

Rational Inattention, Long-run Consumption Risk, and Portfolio Choice*

Yulei Luo[†]

Forthcoming in *Review of Economic Dynamics*

February 4, 2010

Abstract

This paper explores how the introduction of rational inattention (RI) – that agents process information subject to finite channel capacity – affects optimal consumption and investment decisions in an otherwise standard intertemporal model of portfolio choice. We first explicitly derive optimal consumption and portfolio rules under RI and then show that introducing RI reduces the optimal share of savings invested in the risky asset because inattentive investors face greater long-run consumption risk. We also show that the investment horizon matters for portfolio allocation in the presence of RI, even if investment opportunities are constant and the utility function of investors is constant relative risk aversion. Second, after aggregating across investors, we show that introducing RI can better explain the observed joint dynamics of aggregate consumption and the equity return. Finally, we show that RI increases the implied equity premium because investors under RI face greater long-run consumption risk and thus require higher compensation in equilibrium.

JEL Classification Numbers: *D81, E21, G11.*

Keywords: Rational Inattention, Long-run Consumption Risk, Long-term Portfolio Choice, Aggregate Consumption.

*An earlier version of this paper was circulated under the title: “Rational Inattention, Portfolio Choice, and the Equity Premium.” I would like to thank Christopher Sims, Michael Haliassos, Winfried Koeniger, Per Krusell, Urban Jermann (the Editor), Jonathan Parker, Ricardo Reis, Raman Uppal, Wei Xiong, Chi-wa Yuen, and two anonymous referees, as well as seminar and conference participants at Princeton University, European University Institute, University of Warwick, University of Guelph, University of Hong Kong, Guanghua School of Management at Peking University, and the SED conference for helpful comments and suggestions.

[†]School of Economics and Finance, The University of Hong Kong, Hong Kong. E-mail: yluo@econ.hku.hk.

1 Introduction

Optimal asset allocation is a classic problem in finance and macroeconomics. In a canonical two-asset intertemporal portfolio choice model in which investment opportunities are constant and the utility function is power, the optimal share of savings invested in the risky asset is proportional to the expected excess return and the reciprocal of the volatility of the return to the risky asset and the coefficient of relative risk aversion. However, it is independent of either the investment horizon or the time discount factor of investors. Hence, the long-term portfolio choice is the same as the myopic portfolio choice. Furthermore, given the observed mean and volatility of the return to the risky asset and plausible degree of risk aversion, the canonical model generates a counterfactually high stock market participation rate.¹ Another important topic that is closely related to portfolio choice is asset pricing. According to the canonical consumption-based capital asset pricing theory (CCAPM), the expected excess return on any risky portfolio over the risk-free interest rate is determined by the covariance of the excess return with contemporaneous consumption growth and the coefficient of relative risk aversion. Given the observed low covariance between equity returns and contemporaneous consumption growth, the canonical CCAPM theory thus predicts that equities are not very risky. 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 coefficient of relative risk aversion must be very high. Mehra and Prescott (1985) examined this issue in the Lucas-type asset pricing framework and were the first to call it “the equity premium puzzle”.²

An implicit but key assumption in the standard rational expectations (RE) version of the intertemporal portfolio choice model is that investors have *unlimited* information-processing capacity and thus can observe the state variable(s) without error. Consequently, they can react *instantaneously and completely* to any innovations to equity returns. However, the assumption that ordinary agents can observe the relevant state without errors is too strong to be consistent with the reality because complete observation requires costless computing and information processing ability that normal agents may not possess in the real world.³ In fact, ordinary people face many competing demands for their information capacity every period, so the amount

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

²For the history of this puzzle, see Hansen and Singeton (1983), Mehra and Prescott (1985), Hansen and Jagannathan (1991), Cochrane and Hansen (1992), and Jermann (1998). Kocherlakota (1996) and Campbell (2003) gave useful reviews on this topic.

³As shown in Shannon (1948), measuring a real-value stochastic process without error implies an *infinite* amount of information processing capacity.

of capacity devoted to processing financial information could be much lower than their total capacity. Hence, it is reasonable to assume that ordinary people only have *limited* information-processing capacity. As a result, they cannot observe the state(s) perfectly and thus have to react to the innovations *slowly and incompletely*. Sims (2003) first introduced this type of information capacity constraints into economics and called it “rational inattention” (henceforth, RI). In Sims’ RI framework, the *entropy* concept is imported from information theory to measure the uncertainty of a random variable; thus the reduction in the entropy can be used to measure information flow. If there is an input and an output in a channel, the capacity is the maximum amount of information that flows through the channel. For finite channel capacity, the reduction in entropy is bounded by the capacity. Sims also solved linear-quadratic (LQ) optimization problems with finite capacity and showed that the optimal reactions of individuals are slow with respect to fundamental shocks and quick to the endogenous noises. This finding thus simultaneously explains the observed behavior of inertia (slow reaction to fundamentals) and volatility (quick reaction to noises).

The main objective of this paper is to apply the RI idea in an intertemporal model of portfolio choice and examine how RI affects optimal long-term asset allocation and the joint dynamics of aggregate consumption and asset returns.⁴ As the first contribution of this paper, we explicitly solve for optimal consumption and portfolio rules in an RI version of the intertemporal model of portfolio choice. Specifically, we show that consumption adjusts slowly to the innovation to the equity return and quickly to its own shock (the endogenous error). The intuition is simple: ordinary investors only devote finite information capacity to observing the state of the evolution of their financial wealth, and thus the state variable(s) cannot be perfectly observed. In other words, information about changes in the true state cannot be entirely incorporated into decision making instantaneously and completely, so investors have to take some time to digest new information.

Second, given the slow adjustment in consumption to the wealth shocks due to RI, the quantity of the risk of the risky portfolio should be determined by its ultimate (long-term) consumption risk instead of its contemporaneous risk. Parker (2003) and Parker and Jullard (2005) presented convincing empirical evidence to argue that the ultimate consumption risk is a

⁴A number of papers have explored the implications of delayed adjustment for asset pricing and aggregate dynamics from another perspective. For example, Lynch (1996) examined the effects of infrequent decision-making for aggregate consumption and asset returns; Marshall and Parekh (1999) addressed how costs of adjustment help explain the asset return puzzles; Gabaix and Laibson (2001) followed their research line and showed that updating optimal decisions infrequently can better explain the equity premium puzzle in a continuous-time model. We will compare their results to RI’s predictions in Section 2.

better measure for the riskiness of the risky portfolio than the contemporaneous risk because in the data consumption takes many periods to adjust to the innovations to risky assets. Therefore, under RI, the ultimate consumption risk is greater than the contemporaneous risk and investors with finite capacity would choose to invest less in the risky asset. It is irrational for investors to invest a large fraction of their wealth in the risky asset if they cannot allocate enough channel capacity to monitoring their financial wealth, because the innovations to their financial wealth can generate large consumption risk in the long run if the capacity is low.

Third, given the solution of optimal asset allocation, we show that the long-term consumption risk and the demand for the risky asset depend both on the degree of RI and on the discount factor when the coefficient of relative risk aversion is close to 1. In this infinite horizon setting, the discount factor can be used to characterize the effective investment horizon;⁵ hence, the fact that the demand for the risky asset depends on the discount factor also means that the investment horizon matters for portfolio allocation. In addition, if we allow for a heterogeneous degree of inattention in the model, RI could provide an alternative explanation for “the limited stock market participation” observed in the US data because investors with a very low degree of attention might face extremely large long-term consumption risk, which restricts their participation in the stock market.

Finally, after aggregating across all investors, we demonstrate that RI can better explain the observed joint behavior of aggregate consumption and the equity return in the U.S. data in the following dimensions: (i) the smooth process of aggregate consumption, (ii) low contemporaneous correlation and covariance between aggregate consumption and the equity return, (iii) positive autocorrelation of consumption growth; and (iv) positive covariance between aggregate consumption and the equity return. It is also shown that holding the discount factor constant, increasing the degree of inattention is observationally equivalent to increasing the effective coefficient of relative risk aversion that matters for asset allocation. Hence, even if the true value of the coefficient of relative risk aversion is only around 1, the effective coefficient would be higher than 1 if the investor is highly inattentive. In other words, RI has the potential to reconcile low estimates of risk aversion obtained from experimental evidence or introspection, with high estimated values of risk aversion based on asset pricing data.

Some recent papers incorporate explicit information processing constraints into a variety of theoretical models and explore how they affect individual decisions, optimal monetary policy, and equilibrium. For example, Moscarini (2004) derived optimal time-dependent adjustment

⁵That is, the lower the time discount factor, the shorter the investment horizon.

rules from the information constraints in a continuous-time framework. Peng (2004) explored the effects of information constraints on the equilibrium asset price dynamics and consumption behavior under the continuous-time CARA framework. Adam (2005) analyzed the effects of imperfect common knowledge on monetary policy. Huang and Liu (2007) showed that inattention to important news may make investors overinvest or under-invest.⁶ Luo (2008) examined how RI helps explain the excess smoothness and excess sensitivity puzzles in the consumption literature. Maćkowiak and Wiederholt (2009) solved a dynamic general equilibrium model with RI, and found that households and firms devote much larger fractions of channel capacity to monitoring idiosyncratic conditions than to monitoring aggregate technology and monetary policy conditions. Peng and Xiong (2005) discussed investors’ attention and overconfidence. Van Nieuwerburgh and Veldkamp (2010) studied information acquisition and portfolio under-diversification.

This paper is organized as follows. Section 2 solves a long-term portfolio choice model with RI and examines the implications of RI for optimal consumption and investment decisions. Section 3 examines how RI affects the joint dynamics of aggregate consumption and asset returns and the implied equity premium. Section 4 presents some suggestive empirical evidence. Section 5 concludes.

2 An Intertemporal Portfolio Choice Model with Rational Inattention

In this section, we first present and review a standard intertemporal portfolio choice model in the vein of Merton (1969), and then discuss how to incorporate RI into this otherwise standard model. Following the log-linear approximation method proposed by Campbell (1993), Viceira (2001), and Campbell and Viceira (2002), we explicitly solve an RI version of the portfolio choice model after considering the long-term consumption risk facing investors. A major advantage of the log-linearization approach is that we can approximate the original nonlinear problem by a log linear-quadratic (LQ) framework when the coefficient of relative risk aversion is close to 1 and thus can justify the optimality of Gaussian posterior uncertainty under RI. We finally discuss the implications of RI for optimal consumption and investment decisions. In the next

⁶They used “rational inattention” to describe a situation in which information is updated *completely but infrequently*. This alternative formulation is more tractable but is different from Sims’ RI idea – it permits agents to process infinite amounts of information when they choose to do so. However, the essence of RI proposed by Sims (2003) is that agents cannot use all *available* information because most of the time they are not paying that much attention to market signals and thus cannot digest all of the available information.

section, we will examine how RI affects the joint dynamics of aggregate consumption and the asset return as well as the equity premium in equilibrium.

