (1985) also found some evidence for modest predictable variation in US consumption growth, whereas Hall (1988), Cochrane (1997), Lettau and Ludvigson (2001) found that US consumption growth is almost unforecastable.

[54] Note that both expressions do not depend on the optimal share of the risky asset.

[55] Campbell (2002) reported the correlations for different horizons in the quarterly data.

[56] For detailed estimation results, see table 1 in Piazzesi (2001).

2.6. Revisiting the Equity Premium Puzzle

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

$$\widetilde{\gamma} = \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\mathbf{cov}[c_{t+1} - c_t, r_{t+1}^e]} = \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\theta \chi^* \omega_u^2} > \gamma \simeq 1,$$

where $\widetilde{\gamma}$ is the estimated CRRA and γ is the true value that is around to be close to 1. The intuition is quite simple. The RI model can generate very low value of $\mathbf{cov}[c_{t+1} - c_t, r_{t+1}^e]$ and high equity premium simultaneously because the equity premium is determined by the long term consumption risk, $\lim_{S\to\infty} \mathbf{cov}[c_{t+1+S} - c_t, r_{t+1}^e]$. In other words, the estimate of the coefficient of relative risk aversion will be biased upward by a factor measured by the long-term consumption risk. For example, if the true γ is 1 and $1/\vartheta = 10$, he will find that the estimated value $\widetilde{\gamma}$ will be 10. Hence, the estimated high CRRA in the asset pricing literature might arise from a low degree of the average attention in the economy instead of risk aversion itself. This result can reconcile two sets of empirical evidence: 1) relatively low values of risk aversion in introspection and experimental evidence and 2) high values of 'risk aversion' inferred from consumption and asset prices data.

Given the smoothness of consumption and low contemporaneous covariance between consumption growth and asset return, we can see where the equity premium puzzle arises from the following standard pricing equation:

$$\mathbf{E}_t[r_{t+1}^e] - r^f + \frac{1}{2}\omega_u^2 = -\mathbf{cov}_t[m_{t,t+1}, r_{t+1}^e] \tag{2.45}$$

where $m_{t,t+1} = -\gamma \Delta c_{t+1}$ is a one-period stochastic discount factor and the equation says that the equity premium is determined by the negative covariance of the asset with the stochastic discount factor (SDF). Following Hansen and Jagannathan (1991) and Cochrane and Hansen (1992), we can rewrite the equation as follows:

$$\sigma_m \geq \frac{\mathbf{E}_t[r_{t+1}^e] - r^f + \frac{1}{2}\omega_u^2}{\sigma_e} \tag{2.46}$$

where σ_m is the standard deviation of the SDF, σ_e is the standard deviation of the asset return, and here we use the fact that the correlation between the asset return and the SDF, $\rho_{m,e} \geq -1$, that is, $-\mathbf{cov}_t[m_{t,t+1}, r_{t+1}^e] \leq \sigma_e \sigma_m$.

The right hand side of (2.46) is a logarithmic Sharpe ratio for the asset, which is defined as the excess return on an asset, adjusted for Jensen's inequality, divided by the standard deviation of the asset return. Hence, (2.46) says that the standard deviation of the log SDF must be greater than the Sharpe ratio for this risky asset. This inequality can be used to illustrate the equity premium puzzle. In the postwar US data, the estimated low bound for σ_m is greater than 50%. However, given the smoothness of consumption and the low correlation between consumption growth and asset returns, the lower bound can be achieved only if the implied risk aversion coefficients are extremely high or the EIS is extremely low[57]. Lettau and Uhlig (2001) computed the Sharpe ratios in the log normal framework with various types of habit formation and the recursive preference and showed that none of these preferences is able to generate a high Sharpe ratio while keeping risk aversion fairly low and the EIS fairly high, as would be desirable. Later, I will show that how the RI hypothesis can do a better job in generating a high Sharpe ratio.

As shown in Weil (1989), simply accepting high risk aversion would bring about another puzzle, the risk free rate puzzle. Note that the linearized equation for the risk free rate is

$$r^f = -\log\beta + \gamma\mathbf{E}[\Delta c_{t+1}] - \frac{1}{2}\gamma(\gamma+1)\mathbf{var}[\Delta c_{t+1}]. \tag{2.47}$$

Hence, if γ takes a high value, say 30, with $\mathbf{E}[\Delta c_{t+1}] = 1\%$ and $\mathbf{var}[\Delta c_{t+1}] = 0.015^2$, the risk free rate is around 20% even if $\beta = 1$! In other words, in the expected utility function, we cannot resolve the equity premium puzzle by simply setting high value of risk aversion. In the next subsection, we shall see that introducing RI can increase the effective risk aversion without changing the true risk aversion, thus providing a potential explanation for the equity premium puzzle and the risk free rate puzzle.

We can also see how RI can help resolve the equity premium by looking at the Sharpe ratio in the RI economy. (2.27) implies that we use the $S-$period SDF, $m_{t,t+S+1}$ to price the risky asset. As in the previous subsection, we have

$$\sigma_m^S \geq \frac{\mathbf{E}_t[r_{t+1}^e] - r^f + \frac{1}{2}\omega_u^2}{\sigma_e} \tag{2.48}$$

where σ_m^S is the standard deviation of the $S-$period SDF. Since

$$-m_{t,t+S+1} = \gamma\sum_{s=0}^{S}\Delta c_{t+1+s} \text{ and } \Delta c_{t+1+s} = \theta\chi u_{t+1+s} + \theta\chi\frac{((1-\theta)/\phi)u_{t+s}}{1-((1-\theta)/\phi)\cdot L},$$

it is apparent that

$$\sigma_m^S > S\cdot\sigma_m \tag{2.49}$$

[57]See table 4 in Campbell (2003) for a detailed description.

where $\sigma_m = \gamma\theta\chi\sqrt{1+\frac{((1-\theta)/\phi)^2}{1-((1-\theta)/\phi)^2}}\omega_u$ is the standard deviation of the 1−period SDF. Hence, RI can substantially increase the volatility of the SDF and then make the SDF enter the Hansen-Jagannathan volatility bound.

It is worth emphasizing that if the above long-term Euler equation holds, then the actual IES in the RI economy is not $1/\gamma$, but a value less than $1/\gamma$. To see this, we operate on the following equation:

$$C_t^{-\gamma} = \mathbf{E}_t[\beta^{S+1} C_{t+S+1}^{-\gamma} R_{t+1}^p (R^f)^S]. \tag{2.50}$$

Note that this equation is used to price a multiperiod return formed by investing in the portfolio for one period and then transforming to the risk free asset for the next S periods. Because IES is about the willingness to substitute intertemporally, we eliminate the uncertainty here and then have

$$\frac{d(\log C_{t+S+1} - \log C_t)}{d(\log R_{t+1}^p)} = \frac{d(\sum_{s=0}^{S} \Delta \log C_{t+s+1})}{d(\log R_{t+1}^p)} = \frac{1}{\gamma}. \tag{2.51}$$

Therefore, the true IES is equal to

$$\frac{d(\Delta \log C_{t+1})}{d(\log R_{t+1}^p)} < \frac{1}{\gamma}. \tag{2.52}$$

2.7. Implications for the Cross-sectional Returns

In the previous subsection, I focus on the time series properties of the excess returns in the RI model. In this subsection, I will briefly discuss the implications of RI for the cross-sectional pattern of risky returns. In Parker (2003) and Parker and Julliard (2005), they found in the data that although contemporaneous consumption risk has little predictive power for explaining the pattern of average returns across the Fama-French (25) portfolios, the ultimate consumption risk can explain the average returns. In this subsection, I will show how RI can generate larger differences across risky portfolios if we measure the risk correctly.

For simplicity, I assume that there are two risky portfolios and one riskless asset in the economy. We use r_{t+1}^1 and r_{t+1}^2 to represent the returns on these two risky assets, respectively. Further, I assume that r_{t+1}^i is i.i.d. with a mean of μ^i and standard deviation ω_i, where $i = 1, 2$. Denote the market portfolio of consumers as

$$r_{t+1}^p = \chi_1 r_{t+1}^1 + \chi_2 r_{t+1}^2 + (1 - \chi_1 - \chi_2) r^f. \tag{2.53}$$

Using the same procedure as above, we have the following two pricing equations for the two risky portfolios:

$$\mathbf{E}_t[r_{t+1}^1] - r^f + \frac{1}{2}\omega_1^2 = \gamma \mathbf{cov}_t[c_{t+1+S} - c_t, r_{t+1}^1], \tag{2.54}$$

$$\mathbf{E}_t[r_{t+1}^2] - r^f + \frac{1}{2}\omega_2^2 = \gamma \mathbf{cov}_t[c_{t+1+S} - c_t, r_{t+1}^2]. \tag{2.55}$$

Hence, the optimal shares invested in portfolios 1 and 2 are

$$\chi_1 = \frac{\mu^1 - r^f + \frac{1}{2}\omega_1^2}{\widetilde{\gamma}\omega_1^2} \text{ and } \chi_2 = \frac{\mu^2 - r^f + \frac{1}{2}\omega_2^2}{\widetilde{\gamma}\omega_2^2},$$

respectively, where $\widetilde{\gamma} = \gamma/\vartheta$ is the effective CRRA due to RI. It is clear that given the properties of the financial assets (the mean and standard deviation of the assets), RI reduces the optimal shares invested in the two risky assets proportionally.

As usual, we could also think about this problem from the reverse direction. If given the shares invested in the risky assets and the volatility of the equities, we can see how RI change the expected returns on risky assets. Differencing the two equations gives

$$\mathbf{E}_t[r_{t+1}^1 - r_{t+1}^2] = \widetilde{\gamma}[\chi_1\omega_1^2 - \chi_2\omega_2^2] \tag{2.56}$$

Hence, because $\widetilde{\gamma} = \gamma/\vartheta$ is composed of two components: the true CRRA γ and the ultimate consumption risk $1/\vartheta(>1)$, the degree of RI that determines the ultimate risk plays a role in affecting the cross-sectional returns. (2.56) implies that the higher the degree of RI, the higher the expected difference between the two risky assets is. In sum, here RI works as a multiplier that can generate higher differences in the expected returns of the two risky assets.

2.8. Comparisons with Robustness Hypothesis, Recursive Utility, and Habit Formation

Maenhout (2004) derived optimal consumption and portfolio rules that are robust to model misspecification (that is, parameter uncertainty) in the lines of Anderson, et. al (2002). In that framework, optimal decisions are designed such that they not only work well when the structural model holds exactly, but also perform reasonably well when there is some kind of model misspecification. One possible misspecification is based on the assumption that the decision maker worries about some worst-case scenario. In this case, the disparity between the reference model and the worst-case alternative model is constrained by a parameter governing the degree of the preference for robustness. As a result, investors are very conservative or pessimistic when they form a risky portfolio, and part of pessimism can be due to robustness or uncertainty rather than to risk aversion itself.

Maenhout 's model used the continuous-time framework with CRRA utility and derived optimal consumption and portfolio choice (equation (16)[58] and (17) in Maenhout (2004)) as follows

$$C^* = aW_t \tag{2.57}$$

$$\alpha^* = \frac{1}{\gamma + \widetilde{\theta}} \frac{\mu - r^f}{\sigma^2} \tag{2.58}$$

where $a = \frac{1}{\gamma}[\delta - (1-\gamma)r - \frac{1-\gamma}{2(\gamma+\widetilde{\theta})}(\frac{\mu - r}{\sigma^2})^2]$, γ is the CRRA, δ is the discount rate, μ is the mean equity return, r is the risk free rate, σ is the standard deviation of the price of the equity, and $\widetilde{\theta}$ measures the preference for robustness[59]. When $\widetilde{\theta} = 0$, it reduces to the familiar Merton model. Note that in the expression for α^*, the true risk aversion coefficient γ is adjusted by $\widetilde{\theta}$. Hence, robustness amounts to an increase in the effective CRRA, that is, $\gamma + \widetilde{\theta}$, and then reduces the demand for risky assets. Furthermore, the robust consumption rule has a similar structure as that derived from the Merton's model, and the only difference is that the key parameter, a, reflects the difference in portfolio choice. Compare these results with the optimal consumption rule and portfolio choice derived from our RI model, and we have the following proposition:

Proposition 5. *A CRRA investor with RI is observationally equivalent to a CRRA investor with a preference for robustness in the sense that both invest less in the risky asset due to a higher effective CRRA. However, they can be distinguished by their consumption rules.*

Proof. It is straightforward by comparing (2.30) and (2.58) for optimal portfolio choice, and (2.22) and (2.57) for consumption rule. ■

An advantage of this equivalence is that we can infer the values of the parameter governing RI from the values of the preference governing robustness. For the range of the values of preference for robustness, see Maenhout (2004). Maenhout discussed the relationship between robustness and recursive utility and concluded in proposition 2 in his paper that an investor with CRRA($=\gamma$) utility and a preference for robustness ($\widetilde{\theta}$) is observationally equivalent to a Duffie-Epstein-Zin investor with EIS $1/\gamma$ and CRRA $\widetilde{\theta} + \gamma$. In our RI model, the effective CRRA is $\gamma/\vartheta (> \gamma)$ and the true IES is less than $1/\gamma$.

