CESifo WORKING PAPERS

7413
2018

Original Version: December 2018
This Version: July 2019

Optimal Amount of Attention to Capital Income Risk and Heterogeneous Precautionary Saving Behavior

Penghui Yin

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7413*
Category 6: Fiscal Policy, Macroeconomics and Growth

Optimal Amount of Attention to Capital Income Risk and Heterogeneous Precautionary Saving Behavior

Abstract

This paper studies heterogeneous precautionary saving behavior under rational inattention. In the model, consumers face uninsured capital income risk, have CRRA preferences, and suffer from an information-processing capacity constraint. For given attention devoted to capital income risk, I solve for the optimal consumption-saving choices and show that the expected utility from consumption is increasing with the amount of attention. Furthermore, I solve for the optimal attention choice and find that households with more initial wealth would pay less attention to capital income risk. As a result, wealthier households have higher perceived uncertainty in their future capital income and display precautionary motive by saving at higher rates. I also provide empirical evidence to support these results by using data from 2016 wave of the Survey of Consumer Finances.

JEL-Codes: E130, E210, D810, O160.

Keywords: consumption-saving decision, information-processing constraint, capital-income risk.

Penghui Yin
Department of Money and Macroeconomics
Goethe University Frankfurt / Germany
Penghui.Yin@hof.uni-frankfurt.de

I am very grateful to Michael Evers and Mirko Wiederholt for their guidance and support. I also thank Klaus Adam, Stefania Albanesi, Chris Carroll, Martin Ellison, Elisabeth Falck, Michael Haliassos, Michael Hasler, Marten Hillebrand, Joe Kaboski, Ken Kasa, Markus Kontny, Hayk Kamalyan, Yulei Luo, Kristoffer Nimark, Filip Matějka, Giuseppe Moscarini, Evi Pappa, Petr Sedláček, Zexi Sun, Hélène Rey, Laura Veldkamp, Johannes Wohlfart, Jiankang Zhang, Peter Zorn, as well as seminar and conference participants at Goethe University Frankfurt, York University, Renmin University of China, Vanderbilt University, University of Hohenheim, ifo Institute Munich, and China Central University of Finance and Economics for helpful comments and suggestions. All remaining errors are my own.

1 Introduction

Understanding precautionary saving behavior is of vital importance in answering why the rich save so much (see e.g. Carroll (2000) and Dynan et al. (2004)). Standard economic models with decreasing absolute risk aversion preferences show that wealthier households tend to be more risk tolerant and save less for precautionary motive (see, e.g. Kimball (1990a,b)). However, this contradicts to the empirical fact as shown in Figure 1, which shows that households with more wealth, on average, have a larger desired precautionary savings to permanent income ratio[1]. This paper solves this puzzle by introducing rational inattention into a consumption-saving model with capital income risk and investigating the effect of initial wealth on the optimal attention choice and further on the precautionary saving behavior.

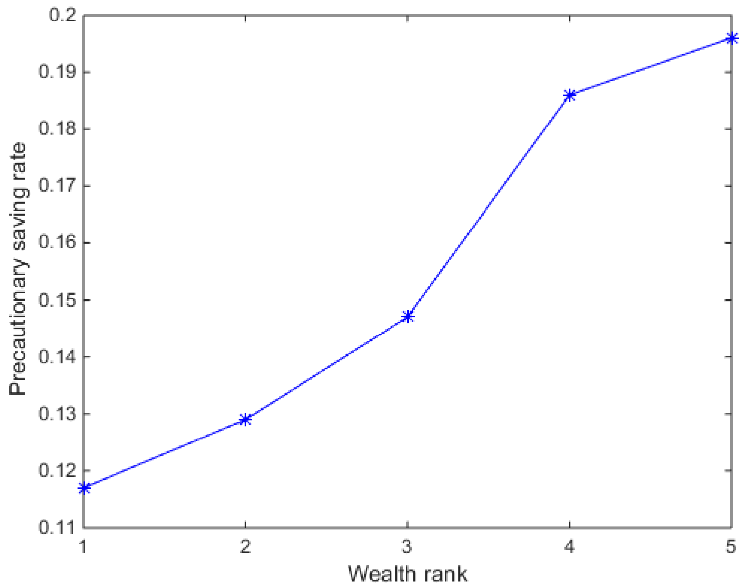


Figure 1: Ratio of desired precautionary savings over permanent income: 2016 wave of Survey of Consumer Finances (SCF)

Although heterogeneous saving behavior has been investigated in models with labor income risk[2], households still take capital income risk into consideration when making consumption-saving decisions. In the U.S., 49% of households invest directly or indirectly in stocks and investment funds, which have highly nondiversified risk (see Bertaut and Starr-McCluer (2002)). In addition, private businesses account for about 50% of aggregate capital and employment in

[1] Following Kennickell and Lusardi (2005) and Deidda (2013), precautionary saving rate is defined as the ratio of desired amount of savings for unexpected emergencies over permanent income. Wealth is defined as the sum of net worth and current income as in Kimball (1990b).

[2] See, e.g. Hubbard et al. (2000), Carroll (2000), Dynan et al. (2004), and etc.

the U.S., and they are subject to uninsurable idiosyncratic capital income risk (see Angeletos (2007)). Moreover, previous studies on temporal resolution of uncertainty, such as Epstein (1980), argue that acquiring more information regarding future capital return affects agents' current consumption choice and increases their welfare. Weil (1990) refers the income effect of an increase in the risk of future capital return on saving rate as the precautionary saving effect. However, little is known about what determines investors' information choice, and how information choice is connected with consumption-saving choice. Since understanding information acquisition is crucial to understanding households' consumption-saving behavior, this paper applies the theory of rational inattention by Sims (2003) to investigate both information acquisition behavior and consumption-saving choices under capital income risk.

In the model, agents choose the amount of attention to future capital return before making consumption-saving decisions subject to a cost of attention. On the one hand, by paying more attention, agents can observe more precise signal, which helps them to make consumption decisions closer to those under perfect foresight. On the other hand, paying more attention causes higher information cost in utility due to the attention capacity constraint. Since agents have scarce consumption resources and limited attention capacity, they need to solve two optimization problems: a consumption choice and an attention allocation problem. A two-period framework with stochastic capital return and a Gaussian signal allows us to solve these two problems analytically. The main findings are as follows.

First, I explicitly show that expected utility from consumption is increasing with the amount of attention devoted to capital return in a Gaussian framework. The intuition is that, by devoting more attention to capital return, agents can observe more precise signal to make better forecast of future capital return. This benefits agents by allowing them to make more efficient intertemporal consumption plans, which are closer to those under perfect foresight scenario. Second, I analytically show that the optimal amount of attention devoted to capital income risk is decreasing with initial endowment. The CRRA utility implies decreasing absolute risk aversion, or increasing absolute risk tolerance, which implies that the poorer households are less tolerant to future capital income risk. This is also the main contribution of my study as it first investigates the role of inequality in attention allocation problem. It can explain why the rich save at a higher rate through a precautionary saving channel. Households with more wealth pay less attention to future capital return and have higher perceived uncertainty in their future capital income, and as a result they will save more to against the potential loss in their consumption tomorrow.

My model also confirms several common themes in the rational inattention literature. First, the optimal amount of attention devoted to capital income risk is decreasing with the marginal cost of information. Information cost can be interpreted as the opportunity cost of devoting some attention to capital return. As a result, agents pay less attention to capital return if paying attention is more costly. Second, the optimal amount of attention devoted to capital income risk is increasing with the prior volatility of capital return. The intuition is that higher prior volatility in capital return makes paying attention to this term more necessary. Considering a counter example, agents do not need to acquire any information if they learn from their prior knowledge that future capital return is not volatile at all. Third, I find that the optimal amount of attention devoted to capital income risk is increasing with the degree of patience. A larger time discount factor makes future consumption more valuable to households today, and therefore, more attention will be paid to future capital income that finances all their consumption tomorrow. Finally, when relaxing the assumption on the ex-post normal distribution, my numerical results show discrete consumption-saving choices under costly information acquisition and bounded prior on capital return[3].

I also provide empirical evidence for my model results above. By using data from 2016 wave of the SCF, I find that: (1) the likelihood of paying attention to investment is negatively correlated with wealth/net worth and information delegation but positively correlated with financial literacy; (2) households with more wealth/net worth are more likely to feel uncertain about their future income, but paying attention to their investment can help to reduce the uncertainty; (3) households with more wealth/net worth desire higher precautionary saving rates.

This paper builds on Sims (2003, 2006). Sims (2003) argues that consumers cannot attend perfectly to all freely available information when decide their life-time consumption plans. He proposes modeling consumers' limited attention as a constraint on information flow, which is measured by entropy as in Shannon (1948)'s information theory. Sims (2006) studies a simple two-period saving problem with labor income risk and CRRA utility. He finds that for higher risk aversion degree, agents pay more attention to a mismatch between consumption and initial wealth when wealth is low[4]. However, a two-period model with capital income risk allows us to solve for both consumption-choice and portfolio choice problem analytically. Moreover, such model also enables us to study the role of wealth inequality in attention choice[5].

