

Portfolio Choice with Information-Processing Limits*

Altantsetseg Batchuluun, Yulei Luo, and Eric R. Young†

In this paper, we examine the joint consumption-portfolio decision of an agent with limited information-processing capacity (rational inattention or RI) in the sense of Sims (2003) within a non-linear-quadratic (non-LQ) setting. Our model predicts that, as processing capacity falls, agents choose to hold less of their savings in the form of risky assets on average; however, they still choose to hold substantial risky assets with some positive probability. Low capacity causes households to act as if they are more risk averse and more willing to substitute consumption intertemporally.

Key Words: Rational inattention; Non-LQ setting; Optimal consumption saving; Portfolio choice.

JEL Classification Numbers: D53, D81, G11.

1. INTRODUCTION

Standard models of portfolio choice typically make two predictions inconsistent with data — they predict essentially a 100 percent participation rate and also a 100 percent portfolio share of risky assets. Even casual observations using US data show these predictions are far off; as noted in Guvenen (2009), the participation rate in equity markets is no greater than 50 percent (and historically has been quite a bit smaller), and as noted in Gabaix and Laibson (2002) the total share of risky assets in total wealth is around 22 percent. See Tables 1 and 2 for some basic accounts of the U.S. household portfolios from the U.S. Survey of Consumer Finances (SCF).

* We thank helpful conversations with Ken Judd and Chris Sims, and seminar participants at University of Virginia and University of Hong Kong. Batchuluun and Young acknowledge the financial support of the Bankard Fund for Political Economy. Luo thanks the Hong Kong General Research Fund (GRF: #HKU17500117) for financial support.

† Batchuluun: Department of Economics, National University of Mongolia. E-mail: ab2hk@virginia.edu; Luo: Faculty of Business and Economics, University of Hong Kong, Hong Kong. CEMA, Central University of Finance and Economics. Email: yulei.luo@gmail.com; Young: Department of Economics, University of Virginia, and Department of Economics, Zhejiang University. E-mail: ey2d@virginia.edu.

Table 1 shows that a significant proportion of the U.S. households do not hold stocks either directly or indirectly through stock mutual funds and other funds in the past ten surveys. Table 2 shows that conditional on holding stocks the risky share of the financial wealth is less than 100%, and increases with income, net worth and education attainment.

TABLE 1.
Percent of Families with Stock Holdings (Directly-held Stocks), SCF.

Year (percent %)	(1) Less than high school	(2) High school degree	(3) Some college	(4) College degree or more
1992	10.9(4.5)	29.8(11.1)	40.2(19.2)	58.5(29.4)
1995	15.4(4.5)	32.9(10.7)	43.0(13.4)	61.9(27.5)
1998	17.9(5.0)	41.2(12.9)	53.7(20.6)	69.1(31.4)
2001	16.8(5.6)	44.9(13.0)	55.1(20.0)	76.4(37.1)
2004	14.4(4.7)	41.1(12.4)	48.8(17.7)	72.9(35.3)
2007	20.1(3.9)	42.3(9.3)	54.2(17.4)	75.5(31.6)
2010	15.0(2.2)	39.1(8.1)	46.4(11.3)	72.1(27.2)
2013	13.2(2.8)	35.3(6.4)	44.6(9.4)	72.0(24.9)

Note: Stock holdings here includes: (1) directly-held stock; (2) stock mutual funds: full value if described as stock mutual fund and 1/2 value of combination mutual funds; (3) IRAs/Keoghs invested in stock: full value if mostly invested in stock, 1/2 value if split between stocks/bonds or stocks/money market, and 1/3 value if split between stocks/bonds/money market; (4) other managed assets w/equity interest (annuities, trusts, MIAs): full value if mostly invested in stock, 1/2 value if split between stocks/MFs & bonds/CDs, or “mixed/diversified,” 1/3 value if “other”; and (5) thrift-type retirement accounts invested in stock full value if mostly invested in stock, 1/2 value if split between stocks and interest earning assets.

One approach to resolve this inconsistency is to introduce fixed costs of participation in equity markets (see Gomes and Michaelides 2008 for an example and references). These models imply that poor households will typically avoid the equity market, while rich ones will enter (Krusell and Smith 1997 obtain a similar result without fixed costs). However, these models typically imply that an agent, once in the equity market, will still hold a portfolio almost entirely composed of risky assets.

We study this issue in a model with rational inattention (Sims 2003). The literature on rational inattention is now quite large, so we refrain from a lengthy citation list and only discuss the papers that deal directly with the issue at hand (Luo 2010 and Luo and Young 2016).¹ Luo (2010) solves a linear-quadratic (LQ) Gaussian portfolio problem with rational inattention

¹Other papers that study limited information and finance include Gennotte (1986), Gabaix and Laibson (2001), Peng and Xiong (2006), Huang and Liu (2007), Lundtofte (2008), Wang (2009), Mondria (2010), van Nieuwerburgh and Veldkamp (2010), and Luo (2017).

TABLE 2.

Stock Share for Families with Stock Holdings in the 2013 SCF

Income percentile	(1) Less than high school	(2) High school degree	(3) Some college	(4) College degree or more
0 – 20%	0.02	0.06	0.07	0.12
20% – 39.9%	0.07	0.12	0.15	0.19
40% – 59.9%	0.13	0.19	0.21	0.26
60% – 79.9%	0.17	0.25	0.28	0.33
80% – 89.9%	0.20	0.29	0.33	0.40
90% – 100%	0.18	0.35	0.36	0.47
Networth percentile	(1) Less than high school	(2) High school degree	(3) Some college	(4) College degree or more
0 – 24.9%	0.03	0.07	0.09	0.13
25% – 49.9%	0.07	0.16	0.19	0.25
50% – 74.9%	0.10	0.20	0.25	0.32
75% – 89.9%	0.13	0.29	0.32	0.40
90% – 100%	0.21	0.32	0.37	0.47

and finds that the model can rationalize the low portfolio shares observed in the data, provided the constraint on information flow is sufficiently tight. When agents cannot learn immediately from signals, they are exposed to “long run risk” and therefore demand more compensation to bear it — that is, they appear “more risk averse” than their preferences would indicate. The problem with Luo (2010) is that the required information flow limit is so tight that it may render the LQ approximation highly inaccurate (see Sims 2005, 2006).²

Luo and Young (2016) address this problem by introducing recursive utility and parameterizing preferences such that agents have a preference for early resolution of uncertainty. In the presence of long-run risk, households who dislike late resolution will demand even more compensation for bearing it than expected utility households (who are indifferent to the timing of uncertainty resolution); this effect amplifies the enhanced risk aversion and thus matching portfolio shares can be done at higher information flow limits (roughly double those in Luo 2010). However, one would still worry that the LQ-Gaussian approximation is inaccurate, and neither model can address the limited participation rate.

Our goal here is to study the portfolio problem with rational inattention in its full nonlinear generality. In general this problem is intractable, even

²Luo (2017) also found that RI only has minor effects on optimal consumption and portfolio rules within a continuous-time constant-absolute-risk-aversion (CARA)-Gaussian framework.

numerically. First, the choice variable for the agents is the joint distribution of states and controls, which is typically very high-dimensional. As shown in Jung, Kim, Matejka and Sims (2015) and Saint-Paul (2011), the optimal distribution is typically discrete and has “holes”, meaning it cannot be described or even approximated by a low-dimensional object; the LQ-Gaussian setup avoids the curse of dimensionality because the optimal distribution is Gaussian (see Sims 2003 and Shafieepoorfard and Raginsky 2013). To combat this problem one either solves a model with only one shock (Tutino 2012) or a very short horizon (Sims 2006, Lewis 2009). Since background risk plays an important role in portfolio, the first approach is undesirable, so we adopt the second and study only a two-period portfolio problem.