2.1 Specification and Solution of the Standard RE Portfolio Choice Model

It is helpful to present the standard portfolio choice model and then discuss how to introduce RI in this framework before setting up and solving the portfolio choice model with RI. Here we consider a simple intertemporal model of portfolio choice with a continuum of identical investors.⁷ We first assume that investors choose consumption and asset holdings to maximize the intertemporal time-separable utility, defined over consumption:

$$\max_{\{C_t, \alpha_t\}_{t=0}^{\infty}} E_0 \left[\sum_{t=0}^{\infty} \beta^t u(C_t) \right], \quad (1)$$

where $u(C_t) = \frac{C_t^{1-\gamma}-1}{1-\gamma}$ is the power utility function, C_t represents individual's consumption at time t , β is the discount factor, and γ is the coefficient of relative risk aversion. When $\gamma = 1$, the utility function becomes logarithmic, $\log C_t$.

For simplicity, we assume that the investment opportunity set is constant. There are two tradable financial assets in the model economy: Asset e is risky, with iid one-period log (continuously compounded) return r_{t+1}^e , while the other asset f is riskless, with constant log return given by r^f . We refer to asset e as the market portfolio of equities, and asset f as savings or checking accounts. Furthermore, we 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 $\text{var}[u_{t+1}] = \omega^2$.

The intertemporal budget constraint for the investor is

$$A_{t+1} = R_{t+1}^p (A_t - C_t), \quad (2)$$

where A_{t+1} is the individual's financial wealth which is defined as the value of financial assets carried over from period t at the beginning of period $t + 1$, $A_t - C_t$ is savings, and R_{t+1}^p is the one-period gross return on savings given by

$$R_{t+1}^p = \alpha_t \left(R_{t+1}^e - R^f \right) + R^f, \quad (3)$$

where $R_{t+1}^e = \exp(r_{t+1}^e)$, $R^f = \exp(r^f)$, and $\alpha_t = \alpha$ is the proportion of savings invested in the

⁷The model is based on Campbell (1993), Viceira (2001), and Campbell and Viceira (2002), and is widely adopted in the macroeconomics and finance literature.

risky asset.⁸ As shown in Campbell (1993), the following approximate expression for the log return on wealth holds:

$$r_{t+1}^p = \alpha(r_{t+1}^e - r^f) + r^f + \frac{1}{2}\alpha(1 - \alpha)\omega^2. \quad (4)$$

Given the above model specification, it is well known that this simple discrete-time model can not be solved analytically.⁹ We then follow the log-linearization method proposed in Campbell (1993), Viceira (2001), and Campbell and Viceira (2002) to approximately solve the model in closed-form. Specifically, the original intertemporal budget constraint, (2), can be written in log-linear form:

$$\Delta a_{t+1} = (1 - 1/\phi)(c_t - a_t) + \psi + r_{t+1}^p, \quad (5)$$

where $c - a = E[c_t - a_t]$ is the unconditional expectation of $c_t - a_t$, $\phi = 1 - \exp(c - a)$, $\psi = \log \phi - (1 - 1/\phi) \log(1 - \phi)$, and lowercase letters denote logs. The optimal consumption and portfolio rules for this standard intertemporal model are as follows:¹⁰

$$c_t = b_0 + a_t, \quad (6)$$

$$\alpha = \frac{\mu - r^f + 0.5\omega^2}{\gamma\omega^2}, \quad (7)$$

where

$$b_0 = \log \left(1 - \exp \left(\left(\frac{1}{\gamma} - 1 \right) E_t [r_{t+1}^p] + \frac{1}{\gamma} \log \beta + \frac{1}{2\gamma} (1 - \gamma)^2 \text{var}_t [r_{t+1}^p] \right) \right). \quad (8)$$

Hence, the optimal 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. Furthermore, combining (5) with (6) gives the expression for the change in consumption:¹¹

$$\Delta c_{t+1} = \alpha u_{t+1}. \quad (9)$$

For an economy with identical consumers/investors, (9) clearly shows that aggregate consumption is unpredictable using past information and the relative volatility of aggregate consumption

⁸ Given iid equity returns and power utility function, the share invested in equities, α_t , is constant over time.

⁹ Of course, it can be solved by using numerical methods adopted widely in the modern consumption and portfolio choice literature.

¹⁰ See the Appendix in Viceira (2001) for detailed derivations.

¹¹ Under RE, individual consumption growth and aggregate consumption growth have the same expression. In addition, here we impose the restriction that $\beta \bar{R}_p = 1$ (i.e., $\log \beta + \alpha(\mu - r^f) + r^f + \frac{1}{2}\alpha(1 - \alpha)\omega^2 = 0$) to eliminate a constant term that is irrelevant to consumption dynamics.

to the equity innovation, $\frac{\text{sd}[\Delta c_{t+1}]}{\text{sd}[r_{t+1}^e]}$, is α .¹² Given that $\omega = 0.16$ and $\pi = \mu - r^f = 0.06$ obtained from the annualized U.S. quarterly data in Campbell (2003), to generate the realistic share invested in the risky asset, $\alpha = 22\%$, we need a risk aversion, γ , of 13, which seems larger than the values most economists find plausible. Furthermore, when the model economy generates $\alpha = 22\%$, the predictions of the standard Merton model for the joint behavior of aggregate consumption and the equity return are not consistent with the empirical evidence:

- (i) In the U.S. data aggregate consumption growth is much smoother than the equity return: As documented in Campbell (2003), the annualized standard deviation of consumption growth is about 1%, while the annualized standard deviation of the equity return is about 16%. By contrast, the standard model predicts that $\text{sd}[\Delta c_{t+1}] = \alpha \text{sd}[r_{t+1}^e] = 3.5\%$, which is larger than its empirical counterpart.
- (ii) In the data the contemporaneous correlation of consumption growth with the equity return is about 0.34 at a 1-year horizon, while the model predicts that the correlation is 1.
- (iii) In the data the contemporaneous covariance between aggregate consumption and the equity return is about $6 \cdot 10^{-4}$, while the model generates a much higher covariance: $5.6 \cdot 10^{-3}$.
- (iv) The first-order autocorrelation of consumption growth is significantly positive and is about 0.2 in the U.S. data. By contrast, the standard RE model predicts that consumption growth is iid.
- (v) The covariance of consumption growth with lagged equity returns is positive in the data, while the model predicts that it is 0.

In the next section, we will show how rational inattention, a consequence of information-processing constraints, can be incorporated into the above standard Merton model and we will discuss how this makes the model fit the data better in the above five aspects.

2.2 Incorporating Rational Inattention into the Standard Model

Following Sims (2003), we introduce rational inattention (RI) into the otherwise standard intertemporal portfolio choice model by assuming consumers/investors face information-processing constraints and have only finite Shannon channel capacity to observe the state of the world. As in Sims (2003), we use the concept of entropy from information theory to characterize the

¹²Here we use $\text{sd}[\cdot]$ to denote standard deviation.

uncertainty about a random variable; the reduction in entropy is thus a natural measure of information flow. Formally, entropy is defined as the expectation of the negative of the log of the density function, $-E[\log(f(X))]$ (see Shannon 1948 and Cover and Thomas 1991 for details). For example, the entropy of a discrete distribution with equal weight on two points is simply $E[\log(f(X))] = -0.5 \log_2(0.5) - 0.5 \log_2(0.5) = 1$, and the unit of information transmitted is called one “bit”. In this case, an agent can remove all uncertainty about X if the capacity devoted to monitoring X is $\kappa = 1$ bit.

With finite capacity $\kappa \in (0, \infty)$, the true state a following a continuous distribution cannot be observed without errors and thus the information set at time $t + 1$, \mathcal{I}_{t+1} , is generated by the entire history of noisy signal $\{a_j^*\}_{j=0}^{t+1}$. Following Sims (2003), Adam (2005), and Maćkowiak and Wiederholt (2009), we assume in this paper that the noisy signal takes the additive form: $a_{t+1}^* = a_{t+1} + \xi_{t+1}$, where ξ_{t+1} is the endogenous noise caused by finite capacity. We further assume that ξ_{t+1} is an iid idiosyncratic shock and is independent of the fundamental shock. Note that the reason that the RI-induced noise is idiosyncratic is that the endogenous noise arises from the consumer’s own internal information-processing constraint (i.e., finite information-processing channel). The investors with finite capacity will choose a new signal $a_{t+1}^* \in \mathcal{I}_{t+1} = \{a_1^*, a_2^*, \dots, a_{t+1}^*\}$ that reduces the uncertainty of the state variable a_{t+1} as much as possible. Formally, this idea can be described by the following information constraint

$$\mathcal{H}(a_{t+1}|\mathcal{I}_t) - \mathcal{H}(a_{t+1}|\mathcal{I}_{t+1}) = \kappa, \quad (10)$$

where κ is the investor’s information channel capacity, $\mathcal{H}(a_{t+1}|\mathcal{I}_t)$ denotes the entropy of the state *prior to* observing the new signal at $t + 1$, and $\mathcal{H}(a_{t+1}|\mathcal{I}_{t+1})$ is the entropy *after* observing the new signal. κ imposes an upper bound on the amount of information – that is, the change in the entropy – that can be transmitted in any given period. Finally, we assume that the noise ξ_{t+1} is Gaussian.¹³ Finally, following the literature, suppose that the ex ante a_{t+1} is a Gaussian random variable.

Following Sims (2003, 2005), we assume that the typical investor maximizes his lifetime utility subject to the budget constraint, (5), as well as the information-processing constraint. The dynamic optimization problem of the investor can be written as

$$\hat{v} = \max_{\{c_t, \mathcal{D}_t\}} E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left[c_t + \frac{1}{2}(1 - \gamma)c_t^2 \right] \right\} \quad (11)$$

¹³As shown in Sims (2003), within the linear-quadratic-Gaussian setting, Gaussian noise is optimal.

subject to

$$\Delta a_{t+1} = (1 - 1/\phi)(c_t - a_t) + \psi + r_{t+1}^p \quad (12)$$

$$a_{t+1} | \mathcal{I}_{t+1} \sim \mathcal{D}_{t+1}, \quad (13)$$

$$a_t | \mathcal{I}_t \sim \mathcal{D}_t, \quad (14)$$

given $a_0 | \mathcal{I}_0 \sim N(\hat{a}_0, \Sigma_0)$, and the information-processing constraint, (10), i.e., 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. The expectation is formed under the assumption that $\{c_t\}_{t=0}^\infty$ are chosen under the information processing constraints. For simplicity, here we 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. Note that in the RI model the *perceived* state variable is not the traditional state variable (e.g., the wealth level a_t in this model), but the so-called *information state*: the distribution of the true state variable, a_t , conditional on the information set available at time t , \mathcal{I}_t .