Luo (2005) showed that in the permanent income hypothesis framework RI and internal additive habit formation (HF) can generate similar dynamics of consumption, savings, and wealth accumulation if the degree of RI and the degree of HF satisfy some condition. Although both

[58] For simplicity here I set $T = \infty$.

[59] Maenhout (2004) used θ to denote the preference for robustness. To avoid the confusion with the optimal observation weight, θ, defined in section 2, here I use $\widetilde{\theta}$ to represent the preference for robustness.

hypotheses are quite different from each other, both of them predict that current consumption not only depends on permanent income, but also on past consumption. In the RI model people have to do so due to finite Shannon capacity, whereas in the HF model people are willing to do so because they like to smooth consumption growth. In this CCAPM framework with CRRA preference, RI and HF also have similar impacts on consumption dynamics. This conclusion is clear if we assume that HF enters the utility function by this way:

$$\frac{(C_t - \alpha C_{t-1})^{1-\gamma}}{1-\gamma}, \tag{2.59}$$

where C_t and C_{t-1} are individual consumption at time t and $t-1$, respectively. As shown in Dynan (2001), after approximating $\Delta \log(C_t - \alpha C_{t-1})$ with $\Delta \log C_t - \alpha \Delta \log C_{t-1}$, we can express the change in logarithmic consumption by a MA(∞) process given i.i.d. equity returns. Compared with (2.39), when $\alpha = (1-\theta)/\phi$, both models predict similar consumption dynamics.

3. RI and Recursive Utility

In the previous section, I showed that RI can reduce the demand for risky asset and could be an alternative explanation for the equity premium puzzle in the time-separable expected utility framework. However, to match the observed high equity premium, we must assume that channel capacity is extremely small. In this section, I consider the effects of RI on asset returns and portfolio choice in the recursive utility framework proposed by Epstein and Zin (1989) and Weil (1989). A crucial attribute of this recursive framework is that it allows us distinguish the CRRA and the EIS and then it provides a suitable framework to examine the interaction among RI, risk aversion, and intertemporal substitution in determining asset returns and portfolio choice[60]. Specifically, in this section I will discuss that how incorporating RI into Epstein-Zin-Weil's recursive utility model can go further in explaining the equity premium puzzle and the risk free rate puzzle.

Following Epstein and Zin (1989), Weil (1989), Campbell and Viceira (1999), and others, the consumer's preference is described by the following utility:

$$U(C_t, E_t U_{t+1}) = \{(1-\beta) C_t^{(1-\gamma)/\rho} + \beta (E_t U_{t+1}^{1-\gamma})^{1/\rho}\}^{\rho/(1-\gamma)} \tag{3.1}$$

where $\beta < 1$ is the discount factor, $\gamma > 0$ is the CRRA, σ is the EIS, and the parameter ρ is defined

[60] Although Epstein and Zin (1991) showed that this recursive utility can better explain the equity premium puzzle and the risk free rate puzzle, Kocherlakota (1996) argued that this conclusion is not correct because Epstein and Zin used the value-weighted return to the NYSE as a proxy to the gross real return to the representative agent's portfolio of assets, but this approximation understated the true level of diversification of this agent.

as $\rho = (1-\gamma)/(1-\sigma^{-1})$. Hence, when $\gamma = \sigma^{-1}$, the above nonlinear recursion becomes linear and the recursive utility reduces to the standard time-separable power utility.

Given i.i.d. equity returns, using the same budget constraint (2.2), we can obtain the optimal consumption and portfolio rules as follows

$$c_t = w_t + b_0$$
$$\chi = \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{[\frac{\rho}{\sigma} + (1-\rho)]\omega_u^2}$$

where $b_0 = \log[1 - \beta^{\frac{1}{\rho}}(\mathbf{E}_t(R_{t+1}^p)^{1-\gamma})^{\frac{1-\rho}{\rho}\frac{1}{1-\gamma}}]$. Furthermore, the value function of the agents with recursive utility in the full-information case also takes the form

$$V(W_t) = \Phi W_t^{1-\gamma}$$

where $\Phi > 0$ is a constant whose value is irrelevant for the present discussion. When γ is close to 1 (but different from 1, otherwise $\rho = 0$ and the model reduces to the static CAPM), as shown in section 2, the loss function is also log-quadratic. Hence, using the same argument in the CRRA case discussed in section 2, it is straightforward to show that the conditional distribution of the true state is also normal approximately in this recursive utility framework, and we choose the noise due to imperfect observations to be normal such that the joint distribution of the true state and the observed state is normally distributed. Following the same procedure as used in section 2, we have

$$c_t = \widehat{w}_t + b_0$$

where $\widehat{w}_t$ follows the same kalman filtering equation (2.23). Given this modified consumption rule, we can easily derive the change in consumption as follows

$$\Delta c_{t+1} = \Delta \widehat{w}_{t+1}, \tag{3.2}$$

Note that here the expression of $\Delta \widehat{w}_{t+1}$ is the same MA(∞) as the one derived in the CRRA case except some constant terms that depend on the mean consumption wealth ratio, b_0, which is a function of risk aversion, intertemporal substitution, and optimal portfolio choice. Similarly, aggregating over all consumers in the economy may eliminate all private noises and only leave the fundamental shocks.

As shown in Epstein and Zin (1991), the recursive utility hypothesis can help resolve the equity premium puzzle. Given the same flow budget constraint (2.2) as in section 2, we have the following linearized pricing equation

$$\mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2 = \frac{\rho}{\sigma}\mathbf{cov}_t[c_{t+1} - c_t, r^e_{t+1}] + (1-\rho)\mathbf{cov}_t[r^p_{t+1}, r^e_{t+1}] \tag{3.3}$$

where $r^p_{t+1} \simeq \chi(r^e_{t+1} - r^f) + r^f$. The equation says that the expected equity premium is a weighted average of two covariances: the first is with consumption growth divided by the IES and gets weight ρ, and the second is with the return on the market portfolio and gets weight $(1-\rho)$[61].

As in the above expected utility maximization model, we now assume that consumers cannot observe the state perfectly due to finite Shannon capacity. The usual standard Euler equation,

$$U_{1,t} = \mathbf{E}_t[U_{2,t}U_{1,t+1}R^e_{t+1}\ ] = \mathbf{E}_t[U_{2,t}U_{1,t+1}R^f\ ], \tag{3.4}$$

does not hold, and instead a long-term Euler equation holds[62], that is,

$$U_{1,t} = \mathbf{E}_t[(\beta_t \cdots \beta_{t+S})U_{1,t+1+S}R^e_{t+1}(R^f)^S] = \mathbf{E}_t[(\beta_t \cdots \beta_{t+S})U_{1,t+1+S}(R^f)^{S+1}]. \tag{3.5}$$

where $U_{i,t}$ denotes the derivative of the aggregate function with respect to its $i-$th argument and $\beta_t = U_{2,t}$ is the discount factor[63]. Following the same procedure used above, we have the following log-linearized pricing equation

$$\mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2 = \frac{\rho}{\sigma}\mathbf{cov}_t[c_{t+1+S} - c_t, r^e_{t+1}] + (1-\rho)\mathbf{cov}_t[r^p_{t+1} + \cdots + r^p_{t+1+S}, r^e_{t+1}] \tag{3.6}$$

Since we assume that equity returns are i.i.d., $\mathbf{cov}_t[r^p_{t+1} + \cdots + r^p_{t+1+S}, r^e_{t+1}] = \chi\omega_u^2$. Using the results in the CRRA case, we can easily derive the optimal share invested in equity as follows

$$\chi = \frac{\mu - r^f + \frac{1}{2}\omega_u^2}{\frac{\rho}{\sigma}\left(\frac{\theta}{1-(1-\theta)/\phi}\right)\omega_u^2 + (1-\rho)\omega_u^2} \tag{3.7}$$

where $\frac{\theta}{1-(1-\theta)/\phi}$ measures the ultimate consumption risk. Note that here we use the facts that $\mathbf{cov}_t[c_{t+1+S} - c_t, r^e_{t+1}] = \frac{\theta}{1-(1-\theta)/\phi}\chi\omega_u^2$ and stock returns are i.i.d. It is clear that when $\rho = 1$

[61] It is easy to see that when $\rho = 1$, it reduces to equation (2.11). When $\gamma = 1$, then $\rho = 0$ and a logarithmic version of the static CAPM pricing equation holds. Furthermore, when σ approaches 1, the coefficient ρ goes to infinity. Giovannini and Weil (1989) showed that in this case the optimal consumption rule is myopic (in the sense that the consumption and wealth ratio is constant), whereas optimal portfolio choice is not myopic unless γ is also 1.

[62] It is equivalent to

$$\mathbf{E}_t[(\beta_t \cdots \beta_{t+S})U_{1,t+1+S}R^e_{t+1}] = \mathbf{E}_t[(\beta_t \cdots \beta_{t+S})U_{1,t+1+S}R^f]$$

[63] Hence, with time additive and expected utility, the discount factor $U_{2,t} = \beta$ is constant and then this Euler equation reduces to the one discussed in section 2.

($\gamma = 1/\sigma$), the above solution reduces to (2.30) derived in the CRRA case.

Expression (3.7) implies that the total risk in the asset allocation problem is a weighted average of the ultimate consumption risk and the innovation to equity return itself. The weights are $\frac{\rho}{\sigma}$ and $1-\rho$, respectively. As discussed above, we assume that γ is slightly greater than 1 such that the approximation is valid. It is straightforward to show that how the interaction between RI, RRA, and EIS affects the optimal demand for the risky asset. First, keeping IES σ unchanged, raising RRA γ can reduce the optimal share χ in the presence of RI since $\frac{\theta}{1-(1-\theta)/\phi} > 1$. Second, if keeping γ unchanged, reducing σ would reduce the weight $\frac{\rho}{\sigma}$ and increase the weight $1-\rho$. Consequently, the related importance of RI on the first weight and the demand for the equity is reduced. As I assumed above, the riskfree asset, bond, is inside bond, the share of the risky asset is 100% in equilibrium. Hence, the above pricing equation means that RI has larger impacts on the equity premium in the case with high RRA and IES.

4. Stochastic Investment Opportunities

So far, I have assumed that stock returns are i.i.d. and have constant expected returns. However, the empirical evidence for the predictability of expected stock returns is strong. See Campbell and Viceira (1999, 2002) for detailed discussions. In this section, as in Campbell and Viceira (1999), I assume that the expected excess log return on the risky asset is dependent on a state variable x_t, such that

$$\mathbf{E}_t[r_{t+1}^e] - r^f = x_t \tag{4.1}$$

and the unique state x_t follows a AR(1) process:

$$x_{t+1} = \mu + \rho_x(x_t - \mu) + \epsilon_{t+1} \tag{4.2}$$

where ϵ_{t+1} is conditionally homoskedastic and $\epsilon_{t+1} \sim N(0, \omega_\epsilon^2)$. Also, I assume that the unexpected log return on the risky asset, u_{t+1}, is conditionally homoskedastic and normally distributed with $N(0, \omega_u^2)$. Finally I assume the two random components are correlated, $\mathbf{cov}_t[u_{t+1}, \epsilon_{t+1}] = \omega_{u\epsilon}$. Given the standard flow budget (2.2) and the recursive utility (3.1) defined in last section, following the same procedure proposed by Cambpell and Viceira (1999)[64], we have the following equation

$$\mathbf{E}_t[r_{t+1}^e] - r^f + \frac{1}{2}\omega_u^2 = \frac{\rho}{\sigma}\mathbf{cov}_t[c_{t+1} - c_t, r_{t+1}^e] + (1-\rho)\mathbf{cov}_t[r_{t+1}^p, r_{t+1}^e] \tag{4.3}$$

[64]See Section III in their paper for detailed derivations.

where $r^p_{t+1} \simeq \chi_t(r^e_{t+1} - r^f) + r^f$.