Our main result is that the nonlinear RI model can replicate the two key facts, namely incomplete participation and low risky shares invested in the stock market conditional on participation, provided the flow constraint is tight enough (unfortunately we cannot compare the flow constraint to those in Luo 2010 and Luo and Young 2016, due to the short horizon). More interesting is that we obtain significant heterogeneity in portfolio holdings, despite all agents being *ex ante* identical. Since it is not optimal to choose a distribution that is degenerate unless information flow is unconstrained, *ex post* the agents can “receive” different portfolios from nature — the key is that the agents are willing to instruct nature to pick from a distribution that puts positive mass on three very different portfolio weights (roughly 0, 30 percent, and 80 percent). Since most of the mass is on the 30 percent share, the average portfolio share is only 37 percent, a bit too large relative to the data but definitely in the ballpark; the model does not replicate the observed 50 percent participation rate, but we could introduce a fixed cost of participation to potentially resolve this inconsistency.

Our results are driven by two factors. First, rational inattention enhances risk aversion, as noted above. However, rational inattention also appears to enhance the willingness of households to intertemporally substitute consumption over time, as the gap between expected current and future consumption is larger as information flow capacity falls. The combination is precisely the configuration of preferences that the long-run risk and recursive utility literature (see Bansal and Yaron 2004) needs to generate realistic asset prices. Thus, our results indicate that rational inattention could be a method for reconciling the high IES needed to match asset market facts with the apparent low estimate evident from consumption data.

This paper is organized as follows. Section 2 presents a full-information rational expectations (FI-RE) two-period model with consumption and portfolio choice. Section 3 introduces information-processing constraint into this otherwise standard model and discusses how to solve the model

numerically. Section 4 presents the main results. Section 5 concludes. Appendices contain the derivations that are omitted from the main text.

2. A STANDARD FI-RE TWO-PERIOD PORTFOLIO CHOICE MODEL

Our main interest in this paper lies with the portfolio problem under limited information-processing capacity. However, it is convenient to draw distinctions between solutions with and without these limitations, so here we present a standard two-period portfolio problem; Samuelson (1969) and Merton (1969) provide complete analyses of this problem in the case of HARA-class utility functions in continuous-time. Consider an agent with an iso-elastic utility function $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$, where $\gamma \geq 0$ is the coefficient of relative risk aversion (CRRA). This agent faces stochastic current income e^1 and stochastic future income e^2 with distributions of $g_1(e^1)$ and $g_2(e^2)$, respectively.

There are two tradable financial assets available to the household, one risky and one risk-free. The return on the risk free asset is r^f and the return on the risky asset over the period is r^e . We consider the risky asset to be a market portfolio of equities with return distribution $\varphi(r^e)$. Letting r be the one period gross return to invested wealth, we obtain

$$r = sr_d^e + (1-s)r^f = s(r_d^e - r^f) + r^f,$$

where s is the proportion of wealth invested in the risky asset. In period 1, wealth w^1 is simply e^1 as the initial wealth is assumed to be 0 and saving (borrowing) is $e^1 - c^1$. In period 2, wealth consists of the return on savings and future income $(w^1 - c^1)r + e^2$. Following Sims (2006), we assume that second period consumption is equal to wealth.³ For reasons we will elaborate on more completely later, we discretize both the state and control space.

The maximization problem for this agent can be written as:

$$\max_{\{c(w_i^1)\}, \{s(w_i^1)\}_{i=1}^I} \left\{ \sum_{i=1}^I u(c(w_i^1)) g_1(e_i^1) + \beta \sum_{i=1}^I \sum_{d=1}^D \sum_{j=1}^J u((e_i^1 - c(w_i^1)) [s(w_i^1)(r_d^e - r^f) + r^f] + e_j^2) \varphi(r_d^e) g_1(e_i^1) g_2(e_j^2) \right\}.$$

In period 1, before income is realized the agent makes a contingent plan for consumption and savings that depends on the realization of e^1 . This

³In the standard model this assumption is without loss of generality provided utility is increasing in consumption. In the model with limited information-processing capacity, households may leave accidental bequests because they are uncertain about their exact wealth. Adding a consumption choice in the second period is computationally costly and would not add any new insight.

specification is equivalent to the usual timing in which the agent makes consumption-savings plan after the realization of e^1 , but it will turn out to be easier to formulate the rational inattention model with this timing. The plan for period 1 then determines consumption in period 2 based on the realizations of e^2 and r^e .

We assume that agents can borrow and that consumption in each date and state must be nonnegative:

$$c(w_i^1) \geq 0, \quad (1)$$

$$(e_i^1 - c(w_i^1)) (s(w_i^1) (r_d^e - r^f) + r^f) + e_j^2 \geq 0, \quad (2)$$

for $\forall i = 1, \dots, I$, $\forall j = 1, \dots, J$, and $\forall d = 1, \dots, D$. Since this problem has a continuous and concave objective function and a convex opportunity set, the following first-order conditions are necessary and sufficient:

$$\begin{aligned} c_1(w_i^1)^{-\gamma} &= \beta \sum_{d=1}^D \sum_{j=1}^J \left\{ (e_i^1 - c(w_i^1)) [s(w_i^1) (r_d^e - r^f) + r^f] + e_j^2 \right\}^{-\gamma} \\ &\quad \times [s(w_i^1) (r_d^e - r^f) + r^f] \varphi(r_d^e) g_2(e_j^2), \\ 0 &= \beta \sum_{d=1}^D \sum_{j=1}^J \left(\begin{array}{l} \{ (e_i^1 - c_1(w_i^1)) [s(w_i^1) (r_d^e - r^f) + r^f] + e_j^2 \}^{-\gamma} \times \\ (e_i^1 - c_1(w_i^1)) (r_d^e - r^f) \varphi(r_d^e) g_2(e_j^2) \end{array} \right), \end{aligned}$$

for $\forall i = 1, \dots, I$. The first condition is the condition on optimal consumption over time: the discounted marginal utilities are equalized. The second condition is the condition for optimal (additional) risk taking. We cannot get a closed-form solution to this problem except in special cases that are not of interest to us (e.g., quadratic or CARA utility). As shown in Merton (1969) and Samuelson (1969), one expects the following properties would hold: (i) optimal portfolio choice s^* is decreasing in γ , (ii) second period consumption and s^* are higher when r^f is low relative to β , and (iii) a mean-preserving spread in the risky asset return decreases s^* .

As in the literature, we can also use a log-linearization method to solve the FI-RE problem. However, because that method does not work when information-processing constraints are imposed, we solve this problem numerically. In the next section, we introduce these constraints and study the portfolio decisions of agents with rational inattention.

3. A RATIONAL INATTENTION NON-LQ MODEL OF PORTFOLIO CHOICE

3.1. Introducing Rational Inattention due to Information-Processing Constraints

We now assume agents have limited information processing capacity in the sense of Sims (2003). Agents choose the optimal joint probability distribution of consumption, portfolio allocation, and wealth to maximize their expected lifetime utility. Hence the choice variables with respect to which we maximize is the joint probability distribution function $f(\cdot)$ of consumption c^1 and the share of savings held in the risky asset s with current income e^1 . The objective function for the agent is

$$\max_{\{f(s_k, c_r^1, e_i^1)\}} \left\{ \sum_{k=1}^K \sum_{r=1}^N \sum_{i=1}^I \left(u(c_r^1) + \sum_{d=1}^D \sum_{j=1}^J \beta \left[u(c_{i,r,k,d,j}^2) \right] \varphi(r_d^e) g_2(e_j^2) \right) f(s_k, c_r^1, w_i^1) \right\} \quad (3)$$

where

$$c_{i,r,k,d,j}^2 = (w_i^1 - c_r^1) [s_k (r_d^e - r^f) + r^f] + e_j^2$$

is second-period consumption.