As shown in Sims (2003, 2005), the linear-quadratic (LQ) specification can rationalize Gaussian *posterior* uncertainty as optimal, while the non-LQ setup easily generates optimal non-Gaussian posterior uncertainty.¹⁴ Therefore, in our approximate LQ setting, ex-post Gaussian uncertainty is optimal, that is, \mathcal{D}_{t+1} is normal:

$$a_{t+1} | \mathcal{I}_{t+1} \sim N(\hat{a}_{t+1}, \Sigma_{t+1}),$$

where $\hat{a}_{t+1} = E[a_{t+1} | \mathcal{I}_{t+1}]$ and $\Sigma_{t+1} = \text{var}[a_{t+1} | \mathcal{I}_{t+1}]$ are the conditional mean and variance of a_{t+1} , respectively. The idiosyncratic error ξ_{t+1} and the noisy signal $a_{t+1}^* = a_{t+1} + \xi_{t+1}$ are both chosen to be Gaussian such that

$$\frac{1}{2} (\log \Psi_t - \log \Sigma_{t+1}) = \kappa, \quad (15)$$

where $\Sigma_{t+1} = \text{var}[a_{t+1} | \mathcal{I}_{t+1}]$ and $\Psi_t = \text{var}[a_{t+1} | \mathcal{I}_t]$ are the posterior and the prior variance of a_{t+1} , respectively. This means that given a finite capacity κ per time unit, the optimizing

¹⁴For example, Sims (2005) showed in a simple two-period saving problem that when the utility function is non-quadratic, the optimal ex post uncertainty follows non-Gaussian distribution. Fully nonlinear versions of the RI problem have either very short horizons or very computation demanding – the state of the world is the distribution of true states and this distribution is not well-behaved (it is not generally a member of a known class of distributions and tends to have ‘holes,’ making it difficult to characterize with a small number of parameters).

consumer would choose a signal that reduces the entropy by $\frac{1}{2}(\log \Psi_t - \log \Sigma_{t+1})$.¹⁵ Note that in the univariate state case this information constraint completes the characterization of the optimization problem.

Given the budget constraint, (5), taking conditional variances on both sides yields:

$$\text{var}_t [a_{t+1}] = \text{var}_t [r_{t+1}^p] + (1/\phi)^2 \Sigma_t. \quad (16)$$

Substituting (16) into (15) then gives $\kappa = \frac{1}{2} [\log (\text{var}_t [r_{t+1}^p] + (1/\phi)^2 \Sigma_t) - \log \Sigma_{t+1}]$, which has a steady state $\Sigma = \frac{\text{var}_t [r_{t+1}^p]}{\exp(2\kappa) - (1/\phi)^2}$, where $\text{var}_t [r_{t+1}^p] = \alpha^2 \omega^2$.

We can now apply the separation principle in this log-LQ model to obtain the following modified consumption function by replacing a_t with \hat{a}_t in (6):¹⁶

$$c_t = \log(1 - \beta) + \hat{a}_t \quad (17)$$

and the perceived state \hat{a}_t is characterized by the following Kalman filter equation:

$$\hat{a}_{t+1} = (1 - \theta) \left[\left(1 - \frac{1}{\beta}\right) c_t + \frac{1}{\beta} \hat{a}_t - \log \beta + \psi \right] + \theta a_{t+1}^*. \quad (18)$$

Combining them gives the evolution of \hat{a}_t :

$$\hat{a}_{t+1} = (1 - \theta) \hat{a}_t + \theta a_{t+1}^*, \quad (19)$$

where $\theta = 1 - \frac{1}{\exp(2\kappa)}$ is the optimal weight on the new observation (i.e., the Kalman gain), $a_{t+1}^* = a_{t+1} + \xi_{t+1}$ is the observed signal, and ξ_{t+1} are the iid noise with $\text{var} [\xi_{t+1}] = \Sigma/\theta$.

Hence, combining equations (5), (17), with (19) gives the expression for individual consumption growth:¹⁷

$$\Delta c_{t+1} = \theta \left[\frac{\alpha u_{t+1}}{1 - ((1 - \theta)/\beta) \cdot L} + \left(\xi_{t+1} - \frac{(\theta/\beta) \xi_t}{1 - ((1 - \theta)/\beta) \cdot L} \right) \right], \quad (20)$$

where L is the lag operator. (See Appendix 6.1 for the derivation.)

¹⁵Note that given Σ_t , choosing Σ_{t+1} is equivalent with choosing the noise $\text{var} [\xi_t]$ since the usual updating formula for the variance of a Gaussian distribution is

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

where Ψ_t is a function of Σ_t .

¹⁶Here we have used the approximation result that $\phi = \beta$. Note that ϕ is independent of the degree of attention, κ , because ϕ approaches to β as γ converges to 1.

¹⁷Note that this MA(∞) expression requires that $(1 - \theta)/\beta < 1$, which is equivalent to $\kappa > \frac{1}{2} \log(1/\beta)$.

Note that none of the above expressions for consumption, the perceived state, and the change in consumption are the final solutions because the optimal fraction of savings invested in the stock market, α , is still *undetermined*. The reason is that we need to use the Euler equation to determine the optimal allocation in risky assets. However, as we will discuss in the next subsection, the standard Euler equation does not hold in the RI model because consumption under information-processing constraints adjusts *gradually and incompletely*. Therefore, we will discuss how to determine the optimal portfolio allocation after considering the long-term Euler equation and the ultimate consumption risk.

2.3 Long-term Consumption Risk and the Demand for the Risky Asset

Parker (2003) and Parker and Julliard (2005) provided convincing empirical evidence to argue 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.¹⁸ Hansen, Heaton, and Li (2006) found a small Long-run risk component in the consumption data. Hence, we need to use the long-term consumption risk to measure the risk of the equity in the RI model because the RI model predicts that consumption reacts gradually and with delay to the innovations to the equity.

In this subsection, we 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 function and the changes in consumption gives us a complete solution to this simple optimal consumption and portfolio choice model with RI. Following Parker's work, we 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 incompletely, it can rationalize the conclusion in Parker's papers that consumption risk should be long term instead of contemporaneous.¹⁹

Given the above analytical solution for consumption growth, it is straightforward to calculate the ultimate consumption risk under RI. Specifically, when agents behave optimally but only have finite channel capacity, we have the following equality for the risky asset e and the risk-free

¹⁸Bansal and Yaron (2004) also documented that consumption and dividend growth rates contain a long-run component. An adverse change in the long-run component lowers asset prices and thus makes holding equity very risky for investors.

¹⁹Note that under RE, the contemporaneous consumption risk is the same as the ultimate consumption risk because consumption adjusts to wealth shocks instantly and completely.

asset f :²⁰

$$E_t \left[R_{t+1}^e C_{t+1+S}^{-\gamma} \right] = E_t \left[R^f C_{t+1+S}^{-\gamma} \right],$$

which can be transformed to the following stationary form:

$$E_t \left[R_{t+1}^e (C_{t+1+S}/C_t)^{-\gamma} \right] = E_t \left[R^f (C_{t+1+S}/C_t)^{-\gamma} \right], \quad (21)$$

where the expectation $E_t[\cdot]$ is conditional on the entire history of the economy up to t and S is the horizon of many periods in the future over which consumption response under RI is studied.²¹ (See Appendix 6.2 for the derivation.) Given the expression for consumption dynamics, it is clear that S is infinitely large because consumption takes infinite periods to react to the exogenous shock. The standard equality, $E_t[R_{t+1}^e C_{t+1}^{-\gamma}] = E_t[R^f C_{t+1}^{-\gamma}]$, does not hold here because consumption reacts slowly with respect to the innovations to equity returns and thus cannot finish adjusting immediately and completely.

Log-linearizing equation (21) yields

$$E_t \left[r_{t+1}^e \right] - r^f + \frac{1}{2}\omega^2 = \gamma \text{covar}_t \left[c_{t+1+S} - c_t, r_{t+1}^e \right], \quad (22)$$

which means that the expected asset return can be written as:

$$E_t \left[r_{t+1}^e \right] - r^f + \frac{1}{2}\omega^2 = \sum_{j=0}^S \text{covar}_t \left[\Delta c_{t+1+j}, r_{t+1}^e \right], \quad (23)$$

where we have used $\gamma \simeq 1$, $c_{t+1+S} - c_t = \sum_{j=0}^S \Delta c_{t+1+j}$, and Δc_{t+1+j} is given by (20). Specifically, the long-run impact of the innovation in the equity return on consumption growth can be rewritten as

$$\lim_{S \rightarrow \infty} \left(\sum_{j=0}^S \text{covar}_t \left[\Delta c_{t+1+j}, r_{t+1}^e \right] \right) = \varsigma \alpha \omega^2, \quad (24)$$

²⁰The 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

$$C_t^{-\gamma} = E_t \left[\beta^{S+1} C_{t+S+1}^{-\gamma} R_{t+1}^e (R^f)^S \right].$$

²¹This measure has some appealing features; see Parker (2003) and Parker and Jullard (2005) for detailed discussions.

where ς is the ultimate consumption risk:

$$\varsigma = \frac{\theta}{1 - (1 - \theta)/\beta} > 1 \quad (25)$$

when the restriction $1 - (1 - \theta)/\beta > 0$, i.e., $\theta > 1 - \beta$ holds. Note that this mild parameter restriction must be imposed to guarantee convergence. (See the MA(∞) process of consumption growth, (20).) In addition, if $\theta < 1 - \beta$, it is clear that the steady state conditional variance Σ does not exist. By simple calculation we obtain $\frac{\partial \varsigma}{\partial \theta} < 0$ and $\frac{\partial \varsigma}{\partial \beta} < 0$ because $\beta, \theta \in (0, 1)$. This ultimate consumption risk should be a better measure of the riskiness of the risky asset than the contemporaneous risk, as consumption adjusts gradually to the shocks to asset returns in the presence of RI. Figure 1 illustrates the relationship between the ultimate consumption risk (ς) and the Kalman filter gain (θ) for different β . (Note that θ also measures the degree of attention.) It clearly shows that given β the ultimate consumption risk decreases with the degree of attention; that is, the more attention of the investor devoted to monitoring the evolution of his financial wealth, the less the ultimate consumption risk of the risky asset held by him.