Based on this equation, I will discuss the implications of RI for equity premium and portfolio selection in several special cases[65].

Case 1 When $\sigma \to 1$ and $\gamma \neq 1$, the optimal portfolio rule is not myopic, while consumption choice is.

As shown in Giovannini and Weil (1989) and Cambpell and Viceira (1999), in this case the consumption rule is myopic, that is, the consumption wealth ratio is constant, $c_t - w_t = \log(1 - \beta)$. The optimal portfolio rule is not myopic and affected by the intertemporal hedging demand. Borrowing the results in Cambpell and Viceira (1999), we have the following optimal portfolio rule

$$\chi_t = a_0 + a_1 x_t \tag{4.4}$$

where

$$a_0 = \frac{1}{2\gamma} - (\frac{1}{1-\sigma})(\frac{\gamma - 1}{\gamma})\frac{\omega_{u\epsilon}}{\omega_u^2}; \tag{4.5}$$

$$a_1 = \frac{1}{\gamma\omega_u^2}. \tag{4.6}$$

Note that when $\omega_{u\epsilon} = 0$, the results becomes to the familiar Merton's solution, $a_0 = \frac{1}{2\gamma}$ and $a_1 = \frac{1}{\gamma\omega_u^2}$. The second term in a_0 measures the hedging demand of the risky asset.

As shown in Epstein and Weil (1989; 1991), the value function per unit of wealth in the recursive utility framework can be expressed as

$$V_t = (1-\beta)^{-\sigma/(1-\sigma)}(\frac{C_t}{W_t})^{1/(1-\sigma)} \tag{4.7}$$

where $V_t = U_t/W_t$. In this case, V_t is a constant in that $\frac{C_t}{W_t} = 1 - \beta$. Consequently, here we do not have a value function that depends on the state, and thus we cannot use the procedure in section 2 to derive the conditional distribution of the state in the presence of RI. Hence, to introduce RI in this model, we need to define a new loss function and assume that the consumers choose the conditional distribution to minimize this loss function subject to information-processing constraints. For simplicity, I assume that the loss function is quadratic. Hence, we can apply the results from the CRRA case here, and it is clear that the true state x_t, given information available at time t,

[65] I use these two extreme cases because they can be solved in close-form and help us get much insights about the implications of RI when investment opportunities are not constant.

follows a normal distribution. Given the true process of x_t, equation (4.2), the estimated process follows the following process

$$\widehat{x}_{t+1} - \mu = (1-\theta)\rho_x(\widehat{x}_t - \mu) + \theta(x_{t+1} - \mu + \xi_{t+1}), \tag{4.8}$$

and the decision rule becomes

$$\chi_{RI,t} = a_0 + a_1\widehat{x}_t \tag{4.9}$$

Proposition 1. *In Case 1, RI reduces the speed of adjusting the optimal portfolio with respect to the innovations in every period.*

Proof. Without RI, the agents adjust their optimal portfolio according to the expected excess return on equities. As this unique state follows an AR(1) process, past information can be used to predict their future values because x_t can written as a MA(∞) process, $x_t = \mu + \frac{\epsilon_t}{1-\rho_x \cdot L}$. With RI, the decision rule depends on the estimated state, $\widehat{x}_t = x_t + \frac{(\theta-1)\epsilon_t + \theta\xi_t}{1-(1-\theta)\rho_x \cdot L}$. Note that here we use the fact that combining the true process (4.2) and the Kalman filtering equation (4.8) implies that $\widehat{x}_{t+1} - x_{t+1} = \frac{(\theta-1)\epsilon_{t+1} + \theta\xi_{t+1}}{1-(1-\theta)\rho_x \cdot L}$. Hence,

$$\begin{aligned}\chi^*_{RI,t} &= a_0 + a_1[x_t + \frac{(\theta-1)\epsilon_t + \theta\xi_t}{1-(1-\theta)\rho_x \cdot L}] \\ &= a_0 + a_1[\mu + \theta\epsilon_t + \sum_{j=1}^{\infty}\rho_x^j(1-(1-\theta)^{j+1})\epsilon_{t-j} + \theta\sum_{j=0}^{\infty}((1-\theta)\rho_x)^j\xi_{t-j}],\end{aligned} \tag{4.10}$$

that is, the optimal portfolio choice also adjusts slowly with respect to the innovations to the expected return on the risky asset. ■

On the other hand, with RI and $\sigma = 1$, the following equation holds

$$\mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2 = \rho\mathbf{cov}_t[c_{t+1+S} - c_t, r^e_{t+1}] + (1-\rho)\mathbf{cov}_t[r^p_{t+1} + \cdots + r^p_{t+1+S}, r^e_{t+1}]. \tag{4.11}$$

Following the same procedure used in the CRRA case, we also need to use the ultimate risk to determine the optimal allocation in risky asset. The above equation implies that

$$\mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2 = \mathbf{cov}_t[r^p_{t+1} + \cdots + r^p_{t+1+S}, r^e_{t+1}] \tag{4.12}$$

where we use the fact that $\Delta c_{t+1} = (c_{t+1} - w_{t+1}) - (c_t - w_t) + \Delta w_{t+1} = \Delta w_{t+1}$.

Therefore, in this special case the expected excess return of equity is determined by the conditional covariance between equity return at time $t+1$ and the portfolio return at the same period as well as many following periods. Because $r^p_{t+1} = \chi_t r^e_{t+1} + (1-\chi_t) r^f$, $r^e_{t+1} = x_t + u_{t+1}$, and x_t cannot be observed perfectly due to RI, we have

$$\begin{aligned}\mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2 &= \chi_t \ \mathbf{cov}_t[r^e_{t+1}, r^e_{t+1}] + \sum_{s=2}^{S} \mathbf{cov}_t[r^p_{t+s}, r^e_{t+1}] \\ &= \chi_t[\omega_u^2 + \sigma_x^2] + \sum_{s=2}^{S} \mathbf{cov}_t[\chi_s r^e_{t+s}, r^e_{t+1}].\end{aligned}$$

where $\sigma_x^2 = \mathbf{var}_t[x_t]$. Rearranging the above equation gives

$$\chi_t = \frac{\mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2}{\omega_u^2 + \sigma_x^2} - \frac{\sum_{s=2}^{S} \mathbf{cov}_t[\chi_{t+s-1} r^e_{t+s}, r^e_{t+1}]}{\omega_u^2 + \sigma_x^2} \tag{4.13}$$

This equation takes a recursive form and has two components, each one capturing different aspects of asset demand. The first term captures the part of asset demand induced by the current risk premium adjusted by the estimation risk σ_x^2 due to RI. The second term captures another part of asset demand: because the expected return is predictable, the correlation between current equity return and the future portfolio returns is not zero, the consumers are willing to hold less equity to hedge against future bad changes in investment opportunities.

Case 2 When $\gamma \to 1$ and $\sigma \neq 1$[66], both the optimal portfolio rule and consumption choice are not myopic.

As discussed in Giovannini and Weil (1989), in the presence of RI, the portfolio rule under this restriction is myopic, that is, portfolio choice $\chi_t = \frac{1}{2\gamma} + \frac{1}{\gamma\omega_u^2} x_t$ and it does not depend on the hedging component, whereas the consumption rule is not myopic in the sense that the consumption wealth ratio is not constant and in fact the log ratio is a linear function of the unique state x_t,

$$\widetilde{c}_t = c_t - w_t = b_0 + b_1 x_t \tag{4.14}$$

If $V_t = (1-\beta)^{-\sigma/(1-\sigma)} (\exp(b_0 + b_1 x_t))^{1/(1-\sigma)}$ can be approximated around $b_1/(1-\sigma) \cdot x_t = 0$ by a linear quadratic function, as above we can introduce RI under these conditions. Specifically,

[66] Hence, it implies that $\rho = 0$.

with RI consumers form an estimated state to make optimal decisions, that is,

$$\widetilde{c}_t = c_t - w_t = b_0 + b_1 \widehat{x}_t, \tag{4.15}$$

$$\chi_t = \frac{1}{2\gamma} + \frac{1}{\gamma \omega_u^2} \widehat{x}_t. \tag{4.16}$$

Proposition 2. *In Case 2, both the ratio of consumption to wealth and equity holdings react slowly with respect to the innovations to the expected return in every period.*
Proof. *(see the proof in Case 1).* ■

Note that in this case, $\rho \to 0$ when $\gamma \to 1$.Hence, we can have the same pricing equation (4.12) and thus have the same optimal portfolio rule (4.13). In other words, in Case 2, the demand for the risky asset is also composed by two components. The first captures the part of asset demand induced by the current risk premium adjusted by both the fundamental risk ω_u^2 and the induced estimation risk σ_x^2 due to RI, whereas the second can capture the intertemporal hedging demand induced by the correlation between current equity return and the future portfolio returns.

5. Incorporating Labor Income Risk

So far I have assumed that labor income can be capitalized into marketable wealth. However, in reality, human wealth[67] is nontradable in the market because it is difficult to sell claims against future labor income. Consequently, consumers would adjust their financial asset holdings to take account of their implicit holdings of human wealth. It is known that riskless labor income is equivalent to an implicit holding of riskless assets that tilts asset allocations towards risky assets, whereas risky labor income that perfectly correlates with risky assets is equivalent to an implicit holding of risky assets and thus tilts the financial portfolio towards safe assets. Recent theoretical works showed that labor income risk can have significant effects on consumption and portfolio decisions. See Campbell and Viceira (2002) for a textbook treatment. Therefore, it is interesting to examine how labor income risk affects both risk premium and asset allocations in the RI economy. In this section, I adopt the same preference specification as section 2, while modelling the flow budget constraint for consumers differently:

$$W_{t+1} = R^p_{t+1}(W_t + Y_t - C_t) \tag{5.1}$$

[67] 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^p_{t+1} is the one-period return on savings given by equation (2.3), and Y_{t+1} is labor income. Following Viceira (2001) and Campbell and Viceira (2002), I specify the process for labor income as follows

$$Y_{t+1} = Y_t \exp(v_{t+1} + g) \tag{5.2}$$

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 of both permanent and transitory shocks. Here we ignore transitory shocks to labor income to make the notations simple. Furthermore, I also assume that the innovation to labor income may be contemporaneously correlated with the innovation to equity return:

$$\mathbf{cov}_t(u_{t+1}, v_{t+1}) = \omega_{uv}. \tag{5.3}$$

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

As in section 2, to derive the optimal consumption and portfolio choice rules, I first log-linearize the flow budget constraint (5.1) around $c - y = \mathbf{E}[c_t - y_t]$ and $w - y = \mathbf{E}[w_t - y_t]$ as follows

$$w_{t+1} - y_{t+1} \simeq \eta + \eta_w(w_t - y_t) - \eta_c(c_t - y_t) - \Delta y_{t+1} + r^p_{t+1} \tag{5.4}$$

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

$$c_t = b_0 + b_1 m_t \tag{5.5}$$

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

$$b_1 = \frac{\eta_w - 1}{\eta_c} \text{ and } b_0 = \frac{1}{\eta_c}\{\eta - g - \frac{\sigma}{b_1}\log\beta + (1 - \frac{\sigma}{b_1})\mathbf{E}[r^p_{t+1}] - \frac{1}{2}\Xi\} \tag{5.6}$$

where Ξ is the precautionary savings term (See Appendix D for a detailed derivation).

As shown in the Appendix D, for the CCAPM model with labor income, Ξ and then b_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 because the certainty equivalence principle does not hold generally. However, we may impose some condition to eliminate precautionary savings and then fit it into the RI framework as we characterized for the standard LQ case. It is clear from the expression for Ξ :

$$\Xi = (1 - \frac{1}{\sigma}b_1)^2 \mathbf{var}[r^p_{t+1}] + (\frac{1}{\sigma}b_1)^2 \lambda \mathbf{var}[v_{t+1}] - 2\frac{1}{\sigma}b_1\lambda(1 - \frac{1}{\sigma}b_1)\mathbf{cov}[r^p_{t+1}, v_{t+1}] \tag{5.7}$$

where $\lambda = \frac{1-\eta_w+\eta_c}{\eta_w-1}$. When the elasticity of intertemporal substitution σ is set to $b_1 \in (0,1)$, $\Xi = \frac{1-\eta_w+\eta_c}{\eta_w-1}\mathbf{var}[v_{t+1}] = \frac{1}{\exp(c-y)-1}\mathbf{var}[v_{t+1}] > 0$. Hence, when $w - y$ is large enough[68] or $\mathbf{var}[v_{t+1}]$ is small enough, the precautionary savings term Ξ should be close to 0. As documented in Viceira (2001) and Campbell and Viceira (2002), $\mathbf{var}[v_{t+1}]$ is around 10% per year, but here we may set it to a very low value because we are only interested in accessing how RI affect risk premium and asset allocation in the presence of labor income risk rather than calibrating the model to the real economy.