As in the FI-RE case, we assume that agents can borrow and the consumption in each period must be nonnegative. Budget constraints in this model enforce the nonnegativity requirement; they are satisfied automatically if agents are saving ($c_r^1 < w_i^1$ for grid points r and i) but impose restrictions whenever agents are borrowing ($c_r^1 > w_i^1$). In the borrowing case the budget constraints for period 2 take the form

$$f(s_k, c_r^1, e_i^1) = 0 \quad \text{if} \quad (w_i^1 - c_r^1) [s_k (r_d^e - r^f) + r^f] + e_j^2 \leq 0, \quad (4)$$

for $\forall d = 1, \dots, D, \forall r = 1, \dots, N, \forall i = 1, \dots, I, \forall j = 1, \dots, J, \forall k = 1, \dots, K$. The choice set is also restricted by the requirement that the probability density must be well-defined:

$$0 \leq f(s_k, c_r^1, w_i^1) \leq 1, \quad (5)$$

where $k = 1, \dots, K, r = 1, \dots, N$, and $i = 1, \dots, I$. Since income is exogenous, the marginal probability of w^1 chosen by the household must be equal to the probability distribution function of e^1 :

$$\sum_{k=1}^K \sum_{r=1}^N f(s_k, c_r^1, w_i^1) = g_1(e_i^1), \quad (6)$$

where $i = 1, \dots, I$. The final constraint is the information processing constraint (IPC). To formulate the IPC we need to define the mutual

information between random variables. The mutual information between current consumption c^1 , the share of the risky asset s , and current income e^1 , is defined as

$$I(w^1; c^1, s) = H(w^1) + H(s, c^1) - H(s, c^1, w^1),$$

where $I(w^1; c^1, s)$ measures the reduction in uncertainty about w^1 made possible by observing c^1 and s , and is always nonnegative. The assumption of limited information processing capacity κ requires mutual information not exceed finite capacity; that is, $I(w^1; c^1, s) \leq \kappa$. Therefore, the IPC is given by the nonlinear inequality

$$\begin{aligned} \sum_{k=1}^K \sum_{r=1}^N \sum_{i=1}^I f(s_k, c_r^1, w_i^1) \log(f(s_k, c_r^1, w_i^1)) - \sum_{i=1}^I g_1(e_i^1) \log(g_1(e_i^1)) \quad (7) \\ - \sum_{i=1}^I \sum_{k=1}^K \sum_{r=1}^N f(c_r^1, s_k, w_i^1) \log(q(c_r^1, s_k)) \leq \kappa, \end{aligned}$$

where the marginal distribution of consumption in period 1 is $q(c^1, s)$ and defined by

$$\sum_{i=1}^I f(s_k, c_r^1, w_i^1) = q(c_r^1, s_k), \quad (8)$$

where $r = 1, \dots, N$ and $k = 1, \dots, K$. (See Appendix A.1 for the derivation of (7).) It is easy to see that this constraint always binds. The following proposition is straightforward.

PROPOSITION 1. *The objective function is continuous and concave and the constraint set is compact. Therefore, (3) has a solution.*

The problem for the information-constrained household is to maximize (3) with respect to (4)-(7). (See Appendix A.2 for the derivation of the first-order conditions.) For many points in the discrete outcome space $f(s_k, c_r^1, w_i^1) = 0$ may be binding (but obviously not for all of them), meaning that (3) is a highly-nonlinear problem; furthermore, even for relatively coarse discretizations the number of choice variables and constraints is very large. Linearization techniques are not applicable, so we solve the problem directly through a “brute force” method (albeit a highly sophisticated one). Before proceeding to the numerical solutions we examine a special case: $\kappa = 0$. A second special case, $\kappa = \infty$, is equivalent to the standard FI-RE model solved in the previous section. We note also that the Maximum Theorem and the Envelope Theorem apply to our model, so that we can state the following results.

PROPOSITION 2. *The policy function for (3) is continuous in $\kappa \geq 0$.*

PROPOSITION 3. *The value function for (3) is differentiable in $\kappa > 0$ and satisfies $Dv(\kappa) > 0$.*

3.2. Zero Information-Processing Capacity Case

We now consider the $\kappa = 0$ case. As noted above, $I(w^1; s, c^1)$ measures the reduction in uncertainty of w^1 after observing c^1 and s , and it is always nonnegative. In other words, knowledge about (c^1, s) cannot increase uncertainty of w^1 :

$$I(w^1; s, c^1) = H(w^1) - H(w^1|c^1, s) \geq 0.$$

On the other hand, zero information processing capacity requires

$$I(w^1; s, c^1) = H(w^1) - H(w^1|c^1, s) \leq 0.$$

Combining these two inequalities yields:

$$I(w^1; s, c^1) = H(w^1) - H(w^1|c^1, s) = 0,$$

which implies that zero information processing capacity requires that current wealth be independent of current consumption and the risky asset share:

$$f(s_k, c_r^1, w_i^1) = q(s_k, c_r^1) g(e_i^1).$$

Agents cannot condition the distribution of the current wealth on the realization of current consumption and the risky asset share, since those observations carry no usable information. We formally state this result in the following proposition.

PROPOSITION 4. *When $\kappa = 0$, the distribution of w^1 and the joint distribution of (c^1, s) are independent.*

This case illustrates that households with low processing capacity will separate their actions from their state, leading to a kind of “inertia” (see Tutino 2012 for a dynamic illustration of this inertial effect); in a model with a longer horizon, household decisions will appear sticky as they do not respond to changes in the state. Note that this collapse looks very much like increased risk aversion, wherein consumption does not vary much across states of the world. We will see this effect manifest itself at positive κ as

well.⁴ Using the continuity of the solution we note that this case is the limit as $\kappa \rightarrow 0$.

3.3. Computation

Our computational method is similar to those used in Sims (2005) and Lewis (2006). As noted above, we discretize the state and outcome spaces and permit the agent to attach probabilities to each of those outcomes, subject to the appropriate restrictions; the choice of probabilities is continuous. We assume that first period income has 16 grid points with values ranging from 0.01 to 0.16 and second period income has 4 grid points ranging from 0.02 to 0.08. The risky return has 8 grid points ranging from 0.79 to 1.35. For simplicity we assume that both second period income and the risky return are uniform, and the distribution of first period incomes normal with mean 0.085 and standard deviation is 0.023. We also assume no correlation between labor income and the risky return in the second period; aggregate data shows little correlation between stock returns and wages at business cycle frequencies, so this assumption seems a natural benchmark.⁵ For the outcome space, current consumption has 32 grid points between $[0.005, 0.16]$ and the risky asset share has 41 grid points between $[0, 1]$. The risk-free rate is $r^f = 1.02$, the expected excess return is 0.05, and the standard deviation of the risky return is 19.6 percent; these values are consistent with the premium of US equities over T-bills. We set $\beta = 0.97$ and consider several plausible values of γ .⁶

The resulting problem has a large number of choice variables and a large number of constraints, many of which are equality constraints. To solve this problem, we use the AMPL programming environment as a gateway to the KNITRO solver.⁷ We then upload our program to the NEOS Server for Optimization, a publicly-available resource that implements the AMPL-KNITRO program (it greatly increases the size of the problem that we can consider as it uses idle supercomputer resources).⁸ In total, our problem

⁴Luo and Young (2016) stress a different manifestation of rational inattention as enhanced risk aversion when agents have a preference for early resolution of uncertainty — since rational inattention delays the resolution of consumption risk, it is unappealing to an agent who prefers to learn early rather than late.