[3]Similar results can be found in previous studies such as Sims (2006), Matějka (2016), Jung et al. (2019), and Ellison and Macaulay (2019).

[4]Following Sims, rational inattention is popularly applied to solve consumption-saving problems. For example, Luo (2008) solves excess smoothness puzzle and excess sensitivity puzzle.

[5]Maćkowiak and Wiederholt (2015) also study a stochastic capital income in their model, but different from

This paper also builds on Van Nieuwerburgh and Veldkamp (2010). The authors solve a static portfolio choice and information choice problems analytically with CARA and CRRA preferences respectively. Similar to their work, agents in my paper also allocate limited attention to process information regarding the stochastic capital return. However, different from their study that utility comes from terminal wealth, in this paper agents make optimal intertemporal consumption allocation under limited attention. Furthermore, I not only solve the model analytically under assumption on Gaussian signal as in this paper, but also check the robustness of main findings numerically without Gaussian assumption.

My work contributes to the literature on rational inattention consumption by studying attention allocation behavior analytically in a non-linear setup. For the reason of tractability, linear quadratic utility or log utility are popularly used in rational inattention studies[6]. However, when investigating saving behavior with capital income risk neither linear quadratic nor log utility can be employed. The former utility specification implies certainty equivalence and the latter one implies that income and substitution effects cancel out. On the contrary, in this paper I can obtain analytical solutions to the consumption-saving choice and attention allocation problems with a stochastic interest rate and CRRA preferences in a Gaussian framework.

It also contributes to the literature studying consumption choices with stochastic capital return. Phelps (1962) first studies optimal saving behavior under capital income uncertainty by employing different utility functions. Levhari and Srinivasan (1969), Sandmo (1970), and Rothschild and Stiglitz (1971) extend the work by Phelps and find that assuming CRRA preferences, a degree of relative risk aversion larger than unity leads to positive effects of mean-preserving increases of riskiness in capital income on demand for savings. A common assumption in these papers is that future capital return is not observable and agents are endowed with a prior belief about the distribution of capital return. In these models, the effect of capital return risk on consumption-saving choice depends on exogenous changes in the prior volatility of capital return with the expected return remaining constant. Epstein (1980) extends these models above by allowing agents to learn about future capital return. Although he solves optimal consumption-saving choice subject to improving information about future capital return, the amount of information is still exogenous to agents. In this paper, I go one step further and allow agents to choose not only how much to consume but also how much information to acquire

my paper, the authors solve the RI model numerically under linear quadratic framework.

[6]For example, Sims (2003), Luo (2008) and Luo and Young (2016) directly work with linear quadratic framework. Luo (2010) works with log-quadratic approximation of CRRA utility. Sims (2006) works with a log utility.

regarding future capital income.

The remainder of this paper is organized as follows. Section 2 presents the two-period RI consumption-saving model. Section 3 solves consumption-saving and attention allocation problems analytically. Section 4 extends the model into a non-Gaussian framework in order to check robustness of the main finding. Section 5 concludes.

2 Information flow and information cost

In this section, I show some important terms in the information theory, such as information channel, entropy, Shannon capacity, information cost, which are the key elements in the RI model that is shown in the next section.

2.1 Noisy information channel

What is a channel in information theory? Cover and Thomas (2006) defines a discrete channel to be a system consisting of an input X and output Y, as well as a probability transition matrix $p(Y = y|X = x)$ that expresses the probability of observing the output symbol y given that we send the symbol x. Sims (2015) provides an example of noisy Gaussian channel as follows. The input of a channel is an arbitrary real number x drawn from a distribution $N(\mu, \sigma_x)$, and the transmission is contaminated by an independent noise ψ distributed $N(0, \sigma_\psi)$. Then if the conditional distribution of the output signal y is $y|x \sim N(x, \sigma_\psi)$, I call it a noisy Gaussian channel.

The Gaussian noisy channel can be equivalently defined by the following signal structure:

$$y = x + \psi$$

where x is a value drawn from a standard normal distribution and ψ is the noise term with also a normal distribution $N(0, \sigma_\psi)$.

From the example above, we can see that the channel only defines the conditional distribution of output given an input. Then we need to quantify the information flows through a channel. According to information theory, we can use the concept of entropy to characterize the uncertainty of a random variable, and the information flow is quantified as reduction in entropy.

2.2 Quantifying information flow

Consider again the example shown above. The entropy of the random variable x is [7]

$$H(x) = \frac{1}{2}\log(2\pi e\sigma_x^2)$$

where σ_x is the variance of x.

Conditional entropy measures the conditional uncertainty, i.e. given another random variable y, and the conditional entropy of $x|y$ is,

$$H(x|y) = \frac{1}{2}\log(2\pi e\sigma_{x|y}^2)$$

where $\sigma_{x|y}^2$ is the conditional variance of x with y given.

Then we can quantify the information flow that one random variable contains about another variable as the reduction in uncertainty:

$$\begin{aligned} I(x;y) &= H(x) - H(x|y) \\ &= \frac{1}{2}\log(2\pi e\sigma_x^2) - \frac{1}{2}\log(2\pi e\sigma_{x|y}^2) \\ &= \frac{1}{2}\log\left(\frac{\sigma_x^2}{\sigma_{x|y}^2}\right) \end{aligned}$$

where $I(x;y)$ is called mutual information, which shows the expected reduction in entropy of x from seeing y.

2.3 Information cost

In my model , agents make the information choice subject to a cost of information processing. In this case the information is freely available to the individual, however, since the capacity to process information is limited, agents have to decide which dimension of uncertainty to control, and therefore, the cost of information is the shadow price of capacity in agents' overall optimization problem[8].

Following Sims (2006), I also assume that in the daily life of my agent there are many other decisions to make in addition to the consumption-saving choice, and as a result, the

[7] Sims (2003) states that the logarithm in the formula can be to any base, because the base only determines a scale factor for the information measure, but conventionally it takes the logarithm to base 2, and as a result the entropy of a discrete distribution with equal weight on two points is 1 or $(-0.5\log(0.5) - 0.5\log(0.5))$, which is the unit of information called a "bit". When the base is e, the unit of information is a "nat".

[8] Sims (2006) also lists two other types of cost. One is physical cost such as costs of telephone lines, Internet connections and other wiring that brings in information from the outside world. The other one cost is due to costly information investigation such as experimentation and surveying exploration.

information processing cost can be interpreted as an opportunity cost of devoting some of the scarce attention resource to the underlying optimization problem.

3 Two-period consumption-saving model

In this section, I incorporate information constraint into a standard two-period consumption-saving model.

Agents. In the model, agents live only two periods $t \in \{0, 1\}$.

Preference. $u(C_t)$ is period t utility. I assume that the utility function belongs to the CRRA family,

$$u(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma} \tag{1}$$

where $\gamma > 1$ parametrizes the degree of relative risk aversion.

Time discounting factor. Let $\beta \in (0, 1)$ denote the agents' subjective discount factor. It represents the degree of agents' patience.

Budget constraints. Let W_0 be the initial endowment[9] which is strictly positive and known in the initial period. In period 0, agents make consumption and saving decisions and savings consist of a single asset, called capital, which is used to produce the consumption good Y_1, which will be consumed in period 1. Then, the period budget constraints in period 0 and 1 are respectively,

$$C_0 + K_1 = W_0 \tag{2}$$

$$C_1 = Y_1 \tag{3}$$

where C_0 and C_1 are consumption in period 0 and 1 respectively. K_1 is the total savings/investment of agents' in period 0, which will be used as the producing inputs in period 1.

Technology and productivity shock. Each agent owns a firm with the production function:

$$Y_1 = A_1 K_1 \tag{4}$$

where A_1 is productivity or return to capital, and K_1 is the capital used for producing consumption goods Y_1. Productivity follows the process:

$$A_1 = \exp(\epsilon_1) \tag{5}$$

[9]This can be interpreted as the present value of the riskless lifetime labor income. Correspondingly, the period 0 is the period when agents work with labor income, and period 1 is the period that agents have no labor income and consume from savings.

where ϵ_1 is the exogenous productivity shock[10].

Agents do not observe the true realization of the shock, but they are endowed with a prior belief about the distribution from which the productivity disturbance is drawn: $\epsilon_t \sim N(-\frac{1}{2}\sigma^2, \sigma^2)$.

Signal structure. I assume that agents learn the exogenous productivity shock by observing a signal[11]:

$$s_0 = \epsilon_1 + \psi_0 \tag{6}$$

where the signal is noisy but unbiased, and $\psi_0 \sim N(0, \sigma_\psi^2)$ is the endogenous noise caused by finite capacity. The variance of signal is $\sigma^2 + \sigma_\psi^2$, and therefore, the precision of the signal is defined $\frac{1}{\sigma^2+\sigma_\psi^2}$. Given prior belief on ϵ_1, the signal precision is only determined by the noise variance σ_ψ^2.