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

$$c_t = b_0 + b_1\widehat{m}_t \tag{5.8}$$

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

$$\widehat{m}_{t+1} = (1-\theta)\widehat{m}_t + \theta(m_{t+1} + \xi_{t+1}) \tag{5.9}$$

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

$$\Delta c_{t+1} = b_1\{[\theta\zeta_{t+1} + \theta\eta_w \frac{(1-\theta)\zeta_t}{1-(1-\theta)\eta_w \cdot L}] + [\theta\xi_{t+1} - \frac{\theta\xi_t}{1-(1-\theta)\eta_w \cdot L}] + \Omega\}$$

where $\zeta_{t+1} = \chi r^e_{t+1} + \lambda v_{t+1}$, L is the lag operator and Ω is constant term. Furthermore, as in section 2, I 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 have the following proposition

Proposition 1. *Aggregate consumption growth in the CCAPM model with RI can be written as*[69]

$$\Delta c_{t+1} = b_1[\theta\zeta_{t+1} + \theta\eta_w \frac{(1-\theta)\zeta_t}{1-(1-\theta)\eta_w \cdot L} + \Omega], \tag{5.10}$$

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

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

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

asset returns is

$$\mathbf{cov}(\Delta c_{t+1}, r^e_{t+1}) = b_1\theta\chi\omega_u^2 + b_1\theta\lambda\omega_{uv}, \tag{5.11}$$

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

$$\mathbf{cov}[\lim_{S\to\infty}(c_{t+1+S} - c_t), r^e_{t+1}] = b_1\frac{\theta}{1-(1-\theta)\eta_w}\chi\omega_u^2 + b_1\theta\lambda\omega_{uv}, \tag{5.12}$$

and 3) the optimal asset allocation is

$$\chi = \vartheta_2[\frac{\pi}{\gamma b_1\omega_u^2} - \frac{\theta\lambda\omega_{uv}}{\gamma\omega_u^2}] \tag{5.13}$$

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

Proof. see Appendix E and F. ■

Here $\vartheta_2 = \frac{1-(1-\theta)\eta_w}{\theta} < 1$ is inversely proportional to the ultimate consumption risk of the 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, whereas 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[70].

In Viceira (2001) and Campbell and Viceira (2003), they analyzed the optimal portfolio choices of long-horizon investors with undiversifiable labor income risk and showed that a 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 a poor hedge against unexpected declines in labor income. Because $\vartheta_2 < 1$ and $\theta < 1$, we can see from the above proposition 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 CCAPM model. Hence, the hedging demand is reduced.

As in section 2, alternatively, we can also write the pricing equation using long-term consumption

[70] 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.

risk as follows:

$$\mathbf{E}_t[r^e_{t+1}] - r^f + \frac{1}{2}\omega_u^2 = \gamma \sum_{s=0}^{S} \mathbf{cov}_t[\Delta c_{t+1+s}, r^e_{t+1}] \tag{5.14}$$
$$= \gamma[b_1 \frac{\theta}{1-(1-\theta)\eta_w} \chi \omega_u^2 + b_1 \theta \lambda \omega_{uv}].$$

where $\frac{\theta}{1-(1-\theta)\eta_w} > 1$ measures the ultimate consumption risk. It is clear that two components affect the equity premium in this case[71]. The first component represents the premium when labor income risk is idiosyncratic, that is, uncorrelated with the risky asset, while the second component measures an income hedging demand. In the full-information case, a positive covariance between the risky asset return and labor income, $\omega_{uv} > 0$, requires a higher equity premium because the risky asset offers a bad hedge against negative income shocks, and this dampens consumers' enthusiasm for stocks. In this RI model, RI, measured by θ appeared in the second component, reduces the impact of the correlation of labor income and asset returns on the equity premium. Further, if this covariance is negative, that is, the risky asset offers a good hedge against bad income shocks, RI also reduces its negative impact on the premium.

6. Equilibrium Asset Pricing under RI

In the previous section, I solved for the optimal consumption and portfolio rule of consumers facing exogenous asset return processes, Lucas (1978), Mehra and Prescott (1985), and others adopted another reversed procedure to derive the equilibrium asset returns by specifying an exogenous aggregate consumption process and then modeling the stock market as paying a dividend proportional to aggregate consumption due to leverage. In this section, I will explore the equilibrium implications of the decision rules derived in the previous section. Specifically, I use the explicit partial equilibrium results and an log-linear accounting identity to examine how RI affects the price-dividend ratio, the equity premium, and the equity volatility in equilibrium.

Following Campbell and Shiller (1988), Campbell (2003), and others, here I use a loglinear approximation to link stock prices, dividends, and returns in an accounting framework. This loglinear approximation starts with the definition of the log return on the risky asset e,

$$r^e_{t+1} = \log(P^e_{t+1} + D^e_{t+1}) - \log(P^e_t)$$

Note that log return is a nonlinear function of log prices p^e_{t+1} and $p^e_t$and log dividends d^e_{t+1}, and

[71] Of course, here I also assume that the share of the risky asset is given when examining the equity premium issue.

can be approximated around the mean dividend-price ratio, $\mathbf{E}[d_t^e - p_t^e]$, by using first order Taylor expansion. Campbell (2003) derived the resulting approximation as follows:

$$r_{t+1}^e = k + \rho p_{t+1}^e + (1-\rho) d_{t+1}^e - p_t^e \tag{6.1}$$

where $\rho = 1/(1 + \exp(\mathbf{E}[d_t^e - p_t^e]))$[72] and $k = -\log \rho - (1-\rho)\log(1/\rho - 1)$.

This equation is a linear difference equation for the log stock price[73]:

$$p_t^e - d_t^e = \frac{k}{1-\rho} + \mathbf{E}_t \sum_{j=0}^{\infty} \rho^j [\Delta d_{t+1+j}^e - r_{t+1+j}^e] \tag{6.2}$$

where we need to impose the terminal condition $\lim_{j\to\infty} \rho^j p_{t+j}^e = 0$.

Combining equation (6.1) with (6.2) yields

$$r_{t+1}^e - \mathbf{E}_t[r_{t+1}^e] = (\mathbf{E}_{t+1} - \mathbf{E}_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j}^e - (\mathbf{E}_{t+1} - \mathbf{E}_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j}^e, \tag{6.3}$$

which says that the unexpected stock returns must be associated with either the revisions in expectations in future dividend growth (with a positive sign) or news about future returns (with negative sign)[74].

Following Lucas (1978), Mehra and Prescott (1985), and others, here I also assume that the aggregate stock market, denoted by e for equity, is a reasonable proxy for the portfolio of total wealth and thus can be priced as if it pays aggregate consumption as its dividend. Further, as in Abel (1999) and Campbell (2003), dividend on equity equals to aggregate consumption raised to a power λ[75]. In logs, we have

$$d_t^e = \lambda c_t \tag{6.4}$$

Note that in the Lucas-type economy, bonds do not exist explicitly. However, the shadow prices of any security that is in zero net supply can be computed by using the intertemporal first order condition of consumers. We can thus use this approach to price the risk free rate. In other words,

[72] In the postwar quarterly data the average price-dividend ratio has been 28.3 on annual basis, which means that ρ should be about 0.966 in annual data.

[73] Campbell (2003) (section 4) discussed how to derive this equation in detail.

[74] For example, a real bond has no dividend uncertainty, so unexpected returns cannot occur without changes in future expected returns. However, dividend news is important for equity.

[75] They interpreted λ as a measure of leverage.

(2.47) is still valid in this economy. Substituting equation (2.47) into (2.29) gives[76]

$$\mathbf{E}_t[r^e_{t+1}] = \mu^e + \frac{1}{\sigma}\mathbf{E}_t[\Delta c_{t+1}] \tag{6.5}$$

where μ^e is an asset-specific constant term[77]:

$$\mu^e \simeq \gamma\vartheta^{-1}\chi\omega_u^2 - \frac{1}{2}\omega_u^2. \tag{6.6}$$

Note that since ϑ^{-1} is used to measure the ultimate consumption risk and is an increasing function of the degree of RI, μ^e is also increasing with inattention.

Substituting (6.4) and (6.5) into (6.2) and (6.3), we obtain

$$p^e_t - d^e_t = \frac{k-\mu^e}{1-\rho} + (\lambda - \frac{1}{\sigma})\mathbf{E}_t\sum_{j=0}^{\infty}\rho^j\Delta c_{t+1+j}, \tag{6.7}$$

$$\text{and } r^e_{t+1} - \mathbf{E}_t[r^e_{t+1}] = \lambda(\Delta c_{t+1} - \mathbf{E}_t[\Delta c_{t+1}]) + (\lambda - \frac{1}{\sigma})\mathbf{E}_t\sum_{j=1}^{\infty}\rho^j\Delta c_{t+1+j} \tag{6.8}$$

Hence, (6.7) implies that the higher the inattention, the lower the price-dividend ratio. We can further simplify these equations by specifying the aggregate consumption process as follows:

$$\Delta c_{t+1} = \rho_c\Delta c_t + (1-\rho_c)g + \epsilon_{c,t+1} \tag{6.9}$$

where $\epsilon_{c,t+1}$ is i.i.d. white noise with mean 0 and standard deviation ω_c. Substituting this process into (6.7) and (6.8) yields

$$\begin{aligned} p^e_t - d^e_t &= \frac{k-\mu^e}{1-\rho} + (\lambda - \frac{1}{\sigma})[\frac{g}{1-\rho} + \frac{\Delta c_t - g}{1-\rho\rho_c}] \\ \text{and } r^e_{t+1} - \mathbf{E}_t[r^e_{t+1}] &= [\lambda + (\lambda - \frac{1}{\sigma})\frac{\rho\rho_c}{1-\rho\rho_c}]\epsilon_{c,t+1} \end{aligned}$$

Now we can give the following definition about a RI equilibrium

Definition 1. *A RI equilibrium consists of a consumption rule c_t, a portfolio rule χ, the price-dividend ratio $p^e_t - d^e_t$, and the risky return r^e_{t+1}, such that simultaneously*

---

[76] Note that with CRRA utility, the CRRA $\gamma = 1/\sigma$, where σ is the EIS.

[77] Here we use the following approximation:

$$-\log\beta - \frac{1}{2}\gamma(\gamma+1)\mathbf{var}[\Delta c_{t+1}] \simeq 0.$$

1) Given exogenous asset returns, each individual solves the optimization problem (2.13) *under the flow budget constraint and the information-processing constraint,*

2) Markets clear in each period, that is, $c_t = d_t^e/\lambda$ *and* $\chi = 1$.

Hence, given the explicit solutions for consumption and portfolio rule derived in the partial equilibrium CCAPM, we can derive the equilibrium price-dividend ratio, equity premium, and the volatility of asset returns. The following proposition gives the main results.

Proposition 2. *In equilibrium, the mean price-dividend ratio is given by*

$$\mathbf{E}[p_t^e - d_t^e] = \frac{k - \mu^e}{1-\rho} + (\lambda - \frac{1}{\sigma})\frac{g}{1-\rho}\ . \tag{6.10}$$

The innovations of asset returns satisfy

$$u_{t+1} = [\lambda + (\lambda - \frac{1}{\sigma})\frac{\rho\rho_c}{1-\rho\rho_c}]\epsilon_{c,t+1} \tag{6.11}$$

Further, based on the partial equilibrium results, we have

$$\epsilon_{c,t+1} = \theta u_{t+1}, \tag{6.12}$$

where $\rho_c = (1-\theta)/\phi$[78]*. Hence,* (6.11) *and* (6.12) *imply that*

$$[\lambda + (\lambda - \frac{1}{\sigma})\frac{\rho\rho_c}{1-\rho\rho_c}]\theta = 1. \tag{6.13}$$

Finally, the expected excess return can then be written as

$$\mathbf{E}_t[r_{t+1}^e] - r^f + \frac{1}{2}\omega_u^2 = \gamma\frac{1}{\theta}\frac{1}{1-\rho_c}\omega_c^2 \tag{6.14}$$

Note that to derive (6.14), we use the facts that

$$\mathbf{E}_t[r_{t+1}^e] - r^f + \frac{1}{2}\omega_u^2 = \gamma\mathbf{cov}_t[\lim_{S\to\infty}(c_{t+1+S} - c_t), \frac{\epsilon_{c,t+1}}{\theta}] \text{ and } \Delta c_{t+1+S} = \frac{\epsilon_{c,t+1+S}}{1-\rho_c\cdot L} + g.$$

From equation (6.14), it is obvious that RI can help resolve the equity premium puzzle in this equilibrium framework. Given the stochastic properties of aggregate consumption, the persistence and volatility of aggregate consumption (ρ_c and ω_c^2), RI increases the equity premium by a factor

[78] *Note that since* γ *is close to* 1, ϕ *is* β*.*

$\frac{1}{\theta}$. This means that the higher the degree of RI, the higher the equity premium in equilibrium given the observed consumption process.