⁵Heaton and Lucas (2000) study how the presence of background risks influences portfolio allocations. They find that labor income is the most important source of wealth and labor income risk is weakly positively correlated with equity returns. Our results are quite similar when the correlation is close to but not exactly equal to zero; results from these experiments are available upon request.

⁶The value for β guarantees that agents will save in some parts of the state space and borrow in others. The value of β plays no important role here.

⁷The KNITRO solver is a commercial solver used for large nonlinear problems that combines automatic differentiation with sophisticated Newton-based iterations and active set methods to handle the constraints.

⁸The website is <http://www.neos-server.org/neos/>.

features 14,226 choice variables, 1328 linear equality constraints, and 1 nonlinear inequality constraint.

4. MAIN FINDINGS

In this section, we will first solve the FI-RE model numerically and then solve the model with limited information processing capacity (IPC).

4.1. Standard FI-RE Model

In a standard FI-RE model agents observe the realization of current wealth with certainty; since their objective functions are strictly concave, their decisions become deterministic functions of this realization. For comparability reasons we solve the FI-RE models using same numerical approach that we apply to the limited information-processing setting.

FIG. 1. The Share Invested in the Risky Asset (FI-RE)

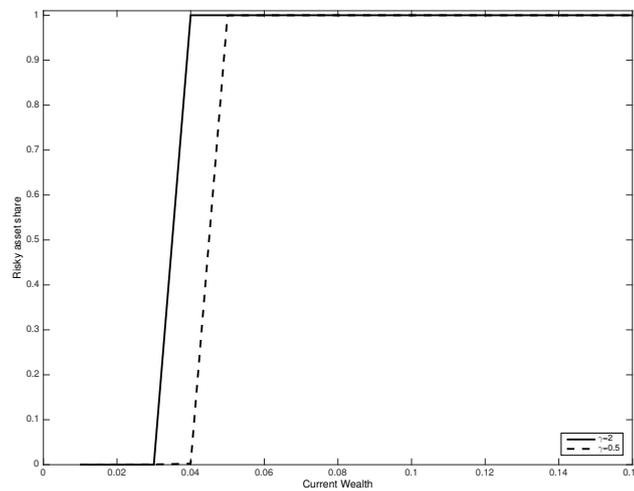


Figure 1 shows the risky asset share over the grid of current period wealth. Agents choose to invest all their savings in the risky asset at the states with relatively high levels of wealth and choose to borrow at the risk-free rate at the states with relatively low levels of wealth. The resulting decision rule is essentially discontinuous; a related but (slightly) less-extreme result can be found in Krusell and Smith (1997) that is driven by the absence of insurance markets. The expected risky asset share is 0.95 for the agent with $\gamma = 0.5$ and 0.97 for the agent with $\gamma = 2$.

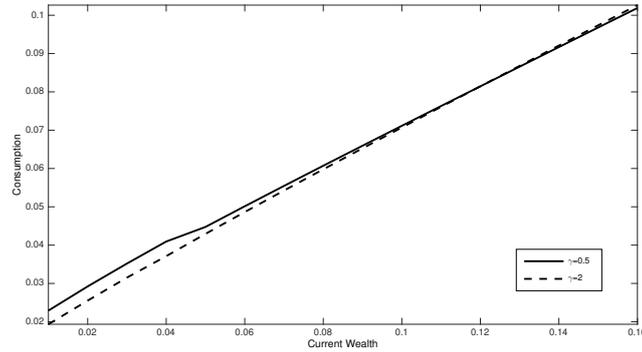
FIG. 2. Optimal Consumption under FI-RE

Table 3 presents expected (average) consumption in each period for both $\gamma = 0.5$ and $\gamma = 2$. The difference between first and second period consumption is smaller for the agent who is less risk averse ($\gamma = 0.5$). Figure 2 shows optimal consumption on the grid points for current period wealth. First period consumption increases as current period wealth increases for both types of agent. However, the $\gamma = 0.5$ agent chooses slightly larger consumption at lower wealth than the $\gamma = 2$ agent and chooses slightly lower consumption mass at higher current wealth.

TABLE 3.

Mean Consumption in the First- and Second-Period

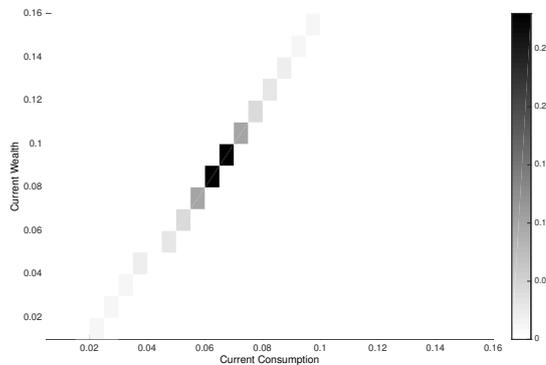
	$\gamma = 0.5$	$\gamma = 2$
$E [c_1]$	0.0633	0.0624
$E [c_2]$	0.0732	0.0742

The more risk-averse agent is more concerned about the states with very low consumption in the future, and therefore chooses to borrow in a smaller portion of the current wealth space and receives lower current consumption than the less risk-averse agent does. At the states with borrowing, both types of agents choose to minimize borrowing costs by setting the risky asset share to be 0, and both types of agents choose to invest all their savings in the risky asset (to maximize the return) at the states with savings. Hence, the choice of the risky asset share is discrete with values of 0 and 1 depending on whether the saving is positive or negative. The overall probability of borrowing is 0.03 for the $\gamma = 2$ agent and 0.05 for the $\gamma = 0.5$ agent, which is why the $\gamma = 2$ agent appears to hold a higher share of risky assets.

4.2. Joint Distribution of Current Wealth and Current Consumption

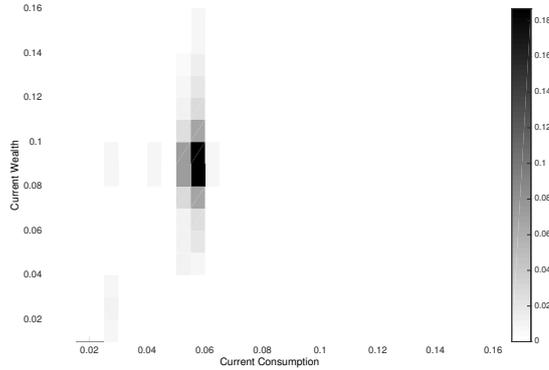
Note that as $\kappa \rightarrow \infty$ the rational inattention model converges to the FI-RE model; we therefore expect that the benchmark model will accurately represent choices of households whose capacity constraints are large. How much capacity is sufficient is model-dependent. For the benchmark model, the exogenous distribution for the current income is centered on its mean 0.085, with the entropy of this process being $H(e^1) = -\sum_{i=1}^I g_1(e_i) \log(g_1(e_i)) = 2.074$ nats.⁹ From the definition of mutual information we know that the knowledge of current consumption and share of risky asset can decrease the uncertainty about current wealth at most by this amount. Hence for $\kappa \geq 2.074$ nats the model gives nearly identical results as the model with unlimited information processing capacity.