Bayes' Law. Agents use Bayes' Law to combine their prior belief and the observed signal in (6) such that $\epsilon_1|s_0 \sim N(\hat{\epsilon}_1, \hat{\sigma}^2)$, where

$$\hat{\epsilon}_1 \equiv E[\epsilon_1|s_0] = \frac{(-\frac{1}{2}\sigma^2)\sigma_\psi^2 + \sigma^2 s_0}{\sigma^2 + \sigma_\psi^2} \tag{7}$$

$$\hat{\sigma}^2 \equiv Var[\epsilon_1|s_0] = \frac{\sigma^2\sigma_\psi^2}{\sigma^2 + \sigma_\psi^2} \tag{8}$$

where $\hat{\epsilon}_1$ is the posterior mean. Notice that the posterior variance $\hat{\sigma}^2$ is determined by the prior variance σ^2 and noise variance σ_ψ^2. Given prior belief, (8) implies that the signal precision can be uniquely determined by the posterior variance. In addition, decreasing posterior variance is equivalent to increasing signal precision because they both are results of reducing noise variance σ_ψ^2.

Information set. Let us now define information sets before and after observing the signal, which are called stage 1 and stage 2 of period 0.

Definition 1 *$\mathfrak{I}^1$ and $\mathfrak{I}^2$ are the information sets in stage 1 and stage 2 respectively.*

$$\mathfrak{I}^1 = \left\{ W_0, \epsilon_1 \sim N\left(-\frac{\sigma^2}{2}, \sigma^2\right), \hat{\epsilon}_1 \sim N\left(-\frac{\sigma^2}{2}, \sigma^2 - \hat{\sigma}^2\right) \right\}$$

$$\mathfrak{I}^2 = \mathfrak{I}^1 \cup \{s_0\}$$

[10] A more complete structure should also include a non-learnable component in capital return as in Grossman and Stiglitz (1980). However, since removing this non-learnable part will not change my results, I only model a learnable component.

[11] Sims (2010) provides two ways to solve models with limited information-processing capacity. The first way is to solve the optimal joint distribution of the control variable and the unobservable state variable. The second way is to assume a signal structure, and then solve for the optimal policy as a function of signal. However, as argued by Sims (2010), the optimal joint distribution can be characterized by many different combinations of signal structure and policy function. I will mainly use the first method to solve for attention choice and consumption-saving choice problems separately and check the robustness of my main result in Section 5.

*where the ex-ante distribution of the posterior mean $\hat{\epsilon}_1$ is derived in the appendix **A.1**. $\mathfrak{I}^1$ implies that agents know their initial endowment, prior beliefs about the distribution of productivity shock and noise in signal.*

Information constraint. According to the RI literature, information constraint is defined as the maximum reduction in uncertainty, which is measured by entropy. Denote $f(\epsilon_1)$ the PDF of the exogenous disturbance, and $f(\epsilon_1|s_0)$ as the conditional PDF of exogenous disturbance conditional on the received signal. The information processing capacity constraint is

$$\begin{aligned} I(\epsilon_1; s_0) &= H(\epsilon_1) - H(\epsilon_1|s_0) \\ &= \frac{1}{2}\log\left(\frac{\sigma^2}{\hat{\sigma}^2}\right) = \kappa \end{aligned}$$

where $H(\epsilon_1)$ is the entropy of productivity shock and $H(\epsilon_1|s_0)$ is the conditional entropy of productivity shock given signal observation;$I(\epsilon_1; s_0)$ is also called the mutual information between productivity shock and signal observation and can be interpreted as how much information about ϵ_1 is contained in s_0; the equality implies that choosing the optimal signal precision is equivalent to allocating the optimal amount of attention κ to learning the capital return.

Optimization problems. In this model, agents need to solve not only a consumption-saving problem but also an information choice problem. Formally the whole optimization problem looks as follows:

$$V = \max_{\{\hat{\sigma}^2\}} E_{\mathfrak{I}^1}[u(C_0^*) + \beta u(C_1^*)] - \lambda\kappa \tag{9}$$

Subject to

$$C_0^* = \arg\max_{C_0} E_{\mathfrak{I}^2}[u(C_0) + \beta u(C_1)]) \tag{10}$$

$$C_1^* = A_1(W_0 - C_0^*) \tag{11}$$

$$\frac{1}{2}\log\left(\frac{\sigma^2}{\hat{\sigma}^2}\right) = \kappa \tag{12}$$

where equation (9) is the objective function for the consumer; $E_{\mathfrak{I}^2}[\cdot]$ is the expectation conditional on the information set $\mathfrak{I}^2$ and $E_{\mathfrak{I}^1}[\cdot]$ is the expectation over all possible signals[12]; λ is the marginal information cost, which can be interpreted as the opportunity cost of devoting some of the scarce resource attention to the stochastic capital return or equivalently to ϵ_1; budget

[12]This idea is close to the study by Dréze and Modigliani (1972) who discuss consumption decisions under timeless uncertainty prospects, and in this case agents optimally select consumption after the resolution of uncertainty, whereas in my model due to limited information processing capacity stochastic capital return is only partially revealed at the initial period.

constraints are incorporated into equations (10) and (11); equation (12) is the information constraint. Later I will show that this strictly positive marginal cost guarantees that the capacity is always finite, and consequently, the signal cannot perfectly reveal future capital return.

4 Analytical solution with Gaussian signal

In this section I solve for optimal consumption-saving and attention allocation decisions under the assumption of a noisy Gaussian signal.

4.1 Solution approach and equilibrium

Given a noisy Gaussian signal, I follow Maćkowiak and Wiederholt (2009) and Van Nieuwerburgh and Veldkamp (2010) and divide the whole optimization problem of period 0 into two stages as shown in the timeline in appendix **A.2** (Figure 10), which displays the information sets and actions on each stage of the initial period.

- Stage 1 optimization: given the initial information set, preference, capacity constraint and the information cost, agents optimally choose how to allocate their information capacity by choosing the optimal posterior variance, or equivalently the noise variance that determines the signal precision.

- Stage 2 optimization: given the signal realization and the optimal signal precision obtained on the first stage, agents now progress to the main consumption-saving decision.

I solve the two-period model with capital income risk backward[13]. For a given signal realization and signal precision, I can solve for the optimal consumption-saving choice, which depends on the posterior mean and variance of the productivity shock ϵ_1. Substituting optimal consumption and savings into the objective function delivers indirect utility as a function posterior variance $\hat{\sigma}^2$. For given signal realization, the problem looks as follows.

$$\max_{C_0, K_1, C_1} U(C_0) + \beta E_{\mathfrak{I}^2} U(C_1) \tag{13}$$
$$\text{s.t. budget constraints}$$

Solving the maximization problem (13) I will obtain the optimal consumption in each period denoted as C_0^* and C_1^*. In Section 3.2 below, I will show that the optimal consumption is a

[13]This method is popularly used in rational inattention literature, such as, Maćkowiak and Wiederholt (2009), Van Nieuwerburgh and Veldkamp (2010), and Matějka and McKay (2015).

function of the realized signal and posterior variance of the capital return, which are contained in the second stage information set $\mathfrak{I}^2$. Therefore, when plugging C_0^* and C_1^* back into the utility function, I have indirect utility as a function of the signal and posterior variance.

The second step is to solve the optimization problem on the first stage: agents choose the optimal signal precision by paying amount of attention κ. Formally agents solve:

$$\max_{\hat{\sigma}^2} E_{\mathfrak{I}^1}\left[U(C_0^*) + \beta E_{\mathfrak{I}^2} U(C_1^*)\right] - \lambda\kappa \tag{14}$$

subject to $C_0^*(\mathfrak{I}^2)$, $C_1^*(\mathfrak{I}^2)$ and

$$\frac{1}{2}\log\left(\frac{\sigma^2}{\hat{\sigma}^2}\right) = \kappa \tag{15}$$

where $\lambda > 0$ is the marginal cost that is associated with the information processing capacity constraint. The more attention is allocated to reducing productivity uncertainty, the higher the total cost $\lambda\kappa$ will be reduced from the utility.

Solving these two optimization problems delivers the equilibrium in this model.

Definition 2 *An equilibrium of this model is characterized by distributions for the posterior variance* $(\hat{\sigma}^2)^*$*, the consumption choices* $C_0^*(\mathfrak{I}^2), C_1^*(\mathfrak{I}^2)$*, and demand for savings* $K_1^*(\mathfrak{I}^2)$ *such that:*

1. *given* W_0*,* $f(\epsilon_1)$ *and signal observation* s_0*, agents choose optimal consumption and savings according to the maximization problem defined in* (13)
2. *and the optimal posterior variance according to the maximization problem defined in* (14).