Furthermore, since $\rho = 1/(1+\exp(\mathbf{E}[d_t^e - p_t^e]))$ and $k = -\log\rho - (1-\rho)\log(1/\rho - 1)$, (6.10) implicitly determines the mean price-dividend ratio, and rearranging it yields an explicit expression for $\mathbf{E}[p_t^e - d_t^e]$:

$$\mathbf{E}[p_t^e - d_t^e] = -\mu^e - (\lambda - \frac{1}{\sigma})g. \tag{6.15}$$

Hence, it is obvious that

$$\frac{\partial \mathbf{E}[p_t^e - d_t^e]}{\partial \kappa} > 0 \tag{6.16}$$

since $\frac{\partial \mu^e}{\partial \kappa} < 0$.

It is also worth noting that when $\lambda = 1$, because γ is set to be close to 1 (that is, $\sigma = 1/\gamma$ is also close to 1) the equilibrium exists only if $\theta = 1$, that is, $\kappa = \infty$. In other words, with CRRA utility and $\lambda = 1$, there is no RI equilibrium. When $\lambda - \frac{1}{\sigma} > 0$[79], there exists a RI equilibrium when equation (6.13) is satisfied.

Further, (6.12) implies that the higher the degree of RI, the larger the volatility of asset returns given consumption volatility. RI can therefore be an alternative explanation for the equity volatility puzzle[80], that is, why is the volatility of real stock returns so high in relation to the volatility of the short-term real interest rate?

In fact, we can also understand the equilibrium by the following way: 1) given asset returns, one can solve for individuals' optimal consumption and portfolio rules; 2) aggregating over all consumers generates the aggregate consumption process, which is a AR(1) process; 3) in equilibrium, aggregate consumption is proportional to aggregate dividend because of leverage (when the leverage ratio is equal to 1, they are the same.); 4) Using the log-linear accounting identity that links returns, prices, and dividends, if taking aggregate consumption process as given, one can determine the price-dividend ratio, the equity premium, and the equity volatility; 5) in equilibrium, when some conditions are satisfied, the resulting returns should be just the initially given returns. This procedure is similar to the computational algorithm used to compute a recursive competitive equilibrium in the heterogenous-agent economy with idiosyncratic risk, e.g., the Bewley-Aiyagari type models, where the equilibrium interest rate and wage rate is also determined by solving individuals' optimization problem, aggregating over all individuals, and then iterating until convergence.

[79] As pointed out in Campbell (2003), an obvious way to generate volatile equity returns is to set a large value for λ, that is, a volatile dividend. However, this result depends on if $\lambda - \gamma > 0$ or not.

[80] For a detailed discussion of this puzzle, see Campbell (2003).

7. Review of Related Empirical Evidence

Some existing survey evidence supports that 1) consumers 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 changes in consumption. For example, Dynan and Maki (2001) 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 whereas the US stock markets rose over 15% per year during this sample period[81]. In the same paper, they also reported that for stockholders with over $10,000 in securities, a 1% increase in the value of security holdings 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 the 10th month to the 18th month, another quarter of which occurs form the 19th month to the 27th month, and the rest occurs from the 28th month to the 36th month. This evidence may be largely captured by our RI model because equation (2.39) implies that

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

where $\lim_{S\to\infty}[c_{t+1+S} - c_t] = \frac{\theta\chi}{1-(1-\theta)/\phi} u_{t+1}$. Consider a numerical example (the time unit here is 3 quarters) in which $\beta = 0.99$, $\phi \simeq \beta$, 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 similar results as those estimated from the US data. Furthermore, using the estimation results from Dynan and Maki (2001), 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 $> \$1,000$, whereas in the right figure, we define SH as households with securities $> \$10,000$. When we plot the profile generated from our model, we calibrate the observation weight θ such that the initial jump of consumption to the shock of 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 dynamic responses reflected in the data. The left figure below shows that the RI model with $\kappa = 0.14$ can fit the empirical results quite well: the responses of consumption to the innovations is muted initially and then increases

[81] Kennickell, et. al (2000) and Starr-McCluer (2001) also reported similar results based on alternative survery sources.

gradually over time. The right figure also shows a similar pattern of the responses, though the fit is not as good as the left one.

[Insert Figure 3 about here]

As in Gabaix and Laibson (2001), we are also interested in whether the values of $\mathbf{cov}[c_{t+1+S} - c_t, r^e_{t+1}]$ 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[82] in the following figure. The figure shows 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. This figure also shows the covariance profiles generated by the RI model with different values of channel capacity. It is obvious that the RI model can capture this apparent empirical feature successfully: the covariance slowly rises 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 of asset returns 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[83] and ignore the accumulating financial wealth until their retirement age 65.

[Insert Figure 4 about here]

Parker (2001) used data from the CEX of the Bureau of Labor Statistics and calculated 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) plotted 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.

[82] 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).

[83] This mental account can be regarded as "asset account" and the MPC from this account is less than the MPC from the "current income" account.

8. Conclusions

Ordinary people do not have unlimited channel capacity and information-processing abilities, instead they only have finite Shannon capacity when processing economic information. Rational inattention, first introduced by Sims (2003), has recently been applied in some economic models as a kind of information-processing constraints. This paper takes such information-processing constraint into account in the CCAPM framework and explores its effects on optimal consumption and portfolio choice as well as the equity premium.

The first contribution of this paper is to solve a simple CCAPM model with RI in closed form approximately. The closed form solutions are useful in aggregating over all agents and calibrating the model to the data. Second, given the solutions, it is shown that introducing RI can generate excess smoothness of consumption growth, the low contemporaneous correlation between consumption growth and asset returns, and implied high equity premium simultaneously. The intuition is that agents with RI cannot digest all relevant information instantly and completely, so they react gradually and with delay to the innovations. Hence, we need to use the ultimate consumption risk rather than the contemporaneous one to measure the true risk of the asset. The larger the degree of RI, the larger the ultimate consumption risk is. Consequently, RI reduces the demand for the risky asset because they face more risk in long run.

Next, it is shown that incorporating RI into this expected utility maximization framework can also disentangle the CRRA from the EIS by increasing the *effective* CRRA and reducing the *effective* EIS. Hence, if individuals are highly inattentive, the model can generate high equity premium by just increasing the *effective* CRRA by increasing the degree of inattention and keeping the *true* CRRA unchanged. RI can then be an alternative explanation for the equity premium puzzle (by increasing the effective CRRA) and the risk free rate puzzle (by keeping the true CRRA and then the true IES unchanged). In other words, RI can reconcile low estimates of risk aversion obtained from experimental evidence or introspection, with high estimated values of risk aversion based on asset pricing data. I then discussed the different implications of RI, recursive preference, and the preference for robustness for optimal consumption and asset holdings and asset returns.

Further, I consider incorporating RI into the recursive utility framework and examine the interactions of RI with the CRRA and the EIS in determining optimal consumption and portfolio choice as well as asset returns. I also consider the dynamic investment problem when consumers face stochastic investment opportunities. It is shown that RI slows the speed of adjusting the portfolio share and the consumption-wealth ratio with respect to the innovations to the expected returns. Also, it affects the hedging demand for the risky asset.

Finally, I propose a general equilibrium Lucas-type asset pricing framework to examine the implications of RI on the equity premium, the mean ratio of price to dividend, and the equity volatility puzzle in general equilibrium.

A number of extensions of this framework considered here seem promising. First, we could consider multiple risky assets in the model and investigate whether RI can resolve the "asset allocation puzzle" of Canner, Mankiw, and Weil (1997). They find that the ratio of risky bonds to equities in the optimal portfolio increases with the CRRA, which is consistent with conventional portfolio advice but inconsistent with static mean-variance analysis. Another related question is whether the investors with a high degree of RI prefer to hold long-term bonds. Second, how does the horizon affect the asset demand in the RI economy? Third, we can examine if time-varying channel capacity can explain the observed undereaction and overreaction phenomena in the stock market in this framework.

9. Appendices

9.1. Appendix A: Deriving the Value Function in the CCAPM Model with RI

Before deriving the value function in the RI economy, it is helpful to first derive the value function in the corresponding full-information case. In the standard full-information PIH model, we have the consumption rule: $c_t = b_0 + w_t$. As usual, we guess that the value function has the following form[84]

$$V(w_t) \simeq A_0 + A_1 w_t + A_2 w_t^2 \tag{9.1}$$

where A_0, A_1, and A_2 are undetermined coefficients. Following the standard procedure, substituting the optimal consumption function into the Bellman equation, $V(w_t) = \max_{c_t} \mathbf{E}_t[u(c_t) + \beta V(w_{t+1})]$, we can then pin down the undetermined coefficients:

$$A_1 = \frac{1}{1-\beta}(1-b_0), A_2 = -\frac{1}{2}\frac{1}{1-\beta}, A_0 = \frac{1}{1-\beta}(b_0 - \frac{1}{2}b_0^2 + \beta A_2 \omega_u^2)$$

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

$$\widehat{V}(\widehat{w}_t) = \max_{c_t, \mathcal{D}_t} \mathbf{E}_t[u(c_t) + \beta \widehat{V}(\widehat{w}_{t+1})] \tag{9.2}$$

where the expectation is formed under the assumption that the current and future consumption are chosen under information processing constraints. Similarly, we guess that $\widehat{V}(\widehat{w}_t) = B_0 + B_1 \widehat{w}_t +$

[84]Note that the exact form of the value function is $V(W_t) = \exp(-\gamma b_0)\frac{W_t^{1-\gamma}}{1-\gamma}$, so when $(\gamma - 1)w_t$ is close to 0, this value function can be approximated by this LQ form.

$B_2\widehat{w}_t^2$. Substituting the optimal consumption rule, the Kalman equation, and the guessed value function into the above Bellman equation, we have the following equality

$$B_0 + B_1\widehat{w}_t + B_2\widehat{w}_t^2 = (b_0 + b_1\widehat{w}_t) - \frac{1}{2}(b_0 + b_1\widehat{w}_t)^2 + \beta(B_0 + B_1\mathbf{E}_t[\widehat{w}_{t+1}] + B_2\mathbf{E}_t[\widehat{w}_{t+1}^2]) \tag{9.3}$$

Using the following facts

$$\begin{aligned}
\mathbf{E}_t[\widehat{w}_{t+1}] &= \widehat{w}_t, \\
\mathbf{var}_t[\widehat{w}_{t+1}] &= \theta^2[(1/\phi)^2\mathbf{var}_t[w_t] + \chi^2\omega_u^2 + \mathbf{var}[\xi_t]], \\
\mathbf{E}_t[\widehat{w}_{t+1}^2] &= \mathbf{var}_t[\widehat{w}_{t+1}] + \widehat{w}_t^2, \\
&= ((1/\phi)^2 - 1)\mathbf{var}_t[w_t] + \chi^2\omega_u^2 + \widehat{w}_t^2,
\end{aligned}$$

we can then easily pin down the coefficients B_0, B_1, and B_2 as follows:

$$B_1 = \frac{1}{1-\beta}(1-b_0), B_2 = -\frac{1}{2}\frac{1}{1-\beta}, B_0 = \frac{1}{1-\beta}[b_0 - \frac{1}{2}b_0^2 + \beta B_2(((1/\phi)^2 - 1)\mathbf{var}_t[w_t] + \chi^2\omega_u^2)]$$

where $\mathbf{var}_t[w_t] = \Sigma_t$. Note that $A_1 = B_1$ and $A_2 = B_2$.