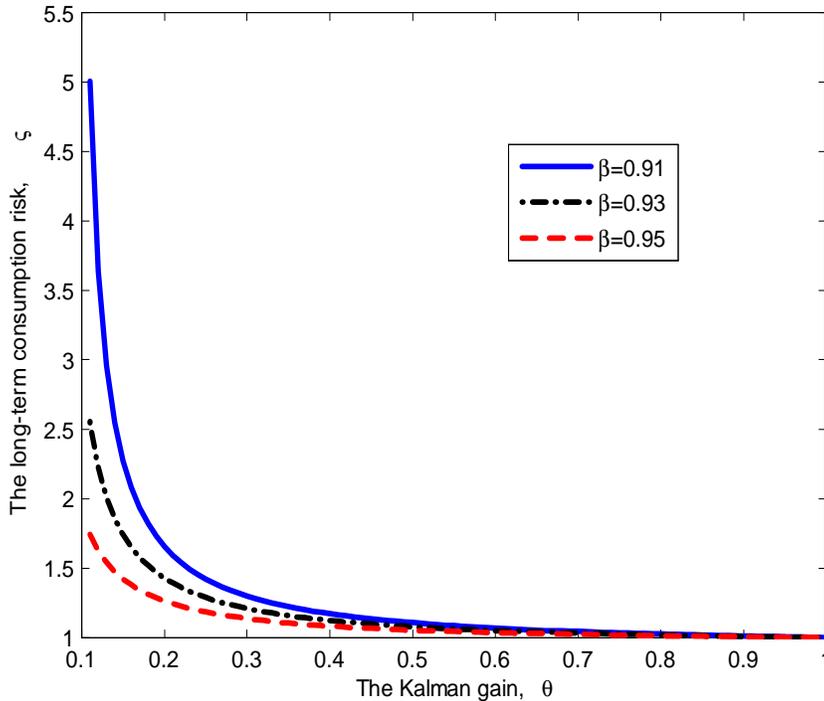


Figure 1: Effects of RI on long-run consumption risk.

Combining equations (23), (24), with (17) gives us optimal consumption and portfolio rules

under RI. The following proposition gives a complete characterization of the model's solution:

Proposition 1 *Consider the optimal consumption and portfolio choice problem (10)-(14). Suppose that γ is close to 1, and $\theta > 1 - \beta$. We then obtain that the optimal share invested in the equity is*

$$\alpha^* = \frac{1}{\varsigma} \frac{\mu - r^f + 0.5\omega^2}{\gamma\omega^2}, \quad (26)$$

where $\frac{1}{\varsigma} = \frac{1-(1-\theta)/\beta}{\theta} < 1$. Furthermore, the consumption function is

$$c_t^* = \log(1 - \beta) + \hat{a}_t, \quad (27)$$

the actual state variable evolves according to

$$a_{t+1} = \frac{1}{\beta}a_t + \left(1 - \frac{1}{\beta}\right)c_t^* + \psi + \left[\alpha^*(r_{t+1}^e - r^f) + r^f + \frac{1}{2}\alpha^*(1 - \alpha^*)\omega^2\right], \quad (28)$$

and the estimated state \hat{a}_t is characterized by the following Kalman filtering equation

$$\hat{a}_{t+1} = (1 - \theta)\hat{a}_t + \theta(a_{t+1} + \xi_{t+1}), \quad (29)$$

where $\psi = \log \beta - (1 - 1/\beta) \log(1 - \beta)$,²² $\theta = 1 - \frac{1}{\exp(2\kappa)}$ is the optimal weight on observation, ξ_t are the iid idiosyncratic noise with $\omega_\xi^2 = \text{var}[\xi_{t+1}] = \Sigma/\theta$, and $\Sigma = \frac{\alpha^{*2}\omega^2}{\exp(2\kappa) - (1/\beta)^2}$ is the steady state conditional variance. Finally, the change in individual consumption is

$$\Delta c_{t+1}^* = \theta \left[\frac{\alpha^* u_{t+1}}{1 - ((1 - \theta)/\beta) \cdot L} + \left(\xi_{t+1} - \frac{(\theta/\beta)\xi_t}{1 - ((1 - \theta)/\beta) \cdot L} \right) \right], \quad (30)$$

Proof. *The proof is straightforward. ■*

It is clear from this proposition that optimal consumption and portfolio rules are *interdependent* in the presence of RI. Expression (26) shows that although the optimal fraction of savings invested in the risky asset is proportional to the risk premium $(\mu - r^f + 0.5\omega^2)$ and the reciprocal of both the coefficient of relative risk aversion (γ) and the variance of the unexpected component in the risky asset (ω^2) as predicted by the standard Merton solution, it also depends on the long-term consumption risk measured by ς . Note that (26) can be rewritten as

$$\alpha^* = \frac{\mu - r^f + 0.5\omega^2}{\tilde{\gamma}\omega^2}, \quad (31)$$

²² Here we use the facts that $\beta = 1 - \exp(c - a)$ and $c - a = E[c_t - a_t]$ is the steady state value.

where the effective coefficient of relative risk aversion is $\tilde{\gamma} = \gamma\varsigma > \gamma$. Hence, even if the true γ is close to 1 as assumed at the beginning of this section, the effective risk aversion that matters for the optimal asset allocation is ς which could be much greater than 1 if the capacity is low. Therefore, both the degree of attention (θ) and the discount factor (β) amount to an increase in the effective coefficient of relative risk aversion. Holding β constant, the larger the degree of attention, the less the ultimate consumption risk. As a result, investors with low attention will choose to invest less in the risky asset. For example, with RI, a 1% percent negative shock in investors' financial wealth would affect their consumption more than that predicted by the standard RE model. Therefore, investors with finite capacity are less willing to invest in the risky asset. It is also worth noting that when the information capacity becomes infinitely large (i.e., $\kappa = \infty$ and $\theta = 1$), the RI model reduces to the standard RE model in which consumption reacts instantaneously and completely to the innovations and the optimal share in the risky asset is the same as the myopic solution.

Equations (27) and (30) show that individual consumption under RI reacts not only to fundamental shocks (u_{t+1}) but also to its own endogenous noises (ξ_{t+1}) induced by finite capacity. The endogenous noise can be regarded as a type of “consumption shock” or “demand shock”. In the intertemporal consumption literature, some ad hoc transitory consumption shocks are modeled to make the model fit the data better. Under RI, the idiosyncratic noise due to RI could endogenize the exogenous ad hoc transitory consumption that lacks economic interpretations. Furthermore, equation (30) also makes it clear that consumption growth adjusts slowly and incompletely to the innovations to asset returns and reacts quickly to the idiosyncratic noise.

It is worth noting that the main difference between Gabaix and Laibson's *6D infrequent-adjustment* model and our RI model is that in their model, investors adjust their consumption plans *infrequently but completely* once they choose to adjust, whereas investors with finite capacity adjust their plans *frequently but incompletely* in every period. In addition, in the *6D* model, the optimal fraction of savings invested in the risky asset is assumed to be *fixed* at the standard Merton solution, whereas optimal portfolio choice under RI reflects the larger long-term consumption risk caused by slow adjustments and thus the share invested in the risky asset is less than the Merton solution.

2.4 Some Implications for Long-term Investment

2.4.1 Investment Horizon Does Matter

In the standard RE portfolio choice model, the investment horizon is *irrelevant* for investors who have power utility functions, have only financial wealth and who face constant investment opportunities. By contrast, it is apparent from proposition 1 that with RI the investment horizon does matter for optimal asset allocations because it affects the long-term consumption risk. As argued in Campbell and Viceira (2002), we may vary the effective investment horizon by varying the discount factor (β) that determines the relative weights investors put on the near future versus the distant future. The investors with large β would place a relatively high weight on the distant future, while those with small β place more weight on the near future. Hence, expression (26) clearly shows that the higher the value of β (the longer the investment horizon), the more fraction of financial wealth would be invested in the risky asset. Figure 2 illustrates how the investment horizon affects optimal asset allocation for given θ . The intuition is that those inattentive investors with larger β face smaller long-term consumption risk; consequently they choose to invest more wealth in the risky asset.

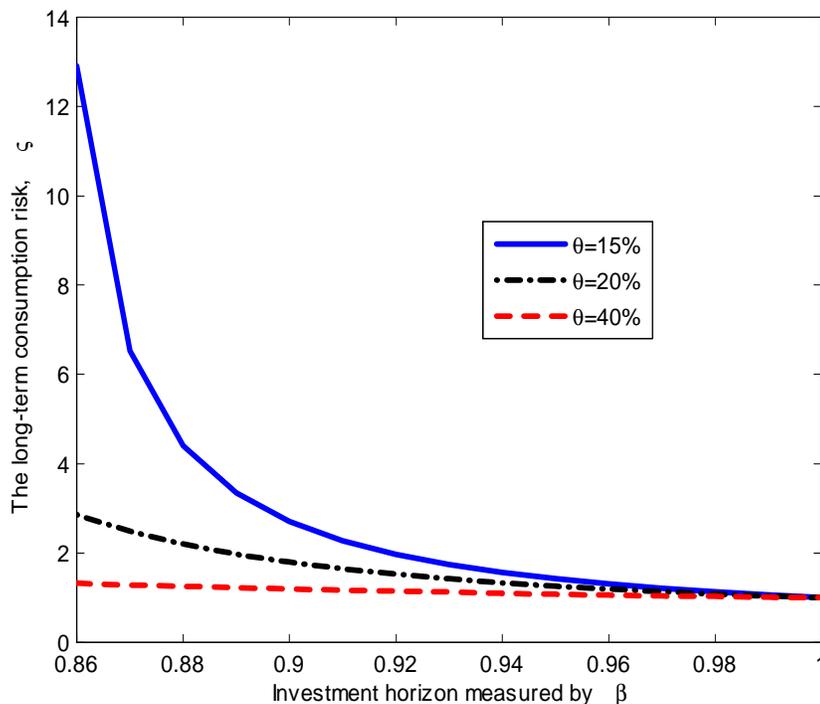


Figure 2: Effects of investment horizon on long-run risk.