FIG. 3. Joint Distribution of Current Wealth and Consumption (Large κ)



Shaded plots of the joint densities of c_1 and w_1 are shown in Figures 3 and 4. The darker the box, the higher the probability weight placed on the corresponding grid point. For sufficiently large κ , Figure 3 shows that the agent chooses a distribution with perfect correlation between current wealth and current consumption and places positive probability only on the FI-RE solutions. However, as κ decreases the agent allocates probability on fewer grid points with low consumption. Small κ does not allow the agent to learn every realization of the current wealth. Thus he wants to be well informed about the grids with lower wealth to prevent future consumption close to zero. Figure 4 shows the joint probability distribution of c_1 and w_1 of an agent with $\kappa = 0.1$, which allows him to resolve only about 5 percent

⁹A ‘nat’ is the unit of information flow when entropy is measured relative to the natural logarithm function. Other measures include the ‘bit’ for base 2 logarithms and the ‘dit’ or ‘hartley’ for base 10. Throughout the paper we use ‘nat’ as the unit of κ .

FIG. 4. Joint Distribution of Current Wealth and Consumption ($\kappa = 0.1$)

of the total uncertainty in current wealth. The agent puts probability on fewer grid points with low consumption and puts probability 0.64 on the grid point with consumption of 0.055; this collapse of the distribution is a reflection of the inertial effects (enhanced risk aversion) discussed already for the $\kappa = 0$ case. Note that the correlation between c_1 and w_1 is 1 under FI-RE and only 0.36 when $\kappa = 0.1$.

4.3. Expected Current and Future Consumption

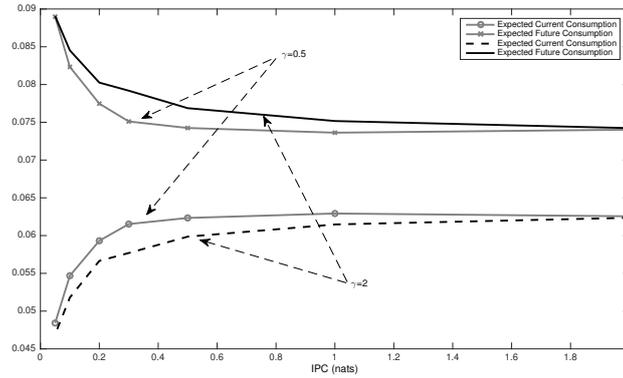
Figure 5 presents expected current and future consumption across two values of γ . The key result here is that the “gap” between current and future average consumption is decreasing in κ . This gap is a measure of the intertemporal elasticity of substitution, as it reflects the willingness to let consumption vary deterministically across time. Thus, households who have low capacity will appear to be more willing to intertemporally substitute than standard FI-RE agents.

TABLE 4.

Expected Saving Rate under RI								
κ	∞	2	1	0.5	0.3	0.2	0.1	0.05
$\gamma = 2$	0.27	0.27	0.28	0.30	0.32	0.33	0.39	0.45
$\gamma = 0.5$	0.27	0.26	0.26	0.27	0.28	0.30	0.36	0.43

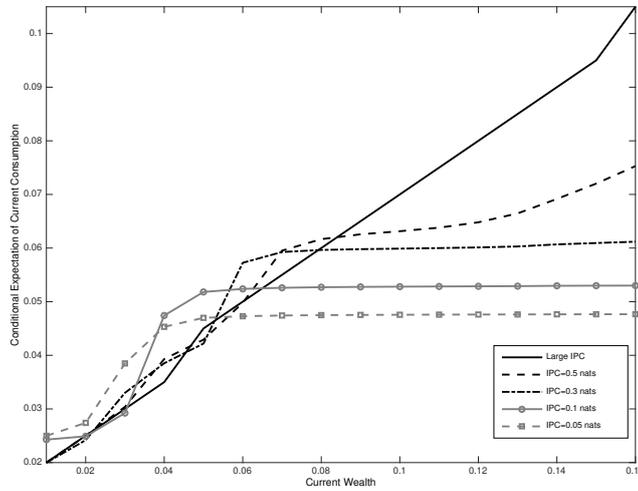
In the previous section, we observed that the agent with limited information processing capacity allocates positive mass to fewer grid points with low values, resulting in a smoother consumption distribution; for example, when $\kappa < 0.2$ the conditional expectation of consumption is low and almost flat across different values of wealth (see Figure 6). As a result of

FIG. 5. Expected Current Consumption and Future Consumption



the consumption smoothing over the current states with higher wealth, the expected “precautionary saving”, the difference between the expected current income and the expected current consumption, increases substantially with smaller κ . Table 4 shows the saving rate, ratio of the expected saving and the expected current income; as κ gets smaller the saving rate increases substantially, leading to high future consumption on average. Note that the increase in future consumption will be smaller for longer horizon models.

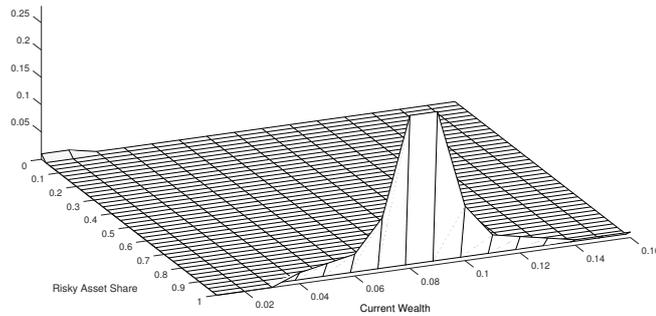
FIG. 6. Conditional Expectation of Current Consumption



4.4. Optimal Share of the Risky Asset

When information processing capacity is large the agent has a lot of information about current wealth, and thus his decision rule is very similar to the FI-RE one — allocate positive mass only to “zero risky share, borrow” and “risky share equal to one, save”. There is only a small change in distribution of the risky asset share in response to a reduction in κ until it gets very small. To choose the optimal distribution of the risky asset share the agent would like to know if the current wealth is higher or lower than the threshold level that triggers the switch from saving to borrowing. This information is available until κ becomes very small. When κ becomes too small this information cannot be obtained and the agent responds by allocating probabilities over grid points with a low risky asset share to decrease the probability of borrowing at a high interest rate; at the same time the agent does not want to forego the excess return entirely, so some probability is still attached to saving with a positive risky share.¹⁰

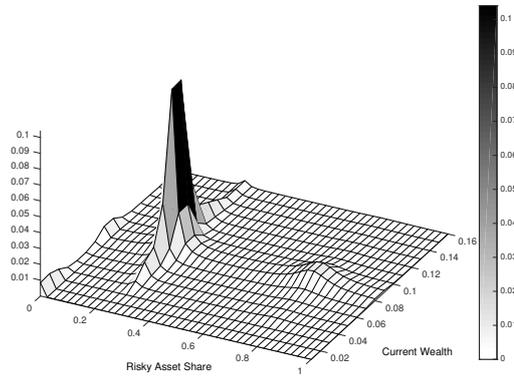
FIG. 7. Joint Distribution of Current Wealth and Portfolio Choice ($\kappa = 2$)



Figures 7 and 8 present the joint distribution of the risky asset share and the current wealth for information processing capacity of 2 and 0.1, respectively. At the capacity $\kappa = 0.1$, the expected risky asset share is 0.37, which is not that far from the average share cited in Gabaix and Laibson (2002) and Luo and Young (2016), while for $\kappa = 2$ the agent is putting all of his savings into equities (note the small mass at zero associated with borrowing). Of particular interest is that, with $\kappa = 0.1$, there is a lot of heterogeneity — the agent puts positive mass in three regions cor-

¹⁰This result is consistent with that obtained in Luo (2010), in which less capacity leads to greater long-run risk and thus reduces the optimal share invested in the risky asset. Luo and Young (2016) find that low shares can be sustained with higher κ if agents are averse to late resolution of uncertainty.