9.2. Appendix B: Deriving the Gaussianity of Optimal Signal

Given the LQG framework, we define the loss function at t due to imperfect information as $\Delta V = V(w_t) - \widehat{V}(\widehat{w}_t)$. Hence, the expected loss function is

$$\begin{aligned}
\mathbf{E}_t[\Delta V] &= \mathbf{E}_t[V(w_t) - \widehat{V}(\widehat{w}_t)] \\
&= \mathbf{E}_t[(A_0 - B_0) + A_1(w_t - \widehat{w}_t) + A_2(w_t^2 - \widehat{w}_t^2)] \\
&= (A_0 - B_0) + A_2\mathbf{E}_t[(w_t - \widehat{w}_t)^2] \\
&= -RA_2\mathbf{E}_t[(w_t - \widehat{w}_t)^2]
\end{aligned}$$

where we use the facts that $V(w_t) = A_0 + A_1 w_t + A_2 w_t^2, \widehat{V}(\widehat{w}_t) = B_0 + B_1\widehat{w}_t + B_2\widehat{w}_t^2, A_1 = B_1 > 0, A_2 = B_2 < 0$, $\widehat{w}_t = \mathbf{E}_t[w_t]$, and $\mathbf{cov}_t[w_t - \widehat{w}_t, \widehat{w}_t] = 0$.

Since $A_2 < 0$, choosing the signal to minimize the above expected loss function above is equivalent to choosing the signal to minimize $\mathbf{E}_t[(w_t - \widehat{w}_t)^2]$. Hence, minimizing this loss function under the IPC can be characterized by the following optimization problem[85]

$$\min_{f(w_t,\widehat{w}_t)} \mathbf{E}_t[(w_t-\widehat{w}_t)^2]$$

subject to:

$$-\mathcal{H}(w_t,\widehat{w}_t)+\mathcal{H}(\widehat{w}_t)+\mathcal{H}(w_t)\leq\kappa$$

where $\mathcal{H}$ is entropy and κ is channel capacity on the mutual information between w_t and $\widehat{w}_t$.

We could reformulate the above optimization problem as follows

$$\min_{f(w_t,\widehat{w}_t)} \int\int (w_t-\widehat{w}_t)^2 f(w_t,\widehat{w}_t)dw_t d\widehat{w}_t \tag{9.4}$$

$$\text{subject to:} \tag{9.5}$$

$$\int\int \log(f(w_t,\widehat{w}_t))f(w_t,\widehat{w}_t)dw_t d\widehat{w}_t - \int \log g_0(w_t)g_0(w_t)dw_t \tag{9.6}$$
$$-\int[\log(\int f(w_t,\widehat{w}_t)dw_t)\cdot\int f(w_t,\widehat{w}_t)dw_t]d\widehat{w}_t\leq\kappa,$$

$$\int f(w_t,\widehat{w}_t)d\widehat{w}_t = g_0(w_t) \tag{9.7}$$

where $g_0(w_t)$ is the prior pdf of the true state, which is assumed to be Gaussian, $f(w_t,\widehat{w}_t)$ is the joint pdf of the true state and the observed signal.

The Lagrangian multipliers for the above two constraints are λ and $\mu(w_t)$, respectively. The Lagrangian function is then

$$L=\int\int(w_t-\widehat{w}_t)^2 f(w_t,\widehat{w}_t)dw_t d\widehat{w}_t \tag{9.8}$$
$$+\lambda\{\int\log g_0(w_t)g_0(w_t)dw_t+\int[\log(\int f(w_t,\widehat{w}_t)dw_t)\cdot\int f(w_t,\widehat{w}_t)dw_t]d\widehat{w}_t$$
$$-\int\int\log(f(w_t,\widehat{w}_t))f(w_t,\widehat{w}_t)dw_t d\widehat{w}_t\}-\mu(w_t)[\int f(w_t,\widehat{w}_t)d\widehat{w}_t-g_0(w_t)]$$

Since there are no derivatives of $f(w_t,\widehat{w}_t)$ involved here, we can solve this problem point by point. Differentiating the Lagrangian function with respect to the pdf $f(w_t,\widehat{w}_t)$ gives us the first order condition

$$(w_t-\widehat{w}_t)^2-\lambda[\log f(w_t,\widehat{w}_t)-\log(\int f(w_t,\widehat{w}_t)dw_t)]-\mu(w_t)=0 \tag{9.9}$$

[85] The following derivations are largely based on Sims's notes on "Optimality of Gaussian Observation Error" that is available from http://sims.princeton.edu/yftp/AdvMacro03/GaussianOptNotes.pdf.

We define the pdf of the signal observed through the limited channel as

$$q(\widehat{w}_t) = \int f(w_t, \widehat{w}_t) dw_t,$$

then the above FOC implies that

$$\begin{aligned} f(w_t|\widehat{w}_t) &= \frac{f(w_t, \widehat{w}_t)}{q(\widehat{w}_t)} \\ &= \exp\{\frac{1}{\lambda}[\mu(w_t) - (w_t - \widehat{w}_t)^2]\} = A(w_t)\exp[B(w_t - \widehat{w}_t)^2] \end{aligned}$$

where $A(w_t) = \exp[\frac{1}{\lambda}\mu(w_t)]$ and $B = -\frac{1}{\lambda}$. By setting the simple function A and B, it is clear that the right-hand side is a Gaussian pdf. Note that given $w_0 \sim N(\widehat{w}_0, \Sigma_0)$, $\widehat{w}_t$ summarizes all current and past observed signals $\{w_s^*, s = 1, \cdots, t\}$.

Hence, if the signals $\{w_s^*, s = 1, \cdots, t\}$ is Gaussian distribution, then the joint distribution of w_t and $\widehat{w}_t$ is also Gaussian.

9.3. Appendix C Deriving the Expression of Consumption Growth

Adding RI in the CCAPM model yields a modified consumption rule (2.22), $c_t = b_0 + \widehat{w}_t$, where $\widehat{w}_t$ can be characterized by the Kalman filtering equation (2.23). Combining these two equations with the log-linearized budget constraint (2.5) yields

$$\Delta\widehat{w}_{t+1} = (\theta/\phi)(w_t - \widehat{w}_t) + (\theta\chi u_{t+1} + \theta\xi_{t+1}) + \theta[\overline{r}^p + \psi + (1 - 1/\phi)b_0]$$

where $w_t - \widehat{w}_t = \frac{(1-\theta)\chi u_t - \theta\xi_t}{1-((1-\theta)/\phi)\cdot L} + \Omega_1$ because[86]

$$w_{t+1} - \widehat{w}_{t+1} = ((1-\theta)/\phi)(w_t - \widehat{w}_t) + [(1-\theta)\chi u_{t+1} - \theta\xi_{t+1}] + (1-\theta)[\overline{r}^p + \psi + (1 - 1/\phi)b_0]$$

where L is the lag operator, $\overline{r}^p = (1-\chi)r^f + \chi\mu$, and $\Omega_1 = \frac{(1-\theta)[\overline{r}^p + \psi + (1-1/\phi)b_0]}{1-(1-\theta)/\phi}$ is a constant term.

Hence, consumption growth can now be written as

$$\Delta c_{t+1} = \Delta\widehat{w}_{t+1} = [\theta\chi u_{t+1} + \theta\chi\frac{((1-\theta)/\phi)u_t}{1-((1-\theta)/\phi)\cdot L}] + [\theta\xi_{t+1} - \theta\frac{(\theta/\phi)\xi_t}{1-((1-\theta)/\phi)\cdot L}] + \Omega_2 \quad (9.10)$$

[86] It is easy to show when γ is close to 1, ϕ is close to β and then Ω_1 is close to 0.

where Ω_2 is the constant term

$$\Omega_2 = \theta[\overline{r}^p + \psi + (1 - 1/\phi)b_0] + (\theta/\phi)\frac{(1-\theta)[\overline{r}^p + \psi + (1 - 1/\phi)b_0]}{1 - (1-\theta)/\phi} \tag{9.11}$$

Note that the endogenous noise terms in the second bracket of the above expression are totally idiosyncratic since they are generated from individual channel, they would be cancelled out when aggregating over all consumers.

9.4. Appendix D Deriving the Consumption Rule in the Presence of Labor Income Risk

We first divide the flow budget constraint (5.1) by Y_{t+1} and log-linearize it around steady state $c - y = \mathbf{E}[c_t - y_t]$ and $w - y = \mathbf{E}[w_t - y_t]$ as follows

$$w_{t+1} - y_{t+1} \simeq \eta + \eta_w(w_t - y_t) - \eta_c(c_t - y_t) - \Delta y_{t+1} + r^p_{t+1} \tag{9.12}$$

where

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

Second, to reduce this multivariate state case to the univariate state case, we need to define a new state variable that is a linear combination of w_t and y_t. Following the same procedure in Luo (2005), we rewrite the log-linearized budget constraint as follows:

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

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

$$m_{t+1} = \eta + \eta_w m_t - \eta_c c_t + r^p_{t+1} - g + \lambda v_{t+1} \tag{9.13}$$

Third, the log-linearized Euler equation is as follows

$$0 = \log\beta - \frac{1}{\sigma}\mathbf{E}_t[c_{t+1} - c_t] + \mathbf{E}_t[r^p_{t+1}] + \frac{1}{2}\mathbf{var}_t[r^p_{t+1} - \frac{1}{\sigma}(c_{t+1} - c_t)]. \tag{9.14}$$

Furthermore, we guess that the optimal log consumption rule take the following form $c_t = b_0 + b_1 m_t$. Hence, $c_{t+1} - c_t = b_1(m_{t+1} - m_t)$.

Combining it with the log-linearized budget constraint yields

$$\begin{aligned}\mathbf{E}_t[c_{t+1} - c_t] &= b_1 \mathbf{E}_t[m_{t+1} - m_t] \\ &= b_1[\eta + (\eta_w - 1 - \eta_c b_1)m_t - \eta_c b_0 + \mathbf{E}[r^p_{t+1}] - g] \end{aligned} \tag{9.15}$$

The log-linearized Euler equation implies that

$$\mathbf{E}_t[c_{t+1} - c_t] = \sigma[\log \beta + \mathbf{E}[r^p_{t+1}] + \frac{1}{2}\Xi] \tag{9.16}$$

$$\begin{aligned}\text{where} \qquad \Xi &= \mathbf{var}_t[r^p_{t+1} - \frac{1}{\sigma} b_1(m_{t+1} - m_t)] \\ &= (1 - \frac{1}{\sigma} b_1)^2 \mathbf{var}[r^p_{t+1}] + (\frac{1}{\sigma} b_1)^2 \lambda \mathbf{var}[v_{t+1}] - 2\frac{1}{\sigma} b_1 \lambda (1 - \frac{1}{\sigma} b_1) \mathbf{cov}_t[r^p_{t+1}, v_{t+1}]\end{aligned}$$

Equalizing the right-hand side of equation (9.15) and (9.16) and identifying coefficients, we obtain two key coefficients in the consumption function:

$$b_1 = \frac{\eta_w - 1}{\eta_c} \text{ and } b_0 = \frac{1}{\eta_c}\{\eta - g - \frac{\sigma}{b_1}\log \beta + (1 - \frac{\sigma}{b_1})\mathbf{E}[r^p_{t+1}] - \frac{1}{2}\Xi\}$$

9.5. Appendix E Deriving the Expression of Consumption Growth with Labor Income Risk

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

$$c_t = b_0 + b_1 \widehat{m}_t \tag{9.17}$$

and substituting it into (9.13) yields

$$m_{t+1} = \eta - g - \eta_c b_0 + (1 - \chi) r^f + \eta_w m_t - (\eta_w - 1)\widehat{m}_t + \chi r^e_{t+1} + \lambda v_{t+1}. \tag{9.18}$$

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

$$\widehat{m}_{t+1} = (1 - \theta)\widehat{m}_t + \theta(m_{t+1} + \xi_{t+1}) \tag{9.19}$$

Combining these three equations yields

$$\Delta\widehat{m}_{t+1} = \theta\eta_w(m_t - \widehat{m}_t) + \theta[\chi r^e_{t+1} + \lambda v_{t+1} + +\xi_{t+1}] + \Pi_1$$

where $\Pi_1 = \theta[\eta - g - \eta_c b_0 + (1-\chi)r^f]$, and

$$m_t - \widehat{m}_t = \frac{(1-\theta)(\chi r^e_t + \lambda v_t) - \theta\xi_t}{1 - ((1-\theta)\eta_w)\cdot L} + \frac{\Pi_2}{1-(1-\theta)\eta_w}$$

since $m_{t+1} - \widehat{m}_{t+1} = ((1-\theta)\eta_w)(m_t - \widehat{m}_t) + (1-\theta)(\chi r^e_{t+1} + \lambda v_{t+1}) - \theta\xi_{t+1} + \Pi_2$, where L is the lag operator and $\Pi_2 = (1-\theta)[\eta - g - \eta_c b_0 + (1-\chi)r^f]$. Hence, consumption growth can now be written as

$$\begin{aligned}\Delta c_{t+1} = b_1\{&[\theta(\chi r^e_{t+1} + \lambda v_{t+1}) + \theta\eta_w\frac{(1-\theta)(\chi r^e_t + \lambda v_t)}{1-((1-\theta)\eta_w)\cdot L}] \\ &+ [\theta\xi_{t+1} - \frac{\theta\xi_t}{1-((1-\theta)/\phi)\cdot L}] + \Pi\}\end{aligned}$$

where Π is the constant term.