Furthermore, this prediction can also be used to explain the investment behavior of the young and old. Financial advisors typically recommend that people shift investments away from the risky asset to the risk-free asset as they age. Some theoretical models have been proposed to evaluate this justification; for example, Jagannathan and Kocherlakota (1996), Viceira (2001), and Campbell and Viceira (2002). In the existing literature, the investment horizon and age are generally closely related. In other words, the old have low *effective* time discount factor because they face high death probability in every period. Hence, RI might provide an economic reasoning for this popular advice, and we can address the model’s predictions for the behavior of the young and old by varying the value of β . Specifically, the old people with low β face large ultimate consumption risk due to slow adjustment; consequently, it would be better for them to invest less in the stock market.

2.4.2 Implications for Limited Stock Market Participation

In the benchmark model, we assumed that every investor has the same level of channel capacity. However, in reality, investors devote different levels of channel capacity to observing and processing information. Hence, in the heterogeneous-attention case, investors have different demands for the risky asset even if they have the same preference and face the common shock to the asset return. Based on (26), it is straightforward to show that “limited stock market participation” can arise endogenously in the RI model.²³ The intuition is that the demand for the equity decreases with the degree of attention (θ). Therefore, the investor whose degree of attention is equal to or lower than the critical value $1 - \beta$,²⁴ will choose not to participate in the stock market because he faces extremely high ultimate consumption risk. Note that because our model abstracts from transaction costs, taxes, and short-sale constraints, among others, it may be very likely that those investors with low capacity (but still greater than the critical value) would also choose not to participate in the stock market if there are other frictions in the model. We do not explore this issue here and leave it for future research. Hence, the model composed of investors with different degrees of RI can generate “limited stock market participation” endogenously. Those investors who know that they cannot devote enough capacity to processing information about the evolution of their financial wealth hit by the aggregate shock choose not to invest in stock market because they face larger consumption risk.

²³Although there is an upward trend in stock market participation in the PSID, the participation rate in the stock market including both direct and indirect investment is still less than 50%. See Vissing-Jørgensen (2002) for a detailed discussion about limited asset market participation.

²⁴Note that it is equivalent to say that κ is equal or lower than $\frac{1}{2} \log(1/\beta)$.

3 Implications for Aggregate Dynamics and Asset Returns

3.1 The Joint Dynamics of Aggregate Consumption and the Equity Return

In this subsection, we provide a complete characterization of the joint dynamics of aggregate consumption and the asset return under RI. Specifically, we analyze several important stochastic properties of the joint behavior of financial markets and the real economy: the smoothness of consumption growth, low contemporaneous correlation and covariance between consumption growth and asset returns, positive autocorrelation of consumption growth, and non-zero covariance of consumption growth and lagged equity returns.

The endogenous noise due to finite capacity, ξ_t , is idiosyncratic because it arises from individuals' own internal information processing channels. Consequently, it can be cancelled out after aggregating across all investors.²⁵ Given that the explicit expression of individual consumption, (30), is in exact aggregation form, aggregating individual consumption growth across all investors yields an expression for the change in aggregate consumption. For simplicity, here we assume that individuals have identical channel capacity, θ . Of course, allowing for heterogeneity in θ can generate more realistic portfolio choices. However, this complication does not significantly affect the main model predictions on the joint dynamics of aggregate consumption and the equity return.²⁶ Denoting aggregate consumption under RI as c_t ,²⁷ the following proposition summarizes the implications of RI for the joint dynamics of aggregate consumption and the asset return:

Proposition 2 *Consider the optimal consumption and investment decisions (26)-(30). Suppose that γ is close to 1, and $\theta > 1 - \beta$. We then obtain the following expression for aggregate consumption growth under RI:*

$$\Delta c_{t+1}^* = \frac{\theta \alpha^* u_{t+1}}{1 - ((1 - \theta) / \beta) \cdot \bar{L}}, \quad (32)$$

where θ is the average degree of inattention in the model economy and α^* is the average optimal

²⁵Note 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. The reason is that the endogenous noises and the innovations to equity returns are uncorrelated.

²⁶See Appendix 6.3 for detailed derivations of the model's predictions on optimal portfolio choice and the joint dynamics of aggregate consumption and the equity return when θ follows a normal distribution and a Bernoulli distribution.

²⁷Since we focus on aggregate behavior and to avoid the notation confusion, in the following equation, we still use c to represent aggregate consumption.

share invested in the risky asset and is given by (26).²⁸ This $MA(\infty)$ process implies that: (1) the standard deviation of consumption growth is

$$\text{sd} [\Delta c_{t+1}^*] = \lambda \alpha \omega, \quad (33)$$

where the excess smoothness ratio $\lambda = (1 - (1 - \theta)/\beta) / \sqrt{1 - ((1 - \theta)/\beta)^2}$, (2) the correlation between consumption growth and equity return is

$$\text{corr} [\Delta c_{t+1}^*, r_{t+1}^e] = \sqrt{1 - ((1 - \theta)/\beta)^2}, \quad (34)$$

(3) the covariance between aggregate consumption growth and the asset return is

$$\text{covar} [\Delta c_{t+1}^*, r_{t+1}^e] = \theta \alpha^* \omega^2, \quad (35)$$

(4) the first-order autocorrelation of consumption growth is

$$\rho_{\Delta c(1)} = \text{corr} [\Delta c_t^*, \Delta c_{t+1}^*] = (1 - \theta) / \beta, \quad (36)$$

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

$$\text{covar} [\Delta c_{t+1}^*, r_{t+1-j}^e] = \theta \alpha^* ((1 - \theta)/\beta)^j \omega^2, \quad (37)$$

where $j \geq 1$.

Proof. Using (32), it is straightforward to obtain the above results. ■

Expression (32) clearly shows that aggregate consumption adjusts gradually to the shock to the equity return; RI affects the amplification and propagation mechanism of the model by two channels: (1) reducing the optimal share invested in the equity return α^* , as $\alpha^* = \frac{1-(1-\theta)/\beta}{\theta} \alpha$, and (2) introducing a slow response of consumption to the equity return due to finite channel capacity, $\frac{\theta}{1-((1-\theta)/\beta)L}$. Expression (32) also implies that to guarantee the existence of solution, $1 - (1 - \theta)/\beta$ must be greater than 0; note that this condition also guarantees that the optimal share in the risky asset, α^* , should be always greater than 0 as $\frac{1}{\zeta} = \frac{1-(1-\theta)/\beta}{\theta} > 0$.

Since $\beta, \theta \in (0, 1)$, it is clear from (33), (34), (35), (36), and (37) that RI can bring the model and the data closer along the following dimensions: (i) RI reduces the relative volatility

²⁸ That is, $\alpha^* = \frac{1-(1-\theta)/\beta}{\theta} \alpha$; as given in (7), $\alpha = \frac{\mu - r^f + 0.5\omega^2}{\omega^2}$ is the optimal share invested in the stock market in the RE model.

of consumption to the equity return, (ii) it reduces the correlation and covariance between consumption and the equity return, (iii) it generates positive serial correlation in consumption growth, and (iv) it generates positive covariances between consumption growth and lagged equity returns. Note that as shown in Section 2.1, the standard RE model fails to match the data well in these aspects.

Figure 3 illustrates how the combination of RI (θ) and the discount factor (β) affects the smoothness ratio of aggregate consumption (λ). It is clear that the relative volatility of consumption growth decreases with the degree of attention (θ) given β . Furthermore, it is also clear from (33) that RI reduces the relative volatility by two channels: (i) reducing the optimal share invested in the equity return $\left(\frac{1-(1-\theta)/\beta}{\theta}\right)$ and (ii) the gradual response of consumption to the equity return due to finite channel capacity $\left(\sqrt{\frac{\theta^2}{1-((1-\theta)/\beta)^2}}\right)$.²⁹ Using (34), by simple calculation we obtain that $\frac{\partial \text{corr}}{\partial \theta} > 0$, that is, given β , the correlation between aggregate consumption and the equity return (corr) increases with θ .

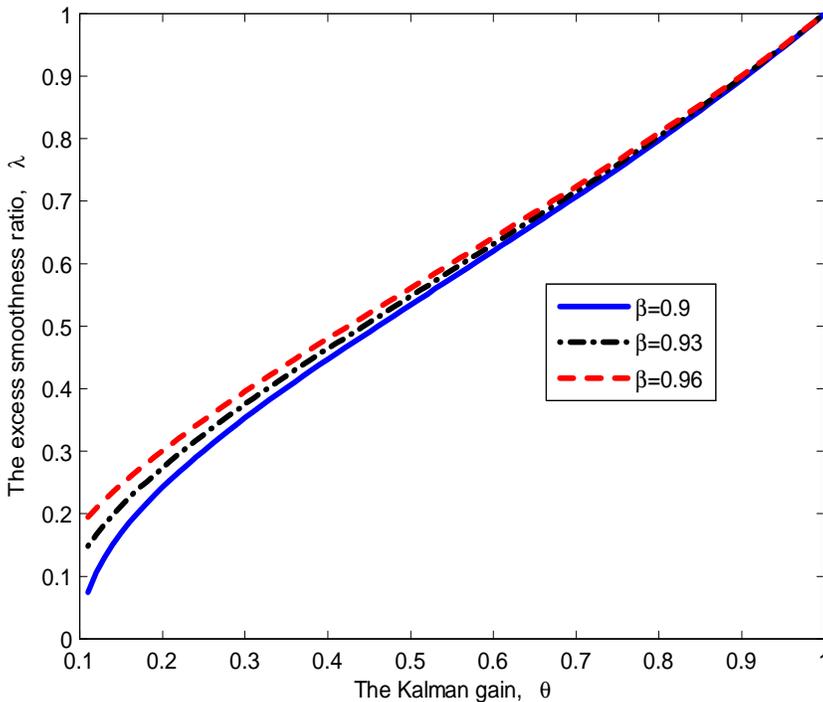


Figure 3: Effects of RI on the excess smoothness ratio.

²⁹It is worth noting that (33) also implies that RI has a potential to explain the equity volatility puzzle because $\lambda < 1$; that is, the same volatility of consumption growth implies higher volatility of the equity in the presence of RI. See Campbell (2003) for a detailed discussion for this puzzle.

As documented in Gabaix and Laibson (2001), the empirical covariances of aggregate consumption and the equity return gradually increase with the horizon, j , in the U.S. data. Expression (37) can thus correctly predict the observation that the covariance slowly rises over time, that is, aggregate consumption reacts to the equity return gradually and with delay. The intuition here is simple: If a large number of consumers/investors in the economy can not perfectly digest the aggregate innovation to their financial wealth due to limited information capacity constraints, aggregate consumption should react to the shock to the equity return with delay and be sensitive to lagged shocks.