4.2 Stage 2 solution: optimal consumption-saving choices

According to the two-step approach, the first step is to solve for the second optimization problem where the signal is already realized, and thus I can rewrite the optimization problem in (13) as:

$$\begin{aligned} U &= u(C_0) + \beta E[u(C_1)|s_0] \\ &= \frac{(W_0 - K_1)^{1-\gamma}}{1-\gamma} + \beta E\left[\left.\frac{(A_1 K_1)^{1-\gamma}}{1-\gamma}\right| s_0\right] \end{aligned} \tag{16}$$

F.O.C:

$$\frac{\partial U}{\partial K_1} = -(W_0 - K_1)^{-\gamma} + \beta E[(A_1 K_1)^{-\gamma} A_1 | s_0] = 0 \tag{17}$$

Solving equation (17), I get optimal consumption and savings conditional on the observed signal[14]:

$$K_1^* = \frac{1}{\left[\beta E[A_1^{1-\gamma}|s_0]\right]^{-\frac{1}{\gamma}} + 1} W_0 \tag{18}$$

$$C_0^* = \frac{1}{\left[\beta E[A_1^{1-\gamma}|s_0]\right]^{\frac{1}{\gamma}} + 1} W_0 \tag{19}$$

$$C_1^* = A_1 K_1^* \tag{20}$$

where $E[A_1^{1-\gamma}|s_0]$ is a signal extraction problem. For different capacity of processing information, agents also interpret the signal differently.

These three equations above imply that if agents pay finite amount of attention to productivity shock, a signal extraction problem or filtering problem $E[A_1^{1-\gamma}|s_0]$ is involved. With the Gaussian assumptions of prior belief and signal about future productivity disturbance, we can easily show that $\epsilon_1|s_0$ follows a normal distribution. And therefore $E[A_1^{1-\gamma}|s_0]$ follows a log-normal distribution.

Given the results in (7), (8), (18), and (19) I can rewrite the consumption rule as :

$$C_0^{\kappa<\infty} = \frac{1}{\left[\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right]^{\frac{1}{\gamma}} + 1} W_0 \tag{21}$$

and the savings decision is given by

$$K_1^{\kappa<\infty} = \frac{1}{\left[\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right]^{-\frac{1}{\gamma}} + 1} W_0 \tag{22}$$

By contrast, if agents pay no attention to the external disturbance, the noise variance will be infinitely large and agents will solve her consumption-saving problem only according to the prior knowledge with $\hat{\epsilon}_1 = -\frac{1}{2}\sigma^2$ and $\hat{\sigma}^2 = \sigma^2$, and consequently the consumption-savings decisions in this case are:

$$C_0^{\kappa=0} = \frac{1}{\left[\beta \exp\left(-\frac{\gamma(1-\gamma)}{2}\sigma^2\right)\right]^{\frac{1}{\gamma}} + 1} W_0 \tag{23}$$

And accordingly, the savings is

$$K_1^{\kappa=0} = \frac{1}{\left[\beta \exp\left(-\frac{\gamma(1-\gamma)}{2}\sigma^2\right)\right]^{-\frac{1}{\gamma}} + 1} W_0 \tag{24}$$

Results above teach us how perceived uncertainty affect consumption-saving decisions. First, when the relative risk aversion coefficient $\gamma = 1$, consumption-savings decisions in both cases

[14]These results are also obtained by Epstein (1980), Gollier (2001) and Miao (2004).

are not influenced by the information acquisition: $C_0 = W_0/(1+\beta)$. This is because the income effect and substitution effect of changing the perceived riskiness will off set with $\gamma = 1$. Furthermore, I find that the consumption (on average) under limited attention is higher than that under zero attention assuming the relative risk aversion coefficient is larger than unity[15]. This result directly shows the role of information constraint in consumer's precautionary saving behavior. Equations (23) and (24) are results that can also be found in the previous literature without learning capital return, such as Sandmo (1970).

4.3 Welfare implications of information processing constraints

In order to examine the welfare effects of capital income fluctuations under information processing constraint, I present how an increase in perceived uncertainty leads to changes in the ex-ante expected utility. Given equations (18)-(20) I can reorganize the expected utility function as:

$$E[U] = \frac{W_0^{1-\gamma}}{1-\gamma} E\left[\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}} + 1\right]^{\gamma} \tag{25}$$

A complete derivation of indirect utility function (25) can be found in appendix **A.4**. Rothschild and Stiglitz (1971) derive a similar expression, but the authors focus on the effect of changing (perceived) capital income risk on savings demand. Here I go one step further and study the effect of changing capital income risk on expected utility. However, calculating the expected utility in equation (25) analytically is not easy, and in the current paper I assume integer γ and provide an approach to analytically study the welfare effect of information constraint by applying the binomial theorem[16].

Proposition 1 *Given an integer degree of relative risk aversion larger than unity, the higher attention paid to future productivity disturbance the higher the expected utility from consumption.*

Proof for Proposition 1 can be found in appendix **A.6**. Although a comparison between expected utility under different timing of uncertainty resolution has been conducted in the literature such as Dréze and Modigliani (1972) and Epstein (1980), here I show an analytical

[15] I show the proof for this result in appendix **A.3**.

[16] The difficulty mainly comes from the expression in the expectation operator in equation (25). As $\hat{\epsilon}_1$ is a function of signal observation, and prior to the signal observation it is a random variable, such that $\left[\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}} + 1\right]^{\gamma}$ is a shifted log-normal variable with power γ. Suppose γ is not a integer, it is not possible to calculate the closed-form solution of its expected value. Appendix **A.5** shows the details of this approach.

comparative statics for the change in expected utility under RI. Proposition 1 suggests that by paying more attention to learning capital return, agents will obtain a more precise signal and therefore can make more efficient consumption-saving decisions. However, due to the assumption on the positive marginal cost of processing information, agents cannot pay infinite amount of attention to future return to savings. As a result, agents need to compare the benefit and cost of paying more attention and solve for the optimal capacity allocation to capital return. In the next subsection I will show how to solve the information choice problem and explain the determinants of attention allocation.

4.4 Stage 1 solution: the optimal information choice

The key mechanism to solve for the optimal attention allocation problem below is the trade-off of paying attention as shown in Section 4.3.

$$V = \max_{\hat{\sigma}^2} \frac{W_0^{1-\gamma}}{1-\gamma} E\left[\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}} + 1\right]^{\gamma} - \lambda\kappa \tag{26}$$

$$\frac{1}{2}\log\left(\frac{\sigma^2}{\hat{\sigma}^2}\right) = \kappa \tag{27}$$

This optimization problem is a well-posed mathematical problem with a concave objective function (26) and a convex constraint set (27). Details are shown in appendix **A.7**. The difficulty of solving this problem is to compute the expectation of indirect utility with non-integer or large relative risk aversion coefficient. To make progress in solving it, (a) I apply binomial theorem on indirect utility with integer $\gamma > 1$, and (b) I conduct comparative statics by applying implicit function theorem on the first order condition.

First, for integer relative risk aversion degree I can obtain the first order condition for the above maximization problem[17]:

$$\frac{W_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[\binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2 + \frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2 \sigma^2\right)\right]$$
$$\times \exp\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\left(1-\frac{q}{\gamma}\right)\right)\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\left(1-\frac{q}{\gamma}\right)\right) + \frac{1}{2}\lambda\frac{1}{\hat{\sigma}^2} = 0 \tag{28}$$

where the first part is the marginal benefit of paying attention to capital return in terms of consumption utility, and the second part is the marginal cost of paying one extra unit of attention to capital return.

Then, in the following comparative statics I show how optimal signal precision characterized by the first order condition above is affected by factors such as prior volatility, information cost,

[17] Details can be found again in appendix A.7.

initial endowment, and time discounting factor[18]. Let us start with the main contribution of the paper

Proposition 2 *Assuming integer $\gamma > 1$, the optimal amount of attention allocated to capital income is decreasing with the initial endowment W_0.*

Proof for this proposition can be found in **A.8**. The key mechanism through which initial endowment operates on attention allocation comes from assumptions on the utility function and marginal cost of paying attention. Agents in this model have CRRA-type preferences, which implies that absolute risk aversion is smaller when initial endowment is higher. A decreasing absolute risk aversion is equivalent to increasing absolute risk tolerance, meaning that agents with very large initial endowment would like to pay less attention to reduce the perceived riskiness in future consumption, or equivalently capital income. A more intuitive explanation is to compare the marginal benefit and the marginal cost of paying attention. The marginal cost λ is a positive constant number and identical to all agents. However, from the first order condition in equation (28) I can observe that the marginal benefit of paying attention in terms of consumption utility is lower for rich agents, and this decreasing benefit of paying attention over initial endowment can also explain the negative relationship between initial endowment and the optimal amount of attention devoted to capital income risk[19].

Proposition 3 *Assuming integer $\gamma > 1$, the optimal amount of attention allocated to capital income is increasing with the patience level denoted by β.*

Proof can be found in **A.9**. Agents with larger β value future utility more, and as a result a reduction in second-period income will negatively affect their welfare more. To avoid a reduction in the future, agents will pay more attention to the return to savings.

This paper also confirms two common themes in the RI literature regarding the effects of information processing cost and prior volatility on information choice[20].

[18]In the literature such as Sims (2006), Maćkowiak and Wiederholt (2009) and Van Nieuwerburgh and Veldkamp (2010), authors tend to discuss the optimal capacity allocation instead of the signal precision, and therefore, in the following text I use attention allocated to capital income risk to substitute the choice of signal precision.

[19]Maćkowiak and Wiederholt (2015) also study the attention allocation of heterogeneous households. However different from the present paper, their focus is on investigating changes in attention allocation to different sources of uncertainty under different income structures, which are measured by labor income to expenditure ratios.