FIG. 8. Joint Distribution of Current Wealth and Portfolio Choice ($\kappa = 0.1$)



responding to nonparticipation ($s = 0$), modestly-risky portfolios (around $s = 0.3$), and highly risky portfolios (around $s = 0.8$); if we suppose that this agent is one of infinitely-many identical agents in an economy (and a suitable law of large numbers applies), then this distribution will also be the cross-sectional distribution, so the model can reproduce both incomplete participation in equity markets and low holdings conditional on participation.¹¹

FIG. 9. Expected Value of the Risky Asset

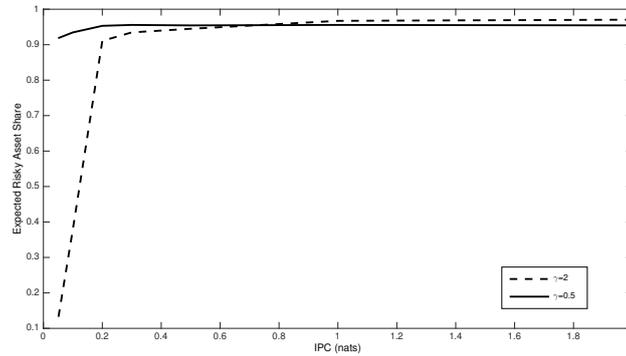
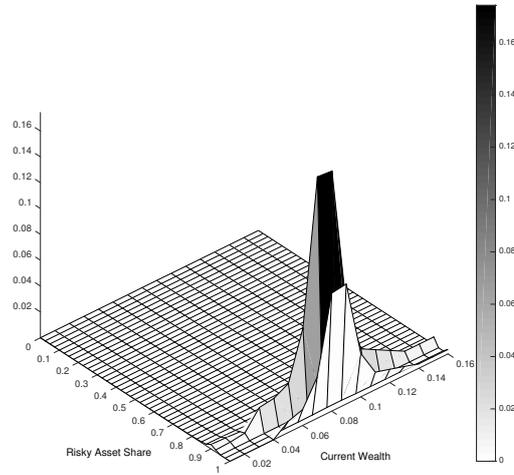


Figure 9 shows the expected value of the optimal share of the risky asset for different levels of information processing capacities. As noted already,

¹¹Models that drive limited participation through fixed costs generally imply that, once the household enters the market, the share of risky assets is too high. Gomes and Michaelides (2008) is one example.

FIG. 10. Joint Distribution of Current Wealth and Risky Asset Share ($\gamma = 0.5, \kappa = 0.1$)

with high κ the risky asset share is very high ($E[s] = 0.975$); because the agent borrows in a small fraction of states, it is below 1. When κ gets low enough (below roughly $\kappa = 0.2$) the average share drops off quickly (for $\gamma = 2$). The level of risk aversion matters a lot when κ is low, which we can see in Figure 9 (the share does not decline significantly until κ gets very small if $\gamma = 0.5$) and Figure 10 (all the mass is placed on a single high value for s instead of being spread across several lower values as when $\gamma = 2$).

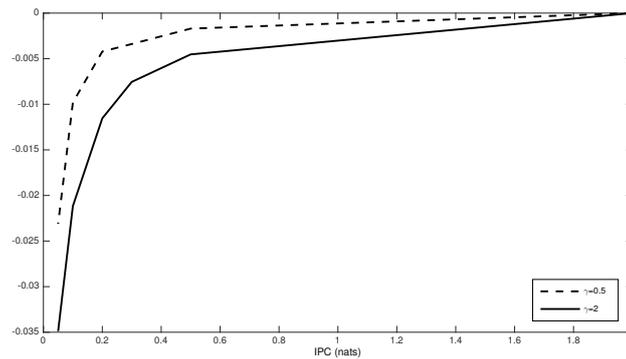
We want to mention here briefly that the average share $E[s]$ depends on the initial distribution of labor income. If we introduce a distribution with more entropy (higher variance), then the curve in Figure 9 shifts down at every value of κ . Higher entropy means more uncertainty about the current state, which translates into more uncertainty about future events and therefore more desire to insure via the risk-free asset. Details of this experiment are available upon request.

4.5. Welfare Implications of RI

Proposition 3 shows that expected welfare is decreasing in the tightness of the constraint. Figure 11 illustrates how the change in welfare increases with the degree of attention for different values of γ . The welfare costs here can be significant, on the order of 0.03 percent of consumption, if κ is very small and γ is large, but generally are modest for relatively high values of κ .

Of course, one must take the numbers from a two-period model with some caution, but they indicate that the general result from the LQ-Gaussian literature that the costs of RI are small (see Luo and Young 2010) may be missing a crucial piece of the picture. Since the large welfare losses come when κ is small, they only arise for parameterizations in which the optimal decisions are decidedly non-Gaussian, and are therefore excluded from the existing studies. However, the large costs also arise precisely in the region of the parameter space where portfolio decisions look most like the data, so we should consider whether they hold up in more general environments.

FIG. 11. Welfare Change



4.6. Correlation Between Income and Risky Asset Returns

In the previous analysis, we assumed zero correlation between labor income and the return to the risky asset. However, the empirical evidence on the correlation between labor income and equity returns shows non-zero or weak negative and positive correlations for various population groups. It is worth noting that estimating the correlation between individual labor income and the equity return is complicated by the lack of data on household portfolio choice that has both time-series and panel dimension, and by the difficulty of identifying households' unanticipated income shocks. Nevertheless, several studies have tried to estimate these effects. For example, as estimated in Davis and Willen (2000), the correlation between aggregate equity returns and labor income shocks ranges from -0.25 for the least educated men to 0.25 or more for college-educated women. Heaton and Lucas (1999) found that the correlation between the entrepreneurial risk and the equity return was about 0.2 . A standard FI-RE portfolio choice model predicts that less-educated investors should be more heavily invested in the stock market while college graduates and entrepreneurs should put less wealth in the stock market. We examined the benchmark model with

FIG. 12. Expected Risky Asset Share

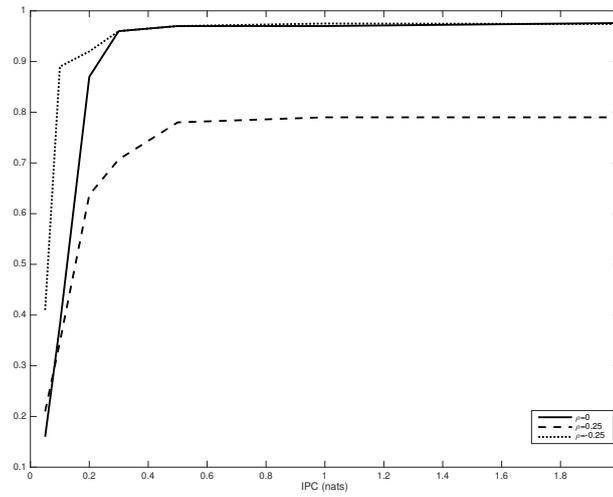
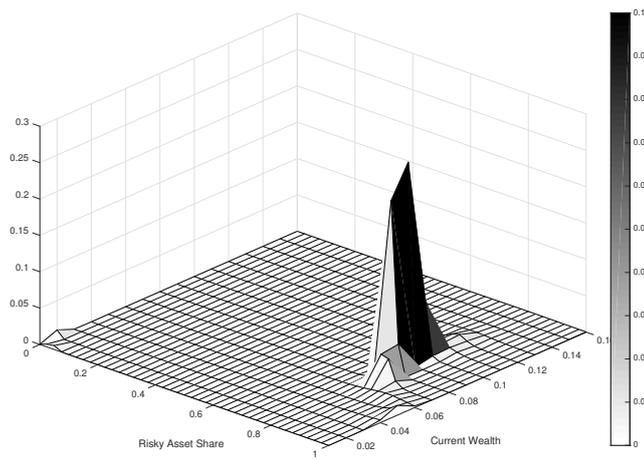