Lynch (1996) solved a discrete-time infrequent-adjustment model numerically, and demonstrated that infrequent adjustment has a potential to generate realistic joint dynamics of aggregate consumption and the asset return. However, the aggregation mechanism in his model is distinct from that in the RI model. Specifically, in the infrequent-adjustment model, aggregation leads to aggregate consumption smoothness because of the synchronization problem caused by infrequent adjustments across individual investors, whereas the RI model generates smooth consumption because the idiosyncratic noises due to finite idiosyncratic channel capacity are cancelled out after aggregating across all investors.

To evaluate our model quantitatively, we calibrate θ first and then use the calibrated θ to compute the model's predictions on the joint dynamics of consumption and the equity return. Specifically, we start with the annualized U.S. quarterly data in Campbell (2003), and assume that $\omega = 0.16$, $\pi = \mu - r^f = 0.06$, $\beta = 0.907$, and $\gamma = 1$. We then calibrate θ to match the observed $\alpha = 22\%$ estimated in Section 5.1 of Gabaix and Laibson (1999) to:³⁰

$$\alpha^* = \frac{1 - (1 - \theta)/\beta}{\theta} \frac{\pi + 0.5\omega^2}{\omega^2} = 0.22,$$

which means that $\theta = 10\%$, that is, $\kappa = 0.08$ bits. (It also means that approximately 10% of the uncertainty is removed upon receiving a new signal about the aggregate shock to the equity return.) Note that if $\gamma = 1$, the RE version of the model generates a highly unrealistic share invested in the stock market: $\alpha = \frac{\pi + 0.5\omega^2}{\omega^2} = 2.84$. As shown in Section 2.1, to match the observed fraction in the U.S. economy (22%), γ must be set to 13.

It is worth noting that even 1 bit of information transmitted is a very low number and is well below the total information-processing ability of human beings. However, it is not implausible for ordinary investors/consumers because they also face many other competing demands on capacity.

³⁰They assume that all capital is identical to stock market capital, and capital income accounts for 1/3 of total income.

For an extreme case, a young worker who accumulates balances in his 401 (k) retirement savings account might pay no attention to the behavior of the stock market until he retires. In addition, in our model for simplicity we only consider an aggregate shock from the equity return, while in reality consumers/investors face substantial idiosyncratic shocks (e.g., labor income shock) that we do not model in this paper. In a recent paper by Maćkowiak and Wiederholt (2009), they showed that given that the total information flow equals 133 bits, the decision-maker of the typical firm only allocates 0.76 bits of information flow to tracking aggregate technology, and 0.41 bits to tracking monetary policy.³¹ Therefore, the exogenous capacity given in our model can be regarded as a shortcut to small fractions of consumers' total capacity used to monitor their financial wealth hit by the aggregate shock.

In addition, low capacity devoted to processing macroeconomic information could be rationalized by examining the welfare effects of finite channel capacity.³² Following Barro (2007), we calculated the welfare losses due to RI in the benchmark model. Specifically, we first define the marginal welfare costs (mwc) due to RI as follows:

$$\text{mwc} = \frac{\partial \hat{v}(\hat{a}_0) / \partial \theta}{(\partial \hat{v}(\hat{a}_0) / \partial \hat{a}_0) \hat{a}_0}, \quad (38)$$

where

$$\begin{aligned} \hat{v}(\hat{a}_0) = & -\frac{\gamma-1}{2(1-\beta)} \hat{a}_0^2 - \left(\frac{\gamma-1}{1-\beta} \log(1-\beta) - \frac{1}{1-\beta} \right) \hat{a}_0 \\ & - \frac{\gamma-1}{2(1-\beta)} \left(\left(\log(1-\beta) - \frac{1}{\gamma-1} \right)^2 + \frac{\theta \beta^3 \alpha^{*2} \omega^2}{(1-\beta)(\beta^2 + \theta - 1)} \right), \end{aligned}$$

is the value function under RI, $\alpha^* = \frac{1-(1-\theta)/\beta}{\theta} \alpha = \frac{1-(1-\theta)/\beta}{\theta} \frac{\pi + 0.5\omega^2}{\omega^2}$, and $\frac{\partial \hat{v}}{\partial \theta}$ is evaluated for given \hat{a}_0 . The welfare costs due to RI are compared with that from a small proportional change in the initial level of the perceived wealth \hat{a}_0 . Using the parameter values set in the paper: $\beta = 0.91$, $\omega = 16\%$, $\pi = 0.06$, and the perceived initial asset holding is set to $\hat{a}_0 = \ln \hat{A}_0 = \ln(10000)$, it is straightforward to calculate that $\text{mwc} = -3.621 \times 10^{-4}$ if $\theta = 10\%$ and $\gamma = 1.001$. That is, to maintain the lifetime expected utility, an increase in θ by, say, 100% (from 10% to 20%) requires

³¹There are some existing estimation and calibration results in the literature. For example, Adam (2005) found $\theta = 40\%$ or $\kappa = 0.255$ bits based on the response of aggregate output to monetary policy shocks. In addition, Luo (2008) found that if $\theta = 50\%$, the otherwise standard permanent income model can generate realistic relative volatility of consumption to labor income; Luo and Young (2009) found that setting $\theta = 57\%$ allows a otherwise standard RBC model to match the post-war U.S. consumption/output volatility.

³²Pischke (1995) solved a permanent income model in which consumers ignore macroeconomic information, and found that the utility loss of a consumer with a coefficient of relative risk aversion $\gamma = 2$ due to this information imperfection only amounts to \$1.5.

a reduction in the initial level of the perceived wealth by approximately 0.03621%. Similarly, $mwc = -1.4484 \times 10^{-5}$ if $\theta = 50\%$ and $\gamma = 1.001$. That is, an increase in θ from 50% to 100% requires a reduction in \hat{a}_0 by approximately 0.0145%. (i.e., \$1.45 if $\hat{A}_0 = \$10000$.) Hence, for plausible assumptions, the welfare losses due to RI are not significant.³³ This conclusion provides some evidence that it is reasonable for ordinary consumers/investors to devote low channel capacity to observing and processing aggregate information because the welfare improvement from adopting the first-best decision rule is not significant.

In the next step, using the calibrated θ (10%), we examine how RI affects the joint dynamics of aggregate consumption and equity returns quantitatively. In particular, we examine the model's quantitative predictions on the dimensions we discussed above:

- (i) Using (33), the relative volatility of aggregate consumption to the equity return (rvc) can be written as

$$rvc = \frac{sd[\Delta c_{t+1}^*]}{sd[r_{t+1}^e]} = \sqrt{\frac{\theta^2 \alpha^{*2}}{1 - ((1 - \theta)/\beta)^2}}, \quad (39)$$

which means that RI can reduce rvc by:

$$\mu_1 = \frac{rvc^{RE}}{rvc^{RI}} = \frac{\sqrt{1 - ((1 - \theta)/\beta)^2}}{1 - (1 - \theta)/\beta}, \quad (40)$$

where rvc^{RE} and rvc^{RI} are the relative volatility of aggregate consumption, respectively. As shown in 2.1, the RE case (i.e., $\theta = 100\%$) generates $rvc^{RE} (= \alpha^*) = 2.84$, which is well above its empirical counterpart in the U.S. data: $1/16 \simeq 6\%$. By contrast, using our calibration choice ($\theta = 10\%$), the RI model implies that $rvc^{RI} = 4\%$, which is much closer to the empirical counterpart.

- (ii) Using $\theta = 10\%$ and $\beta = 0.9$, (34) implies that $\text{corr}[\Delta c_{t+1}^*, r_{t+1}^e] = 0.124$. Therefore, RI can greatly reduce the correlation between Δc_{t+1}^* and r_{t+1}^e and thus improves the model's prediction in this aspect. Note that in the RE model, the correlation is 1, while the empirical counterpart is about 0.34 at a 1-year horizon.

- (iii) Expression (35) implies that RI can reduce the covariance by a factor:

$$\mu_2 = \frac{\text{covar}^{RE}}{\text{covar}^{RI}} = \frac{1}{1 - (1 - \theta)/\beta}, \quad (41)$$

³³This conclusion is also consistent with that in Pischke (1995).

which equals 130 when $\theta = 10\%$, i.e., the covariance can be reduced by 86 times. Note that in the RE case (i.e., $\theta = 100\%$), $\text{covar}^{RE} = \alpha\omega^2 = 2.84 \cdot 0.16^2 = 7.3 \cdot 10^{-2}$, which obviously contradicts its empirical counterpart in the U.S. data: $6 \cdot 10^{-4}$. By contrast, when $\theta = 10\%$, the covariance becomes $5.6 \cdot 10^{-4}$ and is therefore much closer to the empirical counterpart.

- (iv) As reported in Piazzesi (2001) and Campbell (2003), $\rho_{\Delta c(1)}$ is about 0.2 in the U.S. annual data.³⁴ However, the RE model predicts that the autocorrelation of consumption growth is 0, that is, consumption growth is iid. Under our calibration choice ($\theta = 10\%$), (36) implies that $\rho_{\Delta c(1)} = 0.99$. Although the theoretical autocorrelation is significantly lower than the empirical counterpart, it still improves the model's prediction as it is roughly consistent with some empirical evidence that the first-order autocorrelation is significant (i.e., consumption growth is definitely not iid).
- (v) Note that under RE, there is no correlation between Δc_{t+1}^* and r_{t+1-j}^e as Δc_{t+1}^* is iid, and the covariance should initially jump to a plateau and stay there, which definitely contradicts the evidence reported in Gabaix and Laibson (2001) that the empirical covariances of aggregate consumption and the equity return gradually increase with the horizon, j . When $j = 1$ and $\theta = 10\%$, (37) implies that $\text{covar}[\Delta c_{t+1}^*, r_t^e] = 5.6 \cdot 10^{-4}$, which is much closer to the empirical covariance $2 \cdot 10^{-4}$.

Therefore, our above calculations clearly show that a low value of θ or κ is required to roughly capture some important stochastic properties of the joint dynamics of aggregate consumption and the equity return in the US data.