[20]Maćkowiak et al. (2018) state another common theme in the RI literature that agents pays more attention to the variables that are most useful to them.

Proposition 4 *Assuming integer $\gamma > 1$, the optimal amount of attention allocated to capital income is decreasing with the information processing cost λ.*

Proof can be found in appendix **A.10**. The reason that this result is labeled as a common theme is because many other works in the RI literature also find the same result. For example, Wiederholt (2010) solves analytically for the attention choice in a static price setting RI model as well, and concludes that attention allocated to the unobservable total demand is negatively related with information cost. In addition, this finding can be also found in Tutino (2013) who quantitatively shows that when the shadow price of processing information is higher, agents pay less attention to the unobservable labor income history respectively.

Proposition 5 *Assuming integer $\gamma > 1$, the optimal amount of attention allocated to capital income is increasing with the volatility of the exogenous disturbance σ^2.*

Proof can be found in **A.11**. This finding is also in line with many other studies in the vein of RI, such as Maćkowiak and Wiederholt (2009), Van Nieuwerburgh and Veldkamp (2010), and Wiederholt (2010). These works try to tackle different economic problems but all realize that rationally inattentive agents would pay more attention to the variable with higher volatility. The intuition in this model is that for given marginal cost of attention, the higher the prior volatility, the higher the marginal benefit of paying attention.

4.5 Attention choice and precautionary saving effects

Here I show how to use my model to explain the heterogeneous precautionary saving behavior. From previous subsection, I notice that households with more wealth pay less attention to future capital return due to more risk tolerance. As a result, they have higher perceived uncertainty in their future capital return. For relative risk aversion degree larger than unity, I show in Figure 2 that the expected saving rate is increasing with initial wealth[21]. This is because an increase in perceived capital income risk leads to two opposite effects on saving demand, which are income effect and substitution effect.

According to Weil (1990), this income effect is simply a precautionary savings effect. A higher perceived capital income risk implies a higher probability of low consumption tomorrow. Wealthier agents will save more to against this higher perceived risk than poor agents who pay more attention to capital return.

[21] parameter values can be found in appendix C.1.

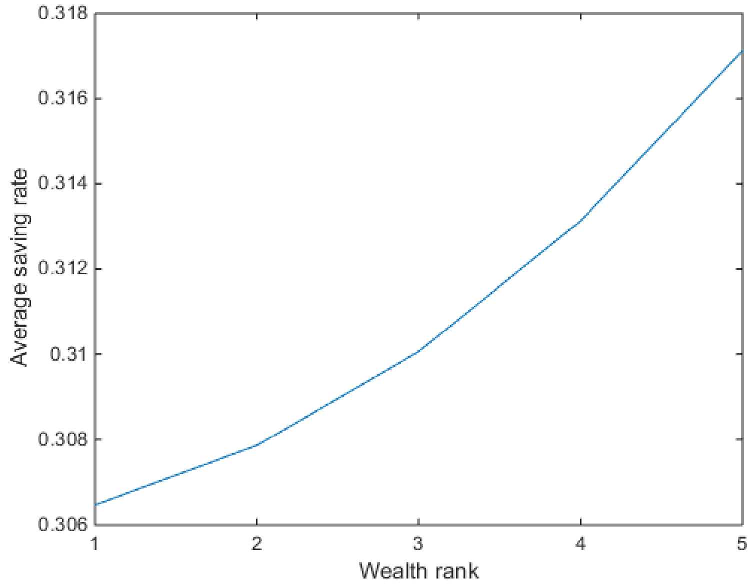


Figure 2: Optimal saving rate and initial wealth

4.6 Additional results

4.6.1 The effect of risk aversion on attention allocation

In the comparative statics studies, I cannot analytically show how relative risk aversion affects attention choice. Here I present a simulation result on the relation between attention allocated to capital income risk and relative risk aversion in Figure 3 (parameter values can be found in appendix C.2.). The intuition is that for higher relative risk aversion, agents dislike the uncertainty in future capital income more, and therefore, in order to reduce the uncertainty they would like to pay more attention to obtain a more precise signal on future capital return[22].

4.6.2 Comparing with rational inattention model with labor income risk

In this subsection, we will compare our model with the rational inattention model with labor income risk from two aspects.

- **Attention choice of heterogeneous agents.** In our model, the main contribution is to study the attention choice problem for heterogeneous agents with different initial endowment. We can show that the optimal amount of attention is decreasing with initial endowment analytically because initial endowment in our model is known when making decisions. However, this cannot be shown in a rational inattention model with labor income risk for the following reasons. First, a model with linear quadratic utility, such as

[22]Tutino (2013) find a similar result in a rational inattention consumption model with labor income risk.

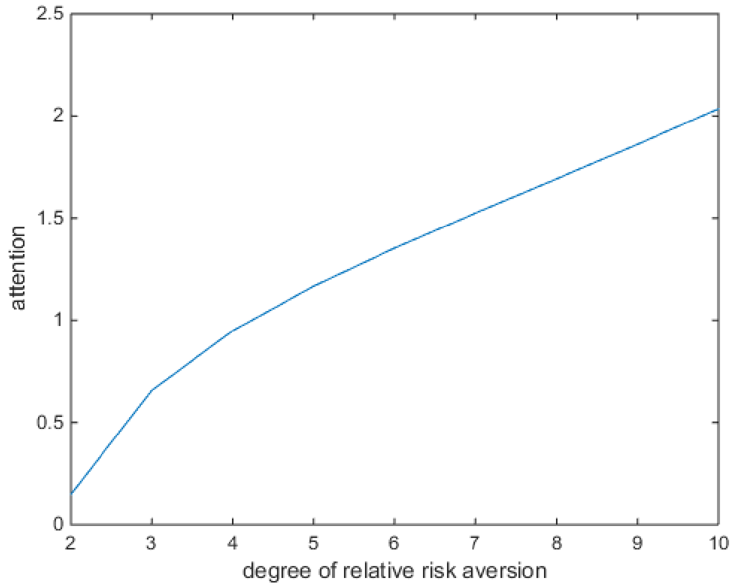


Figure 3: Relative risk aversion and attention

the one in Luo and Young (2016), the optimal amount of attention to labor income risk is independent from initial endowment due the certainty equivalence. Second, in rational inattention works with CRRA utility, such as Sims (2006), initial wealth is unobservable and the author solves for the optimal attention choice from the optimal joint distribution of consumption and initial wealth. Therefore, such framework is not suitable to study the relationship between initial wealth and the optimal attention choice.

- **Posterior normal distribution assumption.** As argued in Sims (2006, 2010), ex-post normal distribution is optimal in rational inattention models only under linear quadratic Gaussian setup. However, for the reason of tractability, the present paper and many studies in the rational inattention consumption-saving models with labor income risk, such as Luo et al. (2017), assume a normal ex post distribution. In order to show that our main findings are valid without such approximation, in the next section we will relax the assumption of ex post Gaussianity and solve the model numerically for an optimal joint distribution of the unobservable state variable and the control variable as in Sims (2006).

5 Moving beyond the Gaussian framework

Sims (2006) and Jung et al. (2019) argue that without linear quadratic utility, although the prior distribution of a state variable is normally distribute, under rational inattention the optimal posterior distribution is not normal or even not continuous. In such scenarios, it is difficult to solve my optimization problems analytically. More precisely, in my model I cannot solve the filtering problem in the ex-post first order condition in (17) any longer. In order to show the robustness of the relationship between initial endowment and the optimal amount of attention devoted to capital income risk, In this section we will solve the optimal joint density model numerically as follows.

5.1 Model

Following the model specification in Sims (2006), I form the optimization problem as the following[23].

$$\max_f \int_{0<K_1<W_0} \left(\frac{(W_0 - K_1)^{1-\gamma}}{1-\gamma} + \beta \frac{(A_1 K_1)^{1-\gamma}}{1-\gamma} \right) f(K_1, A_1) dK_1 dA_1 \tag{29}$$

subject to

$$1 \geq f(K_1, A_1) \geq 0 \tag{30}$$

$$\int_{0<K_1<W_0} f(K_1, A_1) dK_1 = g(A_1) \tag{31}$$

$$\begin{aligned}
&(\lambda :) \\
&\int_{0<K_1<W_0} \log(f(K_1, A_1)) f(K_1, A_1) dK_1 dA_1 \\
&- \int_0^\infty \left(\log \left(\int_{K_2}^\infty f(K_1, A_1) dA_1 \right) \int_{K_1}^\infty f(K_1, A_1) dA_1 \right) dK_1 \\
&- \int_0^\infty \log(g(A_1)) g(A_1) dA_1 = \kappa
\end{aligned} \tag{32}$$

where equation (29) implies that I am still maximizing expected utility, which is the sum of the expected utility in the initial period's consumption, $\frac{(W_0-K_1)^{1-\gamma}}{1-\gamma}$ and the discounted expected utility of next period's consumption, $\beta\frac{(A_1K_1)^{1-\gamma}}{1-\gamma}$. However, in this case the control variable is not consumption or attention as in Section 3, but the joint distribution $f(K_1, A_1)$. Equation (30) shows the boundaries of feasible values for probability densities, which are $f = 0$ and $f = 1$. Equation (31) constraints the marginal distribution of productivity (or capital return), and this constraint also implies that the optimal joint distribution $f(K_1, A_1)$ is optimal if agents' prior capital return is zero. Information constraint is equation (32), and $\lambda > 0$ is the marginal cost of using the capacity to process information.