9.6. Appendix F Proof of Proposition 7

Proof. (1) it is straightforward from the expression (5.10).

(2)

$$\begin{aligned}\mathbf{cov}[\lim_{S\to\infty}(c_{t+1+S} - c_t), r^e_{t+1}] &= \mathbf{cov}_t[\lim_{S\to\infty}(\sum_{s=0}^{S}\Delta c_{t+1+s}), r^e_{t+1}] \\ &= H_1\theta\lim_{S\to\infty}[1 + (1-\theta)\eta_w + \cdots + ((1-\theta)\eta_w)^S]\chi\omega_u^2 + b_1\theta\lambda\omega_{uv} \\ &= b_1\frac{\theta}{1-(1-\theta)\eta_w}\chi\omega_u^2 + b_1\theta\lambda\omega_{uv}.\end{aligned}$$

(3) Since

$$\begin{aligned}\pi &\approx \gamma\mathbf{cov}_t[\lim_{S\to\infty}(c_{t+1+S} - c_t), r^e_{t+1}] \\ &= \gamma[b_1\frac{\theta}{1-(1-\theta)\eta_w}\chi\omega_u^2 + b_1\theta\lambda\omega_{uv}],\end{aligned}$$

we can easily obtain

$$\chi = \vartheta_2[\frac{\pi}{\gamma b_1\omega_u^2} - \frac{\theta\lambda\omega_{uv}}{\gamma\omega_u^2}] \tag{9.20}$$

References

[1] Abel, Andrew (2001), "An Exploration of the Effects of Pessimism and Doubt on Asset Returns," *Journal of Economic Dynamics and Control* 26, 1075-1092.

[2] Adam, Klaus (2004), "Optimal Monetary Policy in the Presence of Imperfect Common Knowledge," European Central Bank, mimeo.

[3] Anderson, E., Lars Hansen, and Thomas Sargent (2000), "Robustness, Detection, and the Price of Risk," Stanford University, mimeo.

[4] Ash, Robert (1990), *Information Theory,* Dover Publications, Inc.

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

[6] Bansal, Ravi, and Amil Yaron (2004), "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles."*Journal of Finance* 59(4), 1481-1509.

[7] Barberis, Nicholas (2000), "Investing for the Long Run When Returns are Predictable," *Journal of Finance* 55, 225-264.

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

[9] Breeden, Douglas T. (1979), "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities," *Journal of Financial Economics* 7: 265–96.

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

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

[12] Campbell, John (2000), "Asset Pricing at the Millennium," *Journal of Finance* 55, 1515-67

[13] Campbell, John (2003), "Consumption-Based Asset Pricing," In George Constantinides, Milton Harris, and Rene Stultz, editors, *Handbook of the Economics of Finance* Volume 1B. Amsterdam: North-Holland, 803-887.

[14] Campbell, John and Robert Shiller (1988), "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors, " *Review of Financial Studies* 1: 195-227.

[15] Campbell, John and John H. Cochrane (1999), "By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205-251.

[16] Campbell, John and Luis Viceira (1999), "Consumption and Portfolio Decisions When Expected Returns are Time Varying," *Quarterly Journal of Economics* 114, 433-495.

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

[18] Canner, Niko, N. Gregory Mankiw and David N. Weil (1997), "An Asset Allocation Puzzle," *American Economic Review*, 87 (1), 181-191.

[19] Cochrane, John (2001), *Asset Pricing.* Princeton University Press.

[20] Cochrane, John (2005), "Financial Markets and the Real Economy," University of Chicargo GSB, mimeo.

[21] Cochrane, John and Lars Peter Hansen (1992), "Asset Pricing Lessons for Macroeconomics," in 1992 *NBER Macroeconomics Annual*, Olivier Blanchard and Stanley Fischer (eds.), 115-1165.

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

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

[24] Dynan, Karen (2000) "Habit Formation in Consumer Preferences: Evidence from Panel Data," *American Economic Review* 90, no. 3 (June): 391-406.

[25] Dynan, Karen and Maki, Dean (2001), "Does Stock Market Wealth Matter for Consumption?" Federal Reserve Board, mimeo.

[26] Epstein, L. and S. Zin (1989), "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica* 57, 937-69.

[27] Epstein, L. and S. Zin (1991), "Substitution, Risk Aversion and the Temporal Behavior of Stock Returns: An Empirical Investigation," *Journal of Political Economy* 99, 263-286.

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

[29] Gali, Jordi (1994), "Keeping Up with the Joneses: Consumption Externality, Portfolio Choice, and Asset Prices," *Journal of Money, Credit, and Banking* 26, 1-8.

[30] Gennotte, Gerard (1986), "Optimal Portfolio Choice Under Incomplete Information," *Journal of Finance* 41, 733-746.

[31] Giovannini, Alberto and Philippe Weil (1989), "Risk Aversion and Intertemporal Substitution in the Capital Asset Pricing Model," *NBER* Working Paper No. 2824.

[32] Grossman, Sanford and Robert Shiller (1981), "The Determinants of the Variability of Stock Market Prices," *American Economic Review* 71, 222-227.

[33] Grossman, Sanford and Guy Laroque (1990), "Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods," *Econometrica* 58, 25–51.

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

[35] Hall, Robert E (1988), "Intertemporal Substitution in Consumption." *Journal of Political Economy* 96(2), 339-357.

[36] Hansen, Lars and Kenneth Singleton (1983), "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *Journal of Political Economy* 91, 249-265.

[37] Hansen, Lars and Ravi Jagannathan (1991), "Implications of Security Market Data for Models of Dynamic Economies," *Journal of Political Economy* 99, 225-62.

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

[39] Hansen, Lars and Thomas Sargent (2003), *Robust Control and Economic Model Uncertainty*, Stanford University, mimeo.

[40] Kandel, Shmuel and Robert Stambaugh (1991) "Asset Returns and Intertemporal Preference," *Journal of Monetary Economics* 27: 39-71.

[41] Kasa, Kenneth (2004), "Robustness and Information Processing," Forthcoming *Review of Economic Dynamics.*

[42] Kennickell, A. B., M. Starr-McCluer, and B. Surette (2000), "Recent changes in U.S. family finances: Results from the 1998 Survey of Consumer Finances," Federal Reserve Bulletin, January 1–29.

[43] Kocherlakota, Narayana (1996), "The Equity Premium: It's Still a Puzzle," *Journal of Economic Literature* Vol. XXXIV, 42-71.

[44] Lettau, Martin and Harald Uhlig (2002), "Sharpe Ratios and Preferences: An Analytical Approach," *Macroeconomic Dynamics*, 6 (2), 242-265.

[45] Lucas, Robert E. (1978), "Asset Prices in an Exchange Economy." *Econometrica* 46(6), 1429-1445.

[46] Luo, Yulei (2005), "Consumption Dynamics under Information Processing Constraint," Manuscript, Princeton University, mimeo.

[47] Luo, Yulei and Eric Young (2005), "Rational Inattention and Aggregate Fluctuations," Princeton University and University of Virginia, mimeo.

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

[49] Maenhout, Pascal (2004), "Robust Portfolio Rules and Asset Pricing", *Review of Financial Studies* Vol. 17 (4), 951-983.

[50] Mankiw, Gregory and Ricardo Reis (2002), "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics* 117(4) 1295-328.

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

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

[53] Merton, Robert (1969), "Lifetime Portfolio Selection under Uncertainty: the Continuous Time Case," Review of Economics and Statistics 50, 247-257.

[54] Moscarini, Giuspeppe (2003), "Limited Information Capacity as a Source of Inertia," *Journal of Economic Control and Dynamics,* 2003-2035.

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

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

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

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

[59] Peng, Lin (2004), "Learning with Information Capacity Constraints," *Journal of Financial and Quantitative Analysis* 40, 307-330.

[60] Peng, Lin and Wei Xiong (2005), "Investor Attention, Overconfidence and Category Learning," *Journal of Financial Economics*, forthcoming.

[61] Piazzesi, Monika (2002), "Comment" on Gabaix and Laibson, "The 6D Bias and the Equity-Premium Puzzle". In *NBER Macroeconomics Annual* 2001, vol. 16, edited by Ben S. Bernanke and Ken Rogoff. Cambridge, MA: MIT Press.

[62] Reis, Ricardo (2003), "Inattentive Consumer," Harvard University, mimeo.

[63] Samuelson, Paul (1969), "Lifetime Portfolio Selection By Dynamic Stochastic Programming," Review of Economics and Statistics 51, 239-246.

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

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

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

[67] Sims, Christopher (2005), "Rational Inattention: A Research Agenda," Princeton University, mimeo.

[68] Starr-McCluer, Martha (2000), "Stock Market Wealth and Consumption," Federal Reserve Board, mimeo.

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

[70] Van Nieuwerburgh, Stijn and Veldkamp, Laura (2005a), "Information Acquisition and Portfolio Under-Diversification," New York University, mimeo.

[71] Van Nieuwerburgh, Stijn and Veldkamp, Laura (2005b), "Information Immobility and the Home Bias Puzzle," New York University, mimeo.

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

[73] Vissing-Jørgensen, Annette (2002), "Limited Asset Market Participation and Elasticity of Intertemporal Substitution,"*Journal of Political Economy* 110(4), 825-853.

[74] Weil, Philippe (1989), "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics* 24, 401-421.

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

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

[77] 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.

Table 1: Ultimate risk and optimal share in equities

	RRA $\gamma = 1.01$		RRA $\gamma = 1.5$	
channel capacity κ (nats)	ultimate risk	optimal share χ^*	ultimate risk	optimal share χ^*
0.025	5.3383	0.3469	5.3383	0.2313
0.04	2.0011	0.9254	2.0011	0.6169
0.1	1.2318	1.5033	1.2318	1.0022
0.5	1.0249	1.8069	1.0249	1.2046
1	1.0066	1.8398	1.0066	1.2265
3	1.0001	1.8517	1.0001	1.2344

Note that the above results are based on the following baseline parameter values: $\beta = 0.96, \mu - r^f + \frac{1}{2}\omega_u^2 = 6\%$, and $\omega_u^2 = 18\%$.