FIG. 13. Joint Distribution of Risky Asset Share and Current Wealth



negative (-0.25) and positive (0.25) correlation between labor income and the risky return. Overall, the results are similar to that obtained in the zero correlation case as the expected risky asset shares and expected cur-

FIG. 14. Joint Distribution of Risky Asset Share and Current Wealth

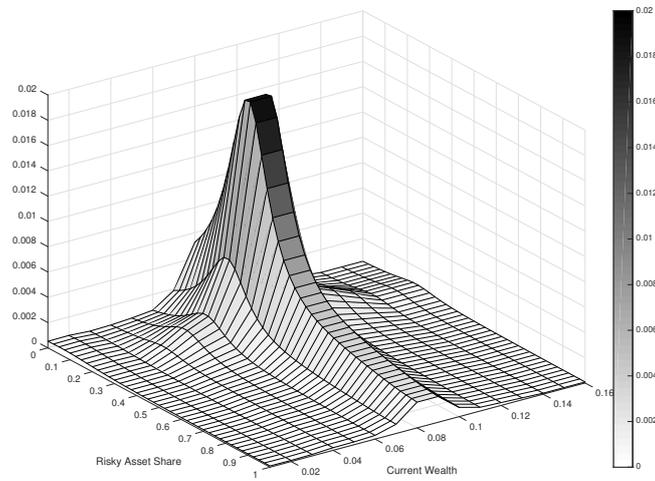
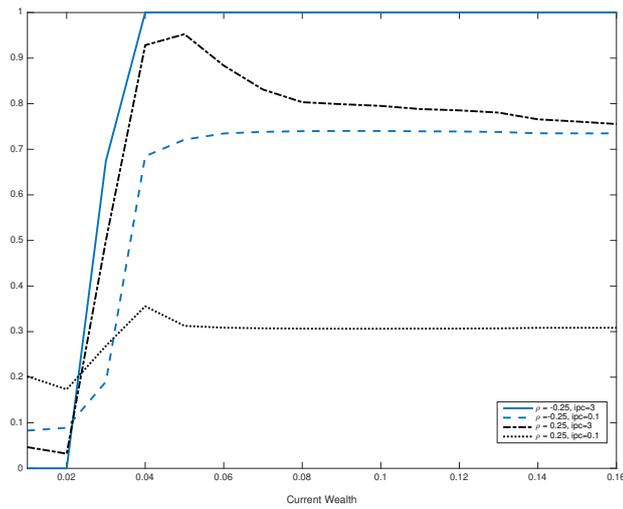


FIG. 15. Expected Risky Asset Share



rent consumptions decrease with the degree of attention. Figure 12 shows that the expected share invested in the risky asset for various degrees of the information processing constraint. When there is a negative correlation

between labor income and the risky asset return, the agent's behavior is very similar to the zero correlation case when κ is sufficiently high (e.g., when $\kappa > 0.4$). However, they will hold much larger share of the risky asset than in the zero correlation case when κ is sufficiently low (e.g., $\kappa < 0.2$). In contrast, when there is a positive correlation between labor income and the return to the risky asset, agents hold smaller share of the risky asset than that in the zero correlation case, which is similar to the FI-RE model. Figures 13 and 14 illustrate the joint distributions of the current wealth and risky asset shares for information processing capacity of 3 and 0.1, respectively. At the capacity $\kappa = 3$, the expected risky asset share is 0.79, where the agent is putting positive mass around 0 and 0.8. Positive correlation between labor income and the return to the risky asset prevents the agent to invest all his saving in risky asset. At the capacity $\kappa = 0.1$, the expected share of the risky asset is 0.35. Capacity constrained agent puts larger probability around 0.3. Figure 15 shows that the share invested in the risky asset for agents with large capacity ($\kappa = 3$) and small capacity ($\kappa = 0.1$).

5. CONCLUSION

We have studied the role of rational inattention — limited information-processing capacity — in a standard two-period portfolio choice problem. Our model is capable of producing an empirically-reasonable share of risky assets in a portfolio for modest levels of risk aversion; in addition, it can produce nonparticipants and households with low risky asset shares. The effect of rational inattention is twofold. First, low processing capacity compresses the distribution of consumption across values of current wealth, an effect similar to an increase in risk aversion. Second, the high precautionary savings that rational inattention generates leads to a large gap between expected current and expected future consumption, making households look like they have high elasticities of intertemporal substitution. The combination therefore reproduces the parameter combination identified as crucial for generating resolutions to the equity premium puzzle.

In this paper, for simplicity, we implicitly assumed the uncertainty to be fully resolved in the future; consequently it may have dampened the effects of rational inattention. To fully understand the impact of RI on individual's and macroeconomic behavior we need a more general model with multiple periods. We are currently exploring the possibility of using “approximate dynamic programming” tools — see Powell (2007) — to break the severe curse(s) of dimensionality that RI problems pose and explore longer horizon problems.

APPENDIX A

A.1. DERIVATION OF IPC

The size of the information produced by the stochastic process is measured by its entropy $H(\cdot)$, following Shannon (1948). $H(X)$, the entropy of a discrete random process X (which is a measure of the expected uncertainty in X), is defined by

$$H(X) = - \sum_x p(x) \log(p(x)).$$

Entropy measures the information provided by a random process about itself.¹ Mutual information is a measure of the information contained in one process regarding another process. Suppose $\{X, Y\}$ is a random process. The average mutual information between X and Y is defined by

$$I(X; Y) = H(X) + H(Y) - H(X, Y)$$

We can also write the mutual information in a more intuitive way using the conditional entropy as

$$I(X; Y) = H(X) - H(X|Y) = H(Y) - H(Y|X),$$

where

$$H(X|Y) = E[-\log(p(X|Y))] = \sum_y H(X|Y = y)p(y)$$

is the conditional entropy of the jointly-distributed X and Y , and $H(X|Y) \leq H(X)$ with equality if and only if X and Y are independent (that is, conditioning on a second random variable can never increase entropy).

In our model, mutual information between the current consumption, savings and the random vector of income in period 1 is given by

$$I(e^1; c^1, s) = H(e^1) - H(e^1|c^1, s)$$

or equivalently

$$I(e^1; c^1, s) = H(e^1) + H(c^1, s) - H(e^1, c^1, s).$$

¹See Cover and Thomas (1991) for a textbook treatment on this topic.

Writing the entropy explicitly yields

$$\begin{aligned} & \sum_{k=1}^K \sum_{r=1}^N \sum_{i=1}^I f(s_k, c_r^1, e_i^1) \log(f(s_k, c_r^1, e_i^1)) - \sum_{i=1}^I g_1(e_i^1) \log(g_1(e_i^1)) \\ & - \sum_{k=1}^K \sum_{r=1}^N q(c_r^1, s_k) \log(q(c_r^1, s_k)). \end{aligned}$$

This expression is constrained to be smaller than the channel capacity, giving rise to (7). Since entropy is a concave function, the resulting constraint set is convex.