[23]See Matějka (2016) for details regarding the equivalence of the joint distribution of savings and signal and the joint distribution of savings and productivity shock.

Similar to Sims (2010), in the optimization problem above, the objective function (29) is linear in the object of choice $f(A_1, K_1)$ and the information constraint is a concave function, which defines a convex feasible choice set. Therefore, there exists an optimal joint distribution that maximizes expected utility[24].

5.2 Numerical technique

In this two-period model, agents choose the optimal amount of attention to capital income risk only once. Therefore, we can easily solve this constrained optimization problem with value function iteration[25]. We first discretize the model with $n_{A_1} = 20$ equi-spaced grid points ranging from 0.8 to 1.2, and $n_{K_1} = 20$ equi-spaced grid points ranging from 10% to 75% of initial endowment W_0. By inserting the budget constraints (2) and (3) into the objective function (29), the optimization above becomes:

$$V = \max_{f(K_1,A_1)} \left[\sum_{A_1 \in \Omega_{A_1}} \left(\sum_{K \in \Omega_{K_1}} \frac{(W_0 - K_1)^{1-\gamma}}{1-\gamma} f(K_1, A_1) \right) + \beta \sum_{A_1 \in \Omega_{A_1}} \left(\sum_{K_1 \in \Omega_{K_1}} \frac{(A_1 K_1)^{1-\gamma}}{1-\gamma} f(K_1, A_1) \right) \right] \tag{33}$$

subject to:

$$1 \geq f(K_1, A_1) \geq 0 \quad \forall (K_1, A_1) \in B \tag{34}$$

$$\sum_{K_1 \in \Omega_{K_1}} f(K_1, A_1) = g(A_1) \tag{35}$$

$(\lambda :)$

$$\kappa = \sum_{A_1 \in \Omega_{A_1}} \sum_{K_1 \in \Omega_{K_1}} f(K_1, A_1) \left(\log \frac{f(K_1, A_1)}{\sum_{K_1 \in \Omega_K} f(K_1, A_1) * g(A_1)} \right) \tag{36}$$

where Ω_{A_1} is the space for productivity, and Ω_{K_1} is the space for savings choice, and $B \equiv \{(K_1, A_1) : K_1 < W_0\}$.

Then I build a prior belief simplex $g(A)$ for the capital return A_1, and it is constructed with uniform random sample, whose sum is one. Next, I initialize the joint distribution $f(K_1, A_1)$ based on the prior belief. Finally, I apply the iterative minimization method to make the value function converge (with about 200 iterations) and solve for the optimal joint density.

[24] For details, see Theorem 4 and its proof in Matějka (2010).

[25] See also Sims (2006) and Ellison and Macaulay (2019).

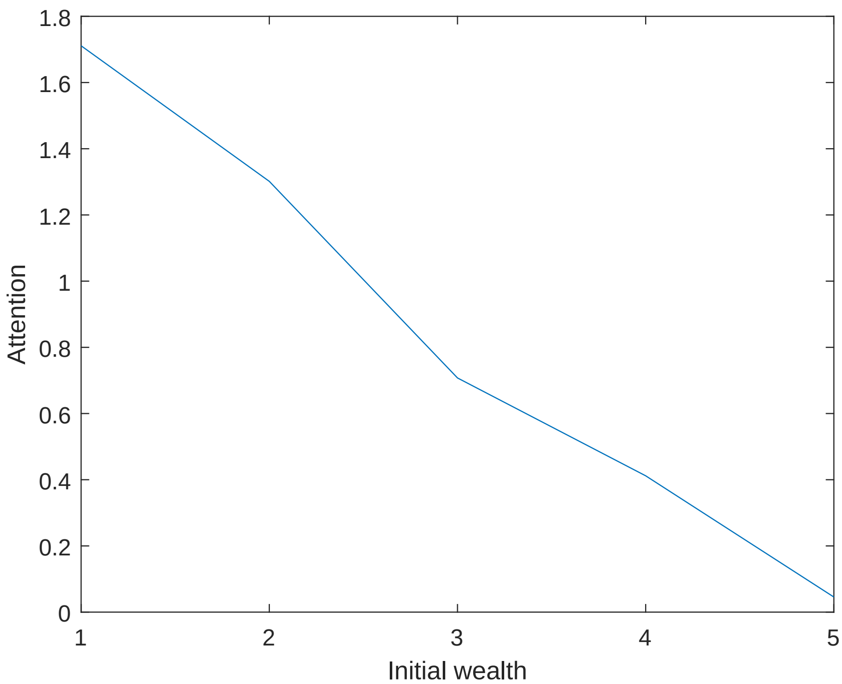
5.3 Main result

In this subsection, I will discuss the relationship between attention choice and initial wealth and joint behavior of consumption-saving and attention choice.

5.3.1 Initial wealth and attention choice

Figure 4 plots the numerical result of average amount of attention devoted to capital income risk for different levels of initial endowment. Especially, on the vertical axis, the average amount of attention is calculated by taking the average of optimal amount of attention for each point in the prior belief simplex. From this quantitative exercise, I can learn that even without making any assumption on the signal distribution, I find that the optimal attention choice is decreasing with W_0, which is in line with the result in the Gaussian case. I argue that the reasons for this results are the same as in the Gaussian case: under CRRA preferences agents have increasing absolute risk tolerance, and as a result the rich households would like to bear more risk in their future capital income. Equivalently, households with less initial endowment do not have so much to lose in future capital income, and they would like to pay more attention to make a better prediction of future capital return.

Figure 4: Attention choice and initial endowment in a non-Gaussian RI model



In order to check the robustness of the optimal joint distribution, I set different initial belief and solve the minimization problem again with the same numerical method. More precisely, I use the optimal joint distribution itself as the initial value and check whether it converges to a different optimal value. I find that it does not converge to a different joint distribution of productivity shock and investment choice. In appendix B, I also show the negative relationship between attention choice and initial wealth with gradient-based search method of Chris Sims.

Next, I discuss the model implications on heterogeneous attention choice behavior and consumption-saving behavior across agents with different initial wealth. Figures 5, and 6 plot the joint densities of saving rate and capital return in two different ways[26]. The left panel present the optimal decision rule in three-dimension meshes, whereas the right panel transform the left panel into two-dimension graph with colormap. First notice that, when marginal information cost $\lambda = 0.001$ these graphs show discretization, meaning that several values of saving rates receive zero probability. This is in line with findings in Sims (2006), Jung et al. (2019), Matějka (2016), and Ellison and Macaulay (2019). Without exposing ex-post normal distribution, agents restrict themselves to discrete choices due to the positive information cost.

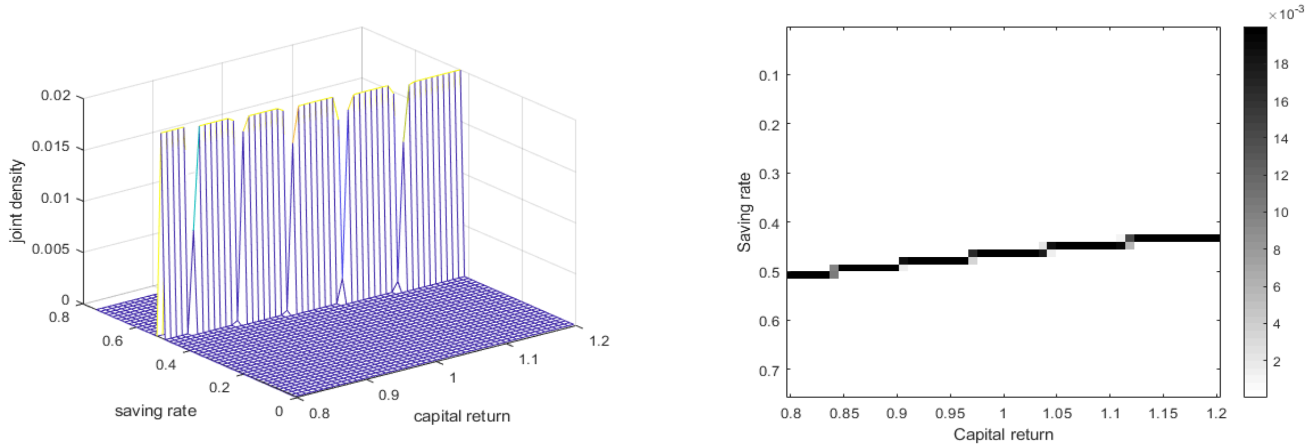


Figure 5: Low wealth ($W_0 = 1$), $\kappa^* = 0.5$

More importantly, these graphs show heterogeneity in the joint behavior between attention choice and consumption-saving problem. As already show in Figure 4, households with more initial wealth pay less attention to capital return. From Figures 5, and 6 I can also observe that, for given prior belief and information cost, households who pay more attention to capital return have more choices than those who pay less attention. This is because lower absolute risk tolerance motivates them to be well informed. In addition, they also allocate their optimal amount of attention to the consumption-saving choices differently. For example, households with low wealth pay more attention to the low (high) saving rate when saving return in high (low). Compare to poor households, people with high initial wealth pay less attention to low

[26] Calibration of parameters can be found in appendix C.3.