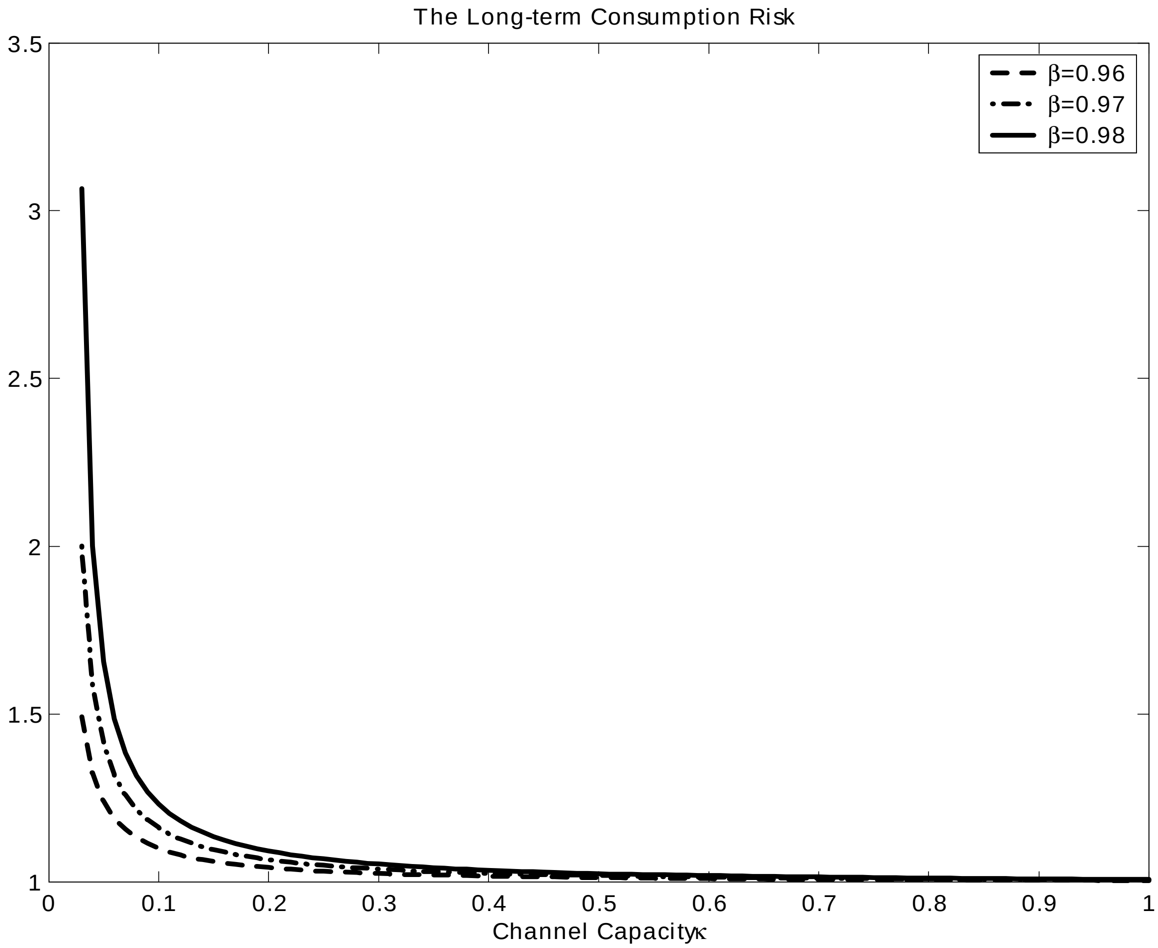


Figure 1: the relationship between long-term consumption risk and channel capacity.

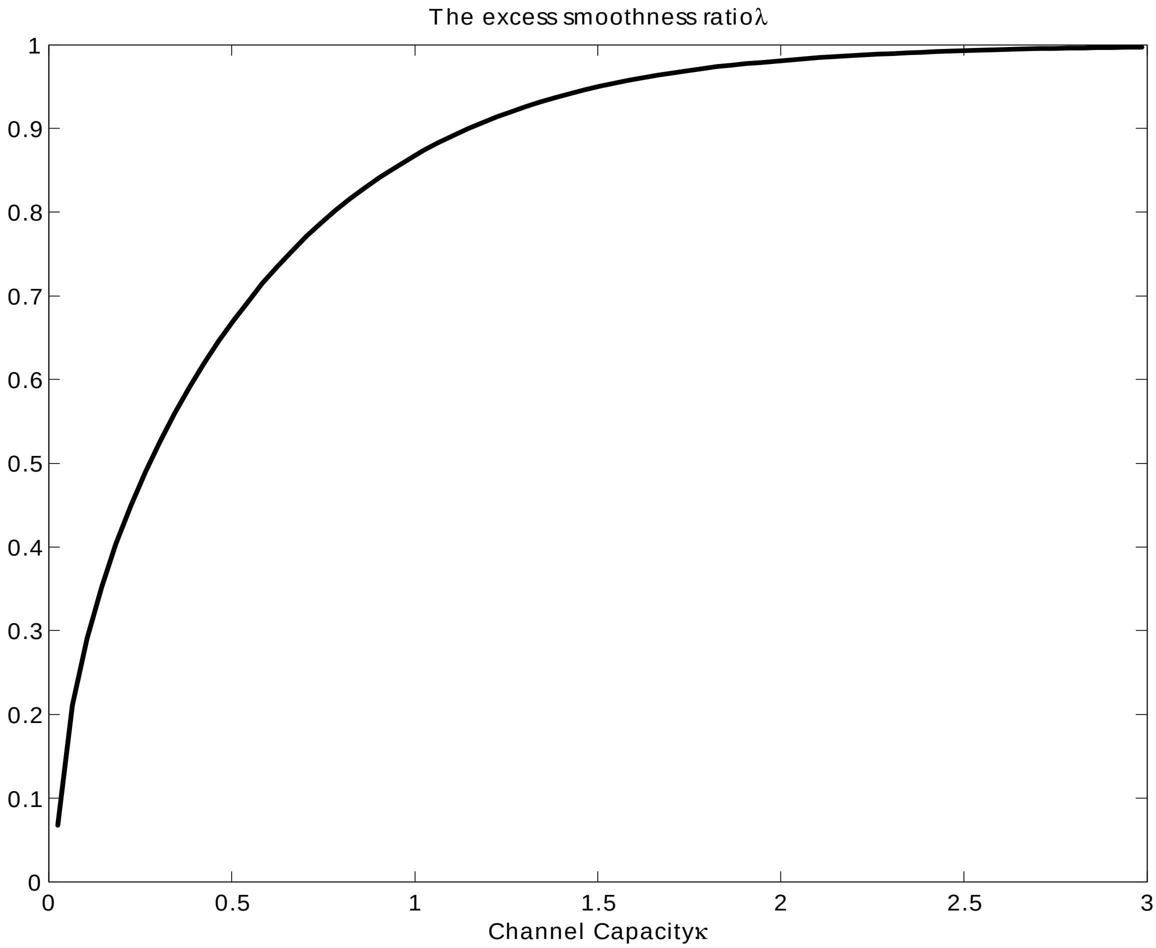


Figure 2: excess smoothness of aggregate consumption

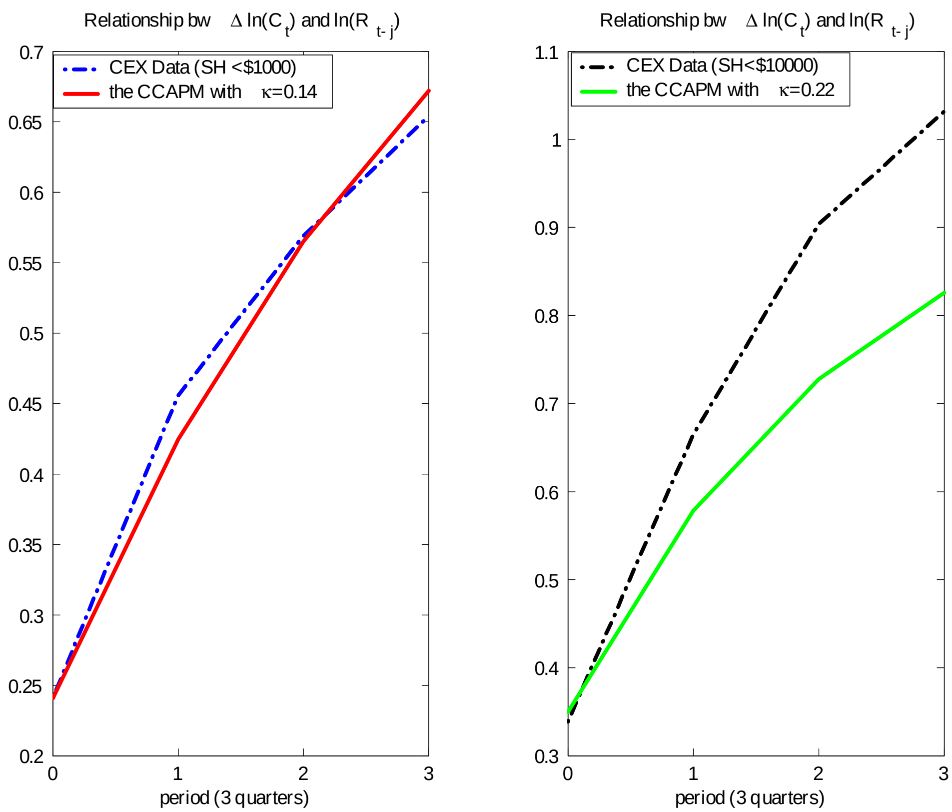
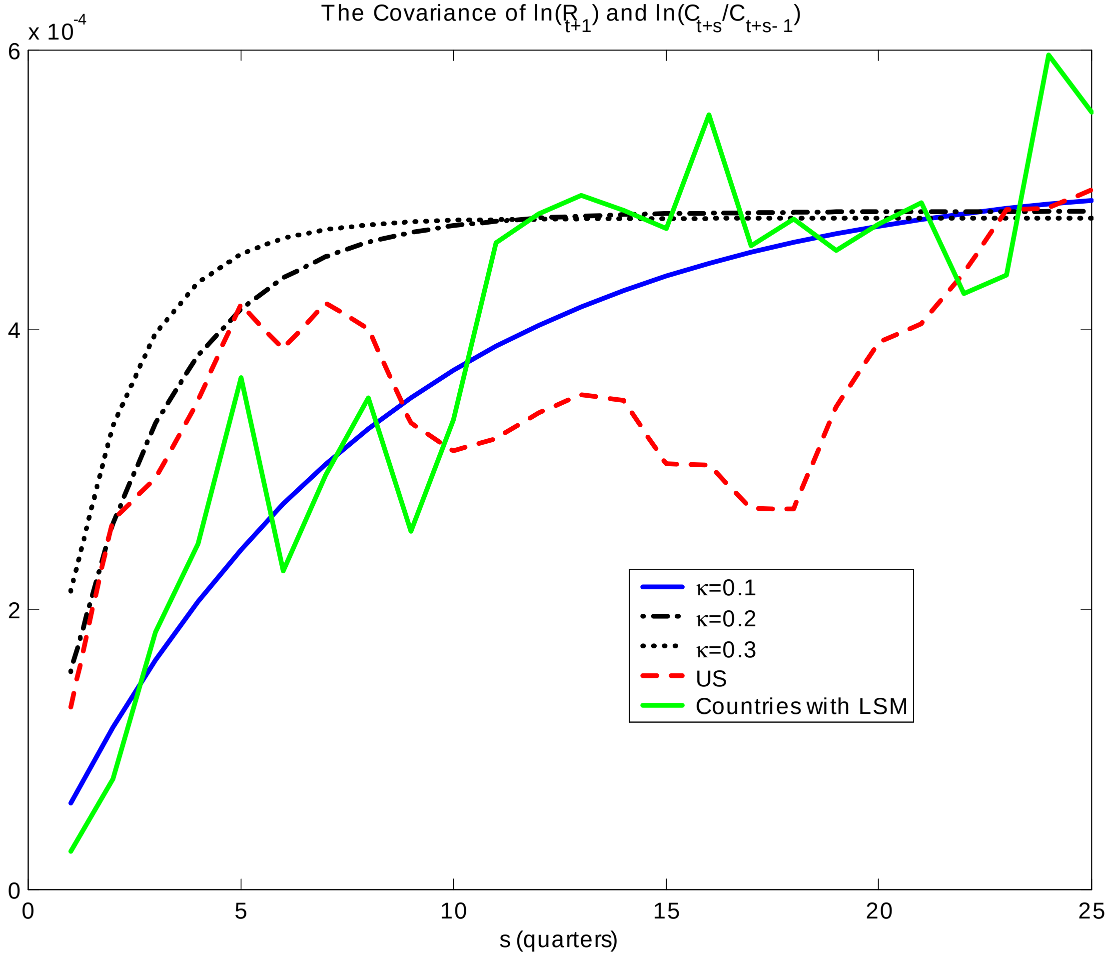

Relationship bw Δ ln(C t) and ln(R t-j)
CEX Data (SH <$1000)
the CCAPM with κ=0.14
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0
1
2
3
period (3 quarters)
Relationship bw Δ ln(C t) and ln(R t-j)
CEX Data (SH<$10000)
the CCAPM with κ=0.22
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0
1
2
3
period (3 quarters)

Figure 3



The Covariance of ln(R_{t+1}) and ln(C_{t+s}/C_{t+s- 1})
x 10^-4
6
4
2
0
0
5
10
15
20
25
s (quarters)
κ=0.1
κ=0.2
κ=0.3
US
Countries with LSM

Figure 4