3.2 The Implications for the Equity Premium

In the benchmark model, because every investor is assumed to have the same degree of RI, the following pricing equation linking aggregate consumption growth and the equity premium holds when $\gamma \simeq 1$:

$$\mu - r^f + 0.5\omega^2 = \zeta\alpha^*\omega^2. \quad (42)$$

Suppose, as in the consumption-based CAPM literature, that the risk-free asset is an inside bond, so that in equilibrium the net supply of the risk free asset is 0 and then the share of the

³⁴Mehra and Prescott (1985) also found some evidence for modest and persistent processes of consumption growth.

risky asset in financial wealth (α^*) is 100%. Accordingly, (42) becomes

$$\mu - r^f + 0.5\omega^2 = \varsigma\omega^2.$$

Hence, RI implies higher equity premium because the ultimate consumption risk $\varsigma = \frac{\theta}{1-(1-\theta)/\beta} > 1$.³⁵ The intuition behind this result is that investors facing higher ultimate consumption risk caused by finite capacity choose to invest less in the stock market and thus require higher risk compensation in equilibrium.

For illustrative purposes, we consider the following question: What will an economist equipped with the aggregate consumption-based 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

$$\tilde{\gamma} = \frac{\mu - r^f + \frac{1}{2}\omega^2}{\text{covar}[c_{t+1} - c_t, r_{t+1}^e]} = \frac{\mu - r^f + 0.5\omega^2}{\theta\alpha^*\omega^2} > \gamma \simeq 1,$$

where $\tilde{\gamma}$ is the estimated coefficient of relative risk aversion and γ is the true value that is close to 1. The intuition is quite simple. The RI model can generate a very low value of $\text{covar}[c_{t+1} - c_t, r_{t+1}^e]$ and a high equity premium simultaneously because the equity premium is determined by the long-term consumption risk, $\lim_{S \rightarrow \infty} \text{covar}[c_{t+1+S} - c_t, r_{t+1}^e]$. In other words, the estimate of the coefficient of relative risk aversion is biased upward by a factor measured by the long-term consumption risk. For example, if the true γ is 1 and $1 - (1 - \theta)/\beta = 0.2$, the economist will find that the estimated value $\tilde{\gamma}$ will be 5. Hence, the estimated high coefficient 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 aggregate consumption and asset prices data.

4 Review of Related Empirical Evidence

Existing survey evidence supports that: (1) a significant fraction of investors do not have enough knowledge about the evolution of their financial wealth and consequently cannot adjust their consumption fully in response to the innovations to the returns, and (2) the innovations to their financial assets can be used to predict their future change in consumption. For example, Ameriks

³⁵By contrast, without RI the pricing equation implies that in equilibrium ($\alpha = 1$) the equity premium is lower because $\mu - r^f + 0.5\omega^2 = \omega^2$.

and Zeldes (2006) used pooled cross-sectional data from Surveys of Consumer Finance (SCF) to examine the dynamics of household portfolio shares. They found some important features of household portfolio behavior; for example, significant non-stockownership, and wide-ranging heterogeneity in asset allocation. As we discussed in Section 2, the optimal portfolio solution under RI, (26), can provide a partial explanation for these features because some investors with low attention face large ultimate consumption risk and thus choose not to invest in the stock market. Further, even if all investors have the same level of risk aversion, they may still choose different investment strategies due to the possible heterogeneity in inattention. In other words, RI introduces another type of heterogeneity.

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.³⁶ In the same paper, they also reported that for stockholders with more than \$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 10 th month to the 18 th month, another quarter of which occurs from the 19 th month to the 27 th month, and the rest occurs from the 28 th month to the 36 th month. This evidence can be largely captured by the RI model because (32) implies that the long-term consumption growth under RI is

$$c_{t+1+S} - c_t = \theta \left[1 + (1 - \theta) / \beta + \dots + ((1 - \theta) / \beta)^S \right] u_{t+1},$$

which implies that $\lim_{S \rightarrow \infty} (c_{t+1+S} - c_t) = \frac{\theta u_{t+1}}{1 - (1 - \theta) / \beta}$. Consider a numerical example (the time unit here is 3 quarters) in which β is set to 0.9 used in the calibration exercise in the last section. When $\theta = 33\%$, i.e., if on average the investors in the sample remove 33% of the uncertainty of the asset returns after observing the new signals about the returns, the long-term impact of the returns on consumption growth is $\lim_{S \rightarrow \infty} (c_{t+1+S} - c_t) = 1.03$, which is exactly the same as that estimated from the data. When $S = 0$, $c_{t+1+S} - c_t = \theta = 33\%$, as corresponds to the estimated 33%; when $S = 1$, $c_{t+1+S} - c_t = \theta [1 + (1 - \theta) / \beta] = 59\%$, which is also close to the estimated impact 66%. When β is set to a more realistic value, 0.98, and $\theta = 40\%$, the long-term impact of the returns on consumption growth is $\lim_{S \rightarrow \infty} (c_{t+1+S} - c_t) = 1.03$, which is exactly the same as that estimated from the data. When $S = 0$, $c_{t+1+S} - c_t = \theta = 40\%$, as corresponds to the

³⁶Kennickell, et. al (2000) and Starr-McCluer (2001) also reported similar results based on alternative survey sources.

estimated 33%; when $S = 1$, $c_{t+1+S} - c_t = \theta [1 + (1 - \theta) / \beta] = 64\%$, which is also close to the estimated impact 66%.³⁷ Thus, our numerical example can generate results similar to those estimated from the US data. Furthermore, using the estimation results from Dynan and Maki (2001), we plot Figure 4 to illustrate the extent to which 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 $\theta = 24\%$ (or $\kappa = 0.2$ bits) can fit the empirical results quite well: the responses of consumption to the innovations is muted initially and then increases 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.

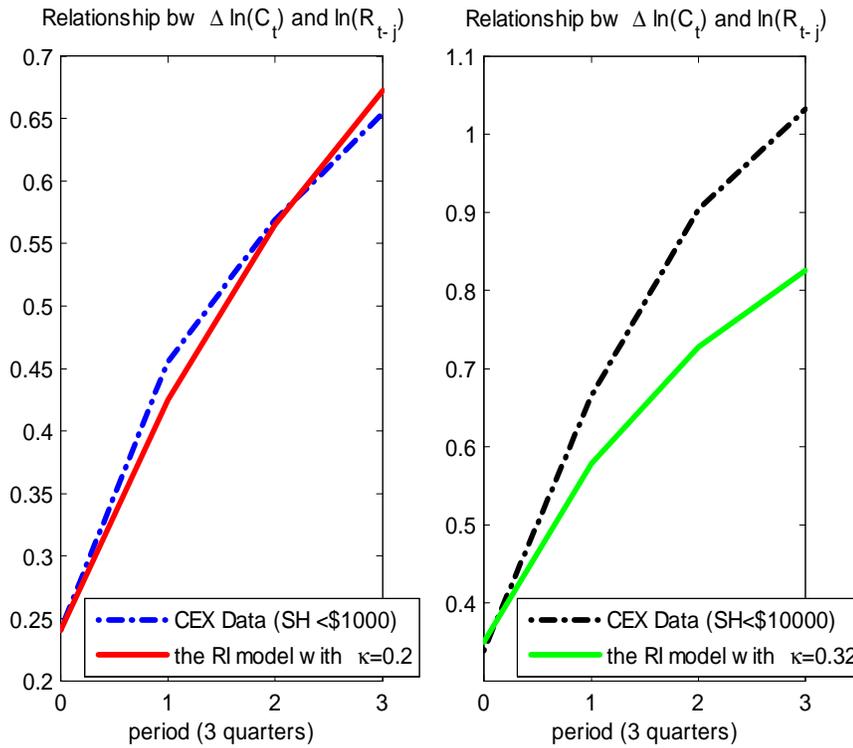


Figure 4: Responses of consumption to the innovations in the CEX.

³⁷Note that in the data, *one third* of the impact increases during the first 9 months and *another third* of which occurs from the 10th month to the 18th month.

5 Conclusion

Ordinary investors do not have infinite information-processing capacity; instead they only have finite capacity when processing available financial information. Rational inattention, first proposed by Sims (2003), has recently been introduced into economics and finance to model this type of information constraints. This paper takes such RI into account in an otherwise standard portfolio choice model and explores its effects on optimal consumption and portfolio rules and the joint dynamics of aggregate consumption and the equity return.

The first contribution of this paper is that we explicitly solve an RI version of the intertemporal portfolio choice model and show that optimal consumption and portfolio rules under RI are interdependent. Second, we show that in the presence of RI, we need to use the ultimate consumption risk rather than the contemporaneous risk to measure the true riskiness of the risky asset because investors adjust their optimal plans slowly and incompletely. Holding the discount factor constant, the larger the degree of inattention, the greater the ultimate consumption risk. Consequently, RI reduces the demand for the risky asset. Hence, under RI, investment horizon does matter for long-term investment and RI could be an alternative explanation for significant non-stockownership. Finally, the explicit solution of consumption and asset holdings is in exact aggregation form. Aggregating across investors therefore depicts a joint dynamics of aggregate consumption and asset returns. We show that introducing RI could reconcile the following observations in the US data simultaneously: (i) smooth aggregate consumption, (ii) the low contemporaneous correlation between aggregate consumption and asset returns, and (iii) the implied high equity premium.

A promising extension is to introduce the risk-sensitivity preference into the benchmark model in this paper. Risk-sensitivity is a special case of Epstein-Zin's recursive utility and can be used to capture a preference for robustness: concern about model misspecification. Hansen, Sargent, and Tallarini (1999) examined how risk-sensitivity alters the choices of consumption, investment, and asset prices in a permanent income model. The most promising part of the Merton model with robustness and rational inattention is that within this framework the two informational frictions have the potential to reduce the optimal share invested in the risky asset, while maintaining the tractability of the LQ setting that rationalizes the optimality of ex post Gaussian distributions due to RI. Another interesting direction for future research is to model multiple risky assets in the RI model and examine whether RI can help resolve the "asset allocation puzzle". That is, the ratio of risky bonds to equities in the optimal portfolio increases with the coefficient of relative risk aversion, which is consistent with conventional

portfolio advice but is inconsistent with static mean-variance analysis.

6 Appendix

6.1 Deriving the Expression of Consumption Growth, (20)

Combining (5), (17), with (19) yields

$$\widehat{a}_{t+1} - a_{t+1} = \frac{1}{\beta}(1 - \theta)(\widehat{a}_t - a_t) - (1 - \theta)\alpha u_{t+1} + \theta \xi_{t+1},$$

which implies that

$$\widehat{a}_{t+1} - a_{t+1} = \frac{-(1 - \theta)\alpha u_{t+1} + \theta \xi_{t+1}}{1 - ((1 - \theta)/\beta) \cdot L}.$$

Furthermore, the change in the perceived state can be written as:

$$\begin{aligned} \Delta \widehat{a}_{t+1} &= \theta(a_{t+1} - \widehat{a}_t) + \theta \xi_{t+1} \\ &= \theta \left[\alpha u_{t+1} - \frac{1}{\beta}(\widehat{a}_t - a_t) \right] + \theta \xi_{t+1} \\ &= \theta \left[\left(\frac{\alpha u_{t+1}}{1 - ((1 - \theta)/\beta) \cdot L} \right) + \left(\xi_{t+1} - \frac{(\theta/\beta) \xi_t}{1 - ((1 - \theta)/\beta) \cdot L} \right) \right], \end{aligned}$$

which leads to (20) in the text.