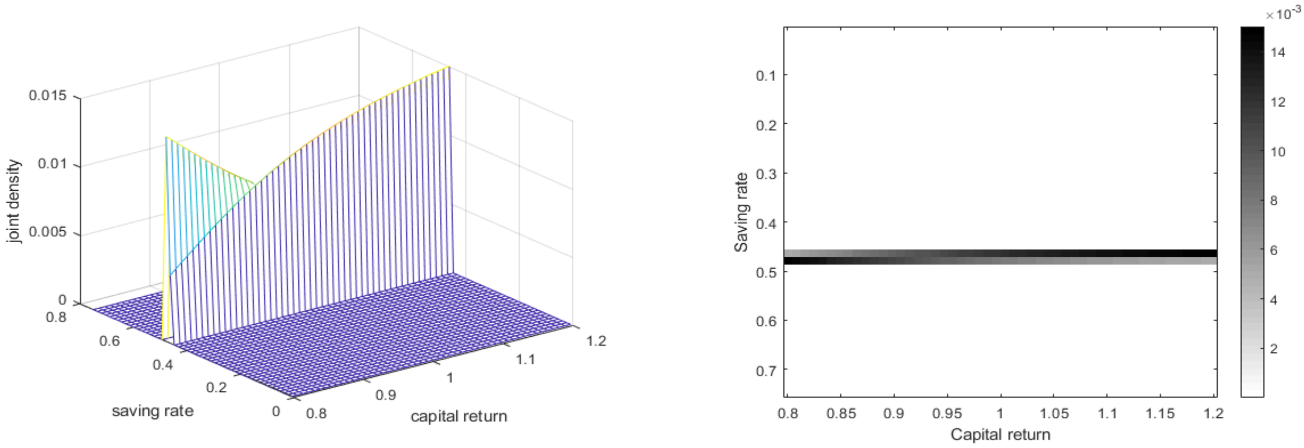
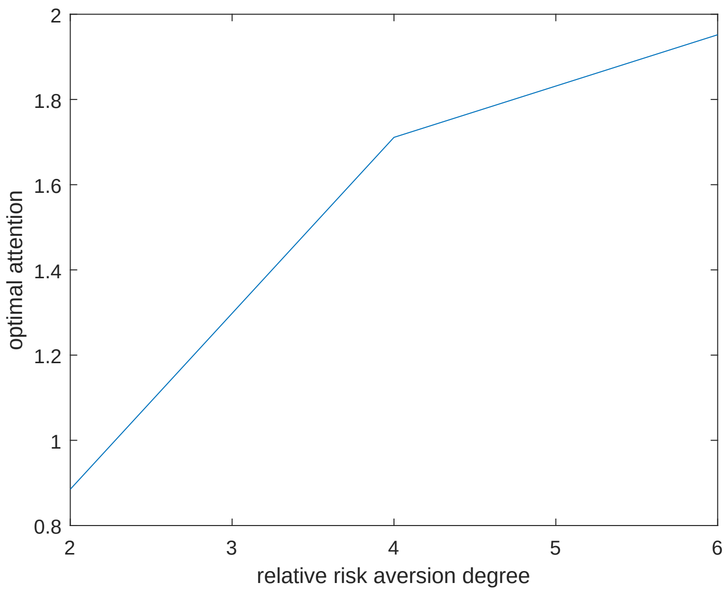


Figure 6: High wealth ($W_0 = 5$), $\kappa^* = 0.05$

saving rate and more attention to high saving rate. These results make sense, with the low wealth solution trading higher probabilities of choosing a low value of saving rate when capital return is in fact high for a lower probabilities of a mismatch between saving and capital return when capital return is low. This is supported by the empirical fact that the rich households on average save at higher rate (see e.g. Dynan et al. (2004)).

5.3.2 Relative risk aversion and attention choice

Figure 7: Attention choice and risk aversion in a non-Gaussian RI model



Similar to the analysis above, here I show how relative risk aversion affects the joint behavior of attention choice and consumption-saving choice. First I notice from Figure 7 that the optimal amount of attention to capital income risk is increasing with the degree of risk aversion. This result is similar to what I present in the Gaussian framework.

From Figure 8 and Figure 9, I can show that high risk averse agents have more choices than low risk averse agents because they pay more attention to future capital return. And I can also observe that low risk averse households put most weight on its lowest saving rate 0.45, whereas the more risk averse households put the most weight on its lowest saving rate 0.43. This is because more risk averse households worry more about fluctuation in their future income, and more risk averse households trade higher probabilities of choosing a lower saving rate when return is high for a lower probabilities of a mismatch between saving rate and capital return when capital return is low.

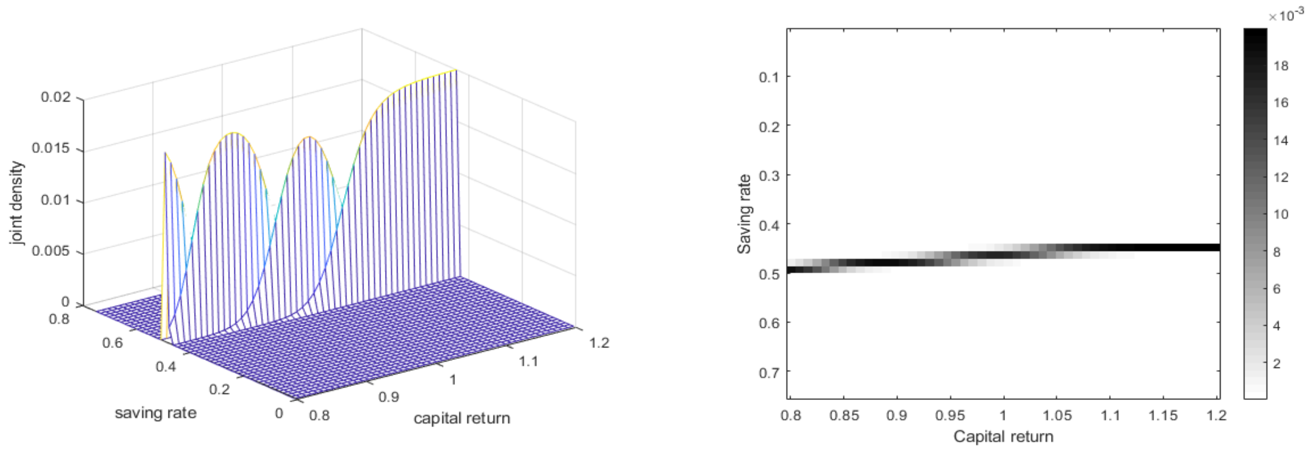


Figure 8: Low risk aversion ($\gamma = 2$), $\kappa^* = 089$

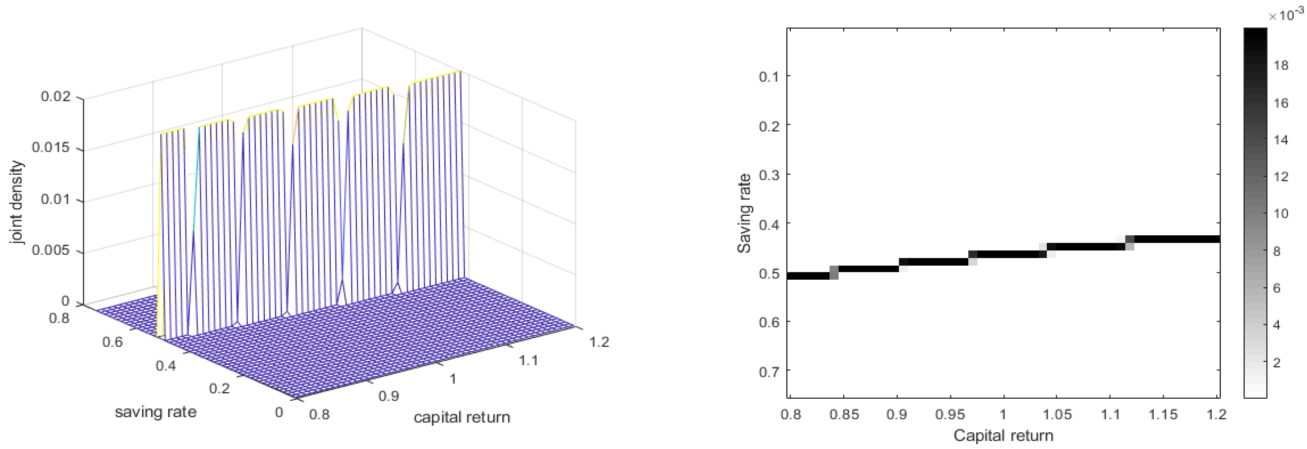


Figure 9: High risk aversion ($\gamma = 4$), $\kappa^* = 1.71$

6 Empirical evidence

The previous section links current saving rate with income inequality through the channel of attention allocation. In order to support this result, I provide empirical evidence on the relations between income and information acquisition in this section.