A.2. DERIVING FIRST-ORDER CONDITIONS

The first-order conditions for $f(\cdot) \in (0, 1)$ are

$$\frac{U_{kri} - v_i}{1 + \log\left(\frac{f(s_k, c_r^1, w_i^1)}{q(c_r^1, s_k)}\right) - \frac{f(s_k, c_r^1, w_i^1)}{q(c_r^1, s_k)}} = \frac{U_{mns} - v_s}{1 + \log\left(\frac{f(s_m, c_n^1, w_s^1)}{q(c_n^1, s_m)}\right) - \frac{f(s_m, c_n^1, w_s^1)}{q(c_n^1, s_m)}}$$

for $\forall k = 1, \dots, K$, $r = 1, \dots, N$, and $i = 1, \dots, I$, where

$$U_{kri} = u(c_r^1) + \beta \sum_{d=1}^D \sum_{j=1}^J u((w_i^1 - c_r^1) [s_k (r_d^e - r^f) + r^f] + e_j^2) \varphi(r_d^e) g_2(e_j^2)$$

is the expected lifetime utility at the state with consumption c_r^1 , the risky asset share s_k , and wealth w_i^1 . The agent chooses optimal probabilities over different states such that the utility per marginal mutual information are equated across different states.

PROPOSITION 5. *The agent allocates higher relative probability over the states with higher expected utility if $v_i = v_s$.*

Suppose $v_i = v_s$ and $U_{kri} > U_{mns}$. Then

$$\frac{f(s_k, c_r^1, w_i^1)}{q(c_r^1, s_k)} - \log\left(\frac{f(s_k, c_r^1, w_i^1)}{q(c_r^1, s_k)}\right) < \frac{f(s_m, c_n^1, w_s^1)}{q(c_n^1, s_m)} - \log\left(\frac{f(s_m, c_n^1, w_s^1)}{q(c_n^1, s_m)}\right).$$

Let $\varphi(x) = x - \log(x)$. Then $D\varphi(x) < 0$ if $0 < x < 1$, so $\frac{f(s_k, c_r^1, w_i^1)}{q(c_r^1, s_k)} > \frac{f(s_m, c_n^1, w_s^1)}{q(c_n^1, s_m)}$.

REFERENCES

- Bansal, Ravi and Amir Yaron, 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance* **59**, 1481-1509.
- Davis, Steven and Paul Willen, 2000. Using Financial Assets to Hedge Labor. Income Risks: Estimating the Benefits. manuscript, University of Chicago and Boston Fed.
- Gabaix, Xavier and David Laibson, 2001. The *6D* Bias and the Equity-Premium Puzzle. *NBER Macroeconomic Annual 2001*, 257-330.
- Genotte, Gerard, 1986. Optimal Portfolio Choice under Incomplete Information. *Journal of Finance* **41**, 733-749.
- Gomes, Francisco and Alexander Michaelides, 2008. Asset Pricing with Limited Risk Sharing and Heterogeneous Agents. *Review of Financial Studies* **21(1)**, 415-448.
- Guvenen, Fatih, 2009. A Parsimonious Macroeconomic Model for Asset Pricing. *Econometrica* **77(6)**, 1711-1750.
- Haliassos, Michael and Carol Bertaut, 1995. Why Do So Few Hold Stocks? *Economic Journal* **105(432)**, 1110-29.
- Heaton, John and Deborah Lucas, 2000. Portfolio Choice in the Presence of Background Risk. *Economic Journal* **110**, 1-26.
- Huang, Lixin and Hong Liu, 2007. Rational Inattention and Portfolio Selection. *Journal of Finance* **62(4)**, 1999-2040.
- Krusell, Per and Anthony A. Smith, Jr., 1997. Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns. *Macroeconomic Dynamics* **1(2)**, 387-422.
- Jung, Junehyuk, Jeong-Ho Kim, Filip Matejka, and Christopher A. Sims, 2015. Discrete Actions in Information-Constrained Tracking Problems. manuscript. PROBLEMS
- Lewis, Kurt F., 2009. The Two-Period Rational Inattention Model: Accelerations and Analyses. *Computational Economics* **33(1)**, 79-97.
- Lundtofte, Frederik, 2008. Expected Life-time Utility and Hedging Demands in a Partially Observable Economy. *European Economic Review* **52(6)**, 1072-1096.
- Luo, Yulei, 2010. Rational Inattention, Long-run Consumption Risk, and Portfolio Choice. *Review of Economic Dynamics* **13(4)**, 843-860.
- Luo, Yulei, 2017. Robustly Strategic Consumption-Portfolio Rules with Informational Frictions. *Management Science* **63(12)**, 4158-4174.
- Luo, Yulei and Eric R. Young, 2010. Risk-sensitive Consumption and Savings under Rational Inattention. *American Economic Journal: Macroeconomics* **2(4)**, 281-325.
- Luo, Yulei and Eric R. Young, 2016. Long-run Consumption Risk and Asset Allocation under Recursive Utility and Rational Inattention. *Journal of Money, Credit and Banking* **48(2-3)**, 325-362.
- Merton, Robert C., 1969. Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case. *Review of Economics and Statistics* **51(1)**, 247-257.
- Mondria, Jordi, 2010. Portfolio Choice, Attention Allocation, and Price Comovement. *Journal of Economic Theory* **145(5)**, 1837-1864.
- Peng, Lin and Wei Xiong, 2006. Investor Attention, Overconfidence and Category Learning. *Journal of Financial Economics* **80(3)**, 563-602.
- Powell, Warren B., 2007. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*, Wiley Series in Probability and Statistics, John Wiley and Sons.

- Saint-Paul, Gilles, 2010. A “Quantum” Approach to Rational Inattention. *CEPR Discussion Papers* 7739.
- Samuelson, Paul A., 1969. Lifetime Portfolio Selection by Dynamic Stochastic Programming. *Review of Economics and Statistics* **51(3)**, 239-46.
- Shafieepoorfard, Ehsan and Maxim Raginsky, 2013. Rational Inattention in Scalar LQG Control. *Proceedings of the IEEE Conference on Decision and Control* **52**, 5733-5739.
- Shannon, Claude, 1948. A Mathematical Theory of Communication. *The Bell System Technical Journal* **27**, 623-56.
- Sims, Christopher A., 2003. Implications of Rational Inattention. *Journal of Monetary Economics* **50(3)**, 665-90.
- Sims, Christopher A., 2005. Rational Inattention: A Research Agenda. Technical Report, Princeton University.
- Sims, Christopher A., 2006. Rational Inattention: Beyond the Linear-Quadratic Case. *American Economic Review* **96(2)**, 158-163.
- Tutino, Antonella, 2012. Rationally Inattentive Consumption Choices. *Review of Economic Dynamics* **16(3)**, 421-439.
- Wang, Neng, 2009. Optimal Consumption and Assets Allocation with Unknown Income Growth. *Journal of Monetary Economics* **56(4)**, 524-34.
- Wachter, Jessica A. and Motohiro Yogo, 2010. Why Do Household Portfolio Shares Rise in Wealth? *Review of Financial Studies* **23(11)**, 3929-3965.
- Van Nieuwerburgh, Stijn and Laura Veldkamp, 2010. Information Acquisition and Under-Diversification. *Review of Economic Studies* **77(2)**, 779-805.