6.2 Deriving the Long-term Euler Equation, (21)

When wealth is allocated efficiently across the two financial assets, the marginal investment in any asset yields the same expected increase in future utility:

$$E_t \left[u'(c_{t+1}) \left(R_{t+1}^e - R^f \right) \right] = 0. \quad (43)$$

Using the Euler equation for the risk free asset between $t + 1$ and $t + 1 + S$, we have

$$u'(c_{t+1}) = E_{t+1} \left[(\beta R_f)^S u'(c_{t+1+S}) \right],$$

which can be substituted into (43) to get the long-term Euler equation:

$$\begin{aligned} &E_t \left[E_{t+1} \left[(\beta R_f)^S u'(c_{t+1+S}) \right] \left(R_{t+1}^e - R^f \right) \right] \\ &= E_t \left[(\beta R_f)^S u'(c_{t+1+S}) \left(R_{t+1}^e - R^f \right) \right] = 0, \end{aligned}$$

which implies

$$E_t \left[u' (c_{t+1+S}) \left(R_{t+1}^e - R^f \right) \right] = 0,$$

which can be transformed to Equation (21) in the text by dividing $u' (c_t)$ on both sides. Note that the expectation $E_t [\cdot]$ is conditional on the entire history of the economy up to t , and S is infinite over which consumption response under RI is studied.

6.3 Model's Predictions when θ is heterogenous

Case 1 (normal distribution) If θ follows a normal distribution:

$$\theta_i \sim N(\theta, \varpi^2),$$

the aggregate optimal allocation in the stock market is

$$\alpha^* = \left[\int_{\theta_{\min}}^1 f(\theta_i) \frac{1 - (1 - \theta_i)/\beta}{\theta_i} d\theta_i \right] \frac{\mu - r^f + 0.5\omega^2}{\gamma\omega^2},$$

where $\theta_{\min} = 1 - \beta$ and $f(\theta_i)$ is the normal pdf of the degree of attention, θ_i . It is straightforward to calculate that the assumption of heterogeneity does not affect α^* significantly. For example, when $\theta = 0.2$, $\varpi = 0.01$, and $\beta = 0.9$, $\alpha^* = 0.55$, which is close to 0.56 obtained from the case with the identical $\theta = 0.2$; when ϖ increases to 0.05, $\alpha^* = 0.51$.

Case 2 (Bernoulli distribution) Another interesting example is that channel capacity (θ) follows a *Bernoulli* distribution. Specifically, we assume that there are only two types of consumers in the economy: type-*I* investors have the identical value of attention $\underline{\theta} < \theta_{\min}$ and type-*II* investors have the identical value of attention $\bar{\theta} > \theta_{\min}$. Drawing investors from the population follows a *Bernoulli* distribution: the probabilities of being type-*I* and type-*II* are $1 - p$ and p , respectively. Given that there are a continuum of agents in the economy, in every period the fractions of type-*I* and type-*II* agents are $1 - p$ and p , respectively. It is obvious that type-*I* investors do not invest in the stock market as they face infinite long-run consumption risk. Hence, the *average* optimal allocation in the stock market is

$$\alpha^* = p \frac{1 - (1 - \bar{\theta})/\beta}{\bar{\theta}} \frac{\mu - r^f + 0.5\omega^2}{\omega^2}.$$

In this case, p is set to 21% to match the observed stock market participation rate in the U.S.

data. If $\bar{\theta} = 95\%$ and $\beta = 0.90$, it is straightforward to calculate that

$$\alpha^* = 0.209 \frac{\mu - r^f + 0.5\omega^2}{\omega^2}$$

Note that in this case the *average* channel capacity is

$$\theta = p\bar{\theta} + (1 - p)\underline{\theta} = 28\%.$$

References

- [1] Adam, Klaus (2005), “Optimal Monetary Policy in the Presence of Imperfect Common Knowledge,” *Journal of Monetary Economics* 54 (2), 267-301.
- [2] Ameriks, John and Stephen Zeldes (2004), “How Do Household Portfolio Shares Vary with Age?” Columbia University, mimeo.
- [3] 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.
- [4] Barro, Robert (2007), “On the Welfare Costs of Consumption Uncertainty,” mimeo, Harvard University.
- [5] Campbell, John (1993), “Intertemporal Asset Pricing without Consumption Data,” *American Economic Review*, vol. 83, 487-512.
- [6] 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.
- [7] Campbell, John and Luis M. Viceira (2002), “*Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*,” Oxford University Press, UK.
- [8] 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.
- [9] Cover, Thomas M. and Joy A. Thomas (1991), *Elements of Information Theory*, New York: John Wiley and Sons.

- [10] Dynan, Karen and Maki, Dean (2001), “Does Stock Market Wealth Matter for Consumption?” Federal Reserve Board, mimeo.
- [11] Gabaix, Xavier and David Laibson (2001), “The 6D bias and the Equity Premium Puzzle,” NBER *Macroeconomic Annual*.
- [12] Hansen, Lars and Kenneth Singleton (1983), “Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns,” *Journal of Political Economy* 91, 249-265.
- [13] Hansen, Lars and Ravi Jagannathan (1991), “Implications of Security Market Data for Models of Dynamic Economies,” *Journal of Political Economy* 99, 225-62.
- [14] Hansen, Lars P., John C. Heaton and Nan Li (2009), “Consumption Strikes Back? Measuring Long-run Risk,” *Journal of Political Economy* 116(2), 260-302.
- [15] Hansen, Lars Peter and Thomas J. Sargent (2007), *Robustness*, Princeton University Press.
- [16] Hansen, Lars Peter, Thomas J. Sargent, and Thomas D. Tallarini, Jr. (1999), “Robust Permanent Income and Pricing,” *Review of Economic Studies* 66(4), 873-907.
- [17] Huang, Lixing and Hong Liu (2007), “Rational Inattention and Portfolio Selection,” *Journal of Finance* 62, 1999-2040.
- [18] Kennickell, Arthur. B., Martha Starr-McCluer, and Brian J. Surette (2000), “Recent changes in U.S. family finances: Results from the 1998 Survey of Consumer Finances,” Federal Reserve Bulletin, January 1–29.
- [19] Kocherlakota, Narayana (1996), “The Equity Premium: It’s Still a Puzzle,” *Journal of Economic Literature* Vol. XXXIV, 42-71.
- [20] Kocherlakota, Narayana and Ravi Jagannathan (1996), “Why Should Older People Invest Less in Stocks than Younger People?” *Federal Reserve Bank of Minneapolis Quarterly Review*.
- [21] Jermann, Urban J. (1998), “Asset Pricing in Production Economies,” *Journal of Monetary Economics*, 257-275.
- [22] Luo, Yulei (2008), “Consumption Dynamics under Information Processing Constraints,” *Review of Economic Dynamics* 11, 366-385.

- [23] Luo, Yulei and Eric Young (2009), “Rational Inattention and Aggregate Fluctuations,” *The B.E. Journal of Macroeconomics*: Vol. 9: Iss. 1 (Contributions), Article 14.
- [24] Lynch, Anthony (1996), “Decision Frequency and Synchronization across Agents: Implications for Aggregate Consumption and Equity Returns,” *Journal of Finance* 51 (4), 1479-1497.
- [25] Maćkowiak, Bartosz and Mirko Wiederholt (2009), “Business Cycle Dynamics under Rational Inattention,” Working Paper, European Central Bank and Northwestern University.
- [26] Marshall, David and Nayan Parekh (1999), “Can Costs of Consumption Adjustment Explain Asset Pricing Puzzles,” *Journal of Finance* 54(2) 623-654.
- [27] Mehra, Rajnish and Prescott, Edward (1985), “The Equity Premium: A Puzzle,” *Journal of Monetary Economics* 15, 145-161.
- [28] Merton, Robert C., (1969), “Lifetime Portfolio Selection under Uncertainty: the Continuous Time Case,” *Review of Economics and Statistics* 51, 247-257.
- [29] Moscarini, Giuseppe (2004), “Limited Information Capacity as a Source of Inertia,” *Journal of Economic Control and Dynamics* 2003-2035.
- [30] Parker, Jonathan (2003), “Consumption Risk and Expected Stock Returns,” *NBER Working Paper* 9548.
- [31] Parker, Jonathan and Christian Julliard (2004), “Consumption Risk and the Cross-section of Expected Returns,” *Journal of Political Economy* 113 (1), 185-222. .
- [32] Peng, Lin (2004), “Learning with Information Capacity Constraints,” *Journal of Financial and Quantitative Analysis* 40, 307-330.
- [33] Peng, Lin and Wei Xiong (2006), “Investor Attention, Overconfidence and Category Learning,” *Journal of Financial Economics* 80 (3), 563-602.
- [34] Piazzesi, Monika (2002), “Comment” on Gabaix and Laibson, “The 6D Bias and the Equity-Premium Puzzle.” In *NBER Macroeconomics Annual* 2001, vol. 16.
- [35] Pischke, J.S. (1995), “Individual income, incomplete information, and aggregate consumption.” *Econometrica* 63 (4), 805–840.

- [36] Shannon, Claude (1948), "A Mathematical Theory of Communication," *The Bell System Technical Journal* 27, 623-656.
- [37] Sims, Christopher (2003), "Implications of Rational Inattention," *Journal of Monetary Economics* 50, 665-690.
- [38] Sims, Christopher (2005), "Rational Inattention: A Research Agenda," Princeton University, mimeo.
- [39] Starr-McCluer, Martha (2000), "Stock Market Wealth and Consumption," Federal Reserve Board, mimeo.
- [40] Van Nieuwerburgh, Stijn and Veldkamp, Laura (2006), "Information Acquisition and Portfolio Under-Diversification," forthcoming, *Review of Economic Studies*.
- [41] Viceira, Luis (2001), "Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income," *The Journal of Finance* 56 (2), 433-470.
- [42] Vissing-Jørgensen, Annette (2002), "Limited Asset Market Participation and Elasticity of Intertemporal Substitution," *Journal of Political Economy* 110(4), 825-853.