6.1 Bring the model to data

To test the model above, I look at attention allocation behaviors and saving rates across agents with different wealth. From propositions 4 and 5, I make the following two testable predictions:

Prediction 1: Households with more wealth pay less attention to capital income risk.

Prediction 2: Households with more wealth feel more uncertain about their future income.

Prediction 3: Wealthier households save at higher rate due to precautionary motivation.

These there hypotheses lead to three challenges, i.e. how to measure attention, posterior uncertainty and precautionary saving motive. First, attention is referring how much uncertainty agents can reduce by processing in order to make economic decisions, but rational inattention theory does not provide a clear way to measure uncertainty reduction. Second, posterior uncertainty about future income is a subjective measure, but it is hard to quantify even in the survey questions. Finally, precautionary saving motive is difficult to disentangle from other motives to save (see Gentry and Hubbard (2000).). In following three subsections I discuss how to deal with these difficulties.

6.2 Measure of attention

Despite the difficulty of measuring attention, few empirical works occurred in the previous literature on attention choice. Mondria et al. (2010) use data on search queries on the internet as the proxy variable of attention devoted to assets of different countries in order to explain home bias puzzle with RI. Recent studies, such as Roth and Wohlfart (2019) and Fuster et al. (2018), conduct survey experiments to investigate what factors determine individual's belief updating rate, which is used as a proxy of attention to study the effect of expectations with endogenous information acquisition on macroeconomic behavior. These two studies above show how prior belief, education (literacy), and etc. affect information acquisition, however, what is missing in these two papers is the role of wealth inequality in information choice.

In previous studies, people already tried to study empirically the relationship between wealth inequality and attention. For example, Sicherman et al. (2016) use online account logins as a proxy for investors paying attention to their financial account. They show that wealthier investors are more likely to engage in financial "ostrich behavior", which makes them to pay less attention to their financial holdings than poorer investors during market declines, because the poor individuals find it harder to suppress monetary concerns. Another work by Shah et al. (2018) studies money in the mental lives of the poor. The authors find that the poor think more carefully about the monetary opportunity cost of economic decisions because those households face more limited budget constraints and they have to pay more attention to opportunity costs in order to be less susceptible to potential losses.

Different from these studies, I measure paying attention to capital return by the information used for investing decisions. In the 2016 SCF (Survey of Consumer Finances) data, I can learn respondents have two ways to obtain information for investment: self information-processing (such as call around, internet, newspapers, TV) and information delegation (such as professional advisers (banker, broker), financial planner, and etc.). Similar to my model setup, here I also define paying attention as whether agents process information to make investing decisions. I choose this proxy variable of paying attention to capital income risk based on the fundamental assumption of rational inattention theory. Sims (2015) argues that agents do not pay full attention even to information that is easily attainable with a negligible cost, e.g. to the large amount of data on the internet, in finance magazines, and etc. If respondents use some of these information sources, I can say that they pay more attention to those who do not use any of those sources.

6.3 Measure of uncertainty

One way to measure the posterior uncertain is like in Roth and Wohlfart (2019), where the authors ask directly how sure respondents about their guess. In such situation, people who answer that they are very confident about their guess have less uncertainty than those who are not confident about their guess. In this paper, I use a similar measure, which is from survey question "do you have a good idea about your income in the next year?" According to the theory of rational inattention, agents who pay more attention to a term (here, capital return) have lower perceived uncertainty regarding this variable. Therefore, people pay more attention to capital income risk will have lower posterior uncertainty.

6.4 Measure of precautionary saving motive

Although savings due to precautionary motive are difficult to disentangle from other motives, I can elicit the amount of desired amount of precautionary savings from SCF data since the surveys of 1995 and onward. In the survey, respondents were explicitly asked to report how much savings they need for unanticipated emergencies and other unexpected things that may come up[27]. As argued in Kennickell and Lusardi (2005), one advantage of such question is that households were asked what they "need to have in savings", instead of what they have for that purpose.

According to my model setup in Section 3 and model results in Section 4, households save at different rates because of their various perceived uncertainty in future capital income. This means that precautionary savings in my model are motivated by the unexpected risks in the future. Therefore, in this paper I follow the study of Kennickell and Lusardi (2005) and investigate heterogeneous precautionary saving behavior by using SCF 2016 data.

6.5 Data

In 2016 SCF, about 6254 households in US are interviewed about their income, asset holding, debt holding, expenditure, saving motive, as well as ways to obtain information for borrowing and investment[28]. This dataset also includes respondents demographic characteristics, such as gender, age, education, financial literacy. According to my model setup, the uncertainty of income comes from investment, I focus on a subsample of high wealth rank (wealth quintile 3, 4 and 5) in my empirical exercise.

In order to measure attention to capital income risk, I create a dummy variable "attention" by using answers to the following question.

- *Information used for investing decisions: call around*
- *Information used for investing decisions: friends, material from work/business contacts*
- *Information used for investing decisions: internet/online service*
- *Information used for investing decisions: magazines, newspapers, books*

[27] This measure of precautionary saving motive can also be seen in other surveys, such as Dutch Socio-Economic Panel, German SAVE survey.

[28] I take out respondents whose annual income is less than 1000 dollar, and respondents are too old (above 90) or too young (below 30). As a result, I obtain data set with 5148 observations.

- *Information used for investing decisions: material in mail, tv, radio, advertisements, telemarketer*
- *Information used for investing decisions: self, shop around, other personal research*

The dummy variable "attention" equals one if respondents use at least one of these sources, zero otherwise. However, among individuals who process information, I unfortunately cannot measure how much information do they actually process.

In 2016 SCF, respondents were also asked:

- *At this time, do you have a good idea of what's your (family's) income for the next year will be?*

From answers to this question, I can create a dummy variable *high_uncertainty*, which takes value 1 if people answer **No**, and takes value 0 if they answer **Yes**.

In order to measure precautionary savings, I follow Kennickell and Lusardi (2005) and use the information from question:

- *About how much do you think you and your family need to have in savings for unanticipated emergencies and other unexpected things that may come up?*

Answers to this question directly quantify the amount of desired precautionary savings. Then I calculate precautionary saving rate, which is the ratio of desired precautionary savings over income[29].

One of the most important results in Section 4 is that households with higher initial wealth pay less attention to capital income risk. Therefore, my main explanatory variable is wealth that is defined by the wealth (sum of income and total net worth) and total net worth itself[30]. I also follow Kennickell and Lusardi (2005) by taking reported "normal" income as a measure of permanent income. The precautionary saving rate is defined as the ratio of desired precautionary savings over permanent income.

In the survey data I also find that some investors delegate information procession to financial planners and other professional advisers (e.g. accountant, lawyer, banker and etc.). Usually,

[29]Different from the work of Kennickell and Lusardi (2005) who focus on the amount of desired precautionary savings, here I study the heterogeneous precautionary saving rate across different wealth levels, as predicted by my model in Section 4.

[30]Total net worth includes financial asset/debt and non-financial assets/debt (house, mortgage debt, vehicles, debt on vehicles). In the appendix D.2 - D.4, I also check the robustness of the main results with only financial net worth.

I think that people who purchase information from these professional advisers know better about their investment returns (see Peress (2004).). Here I use information delegation as a proxy variable for low prior volatility of capital return.

Financial literacy is often used as a proxy of marginal cost of processing information (see Fuster et al. (2018).). According to Sims (2003, 2006), the marginal cost of paying attention is measured how costly of acquiring information, and obviously agents with higher financial literacy are easily access and process freely available information.

In addition, I also take the data of willingness of taking financial risk, age, gender, number of kids, marital status, occupation and etc. A summary of variables used in this paper can be seen in Table 4.

6.6 Estimation model

In order to test the hypotheses above, I consider empirical specifications as follows.

$$\begin{aligned}
\text{Atention} &= \beta_0 + \beta_1 * \log(\text{wealth}) + \beta_2 * \text{delegation} + \beta_3 * \text{financial literacy} \\
&\quad + \beta_4 * \text{low risk aversion} + \boldsymbol{\theta} * \textbf{Interactions} + \boldsymbol{\alpha X} \qquad (37) \\
\text{High uncertainty} &= \beta_0 + \beta_1 * \log(\text{wealth}) + \beta_2 * \text{delegation} + \beta_3 * \text{financial literacy} \\
&\quad + \beta_4 * \text{low risk aversion} + \boldsymbol{\theta} * \textbf{Interactions} + \boldsymbol{\alpha X} \qquad (38) \\
\text{Precautionary saving rate} &= \beta_0 + \beta_1 * \log(\text{wealth}) + \beta_2 * \text{delegation} + \beta_3 * \text{financial literacy} \\
&\quad + \beta_4 * \text{low risk aversion} + \boldsymbol{\theta} * \textbf{Interactions} + \boldsymbol{\alpha X} \qquad (39)
\end{aligned}$$

where **Interactions** in equation (37) refer to interaction terms of measures of rich with other explanatory variables, such as delegation, financial literacy, willingness of taking financial risks (low risk aversion). **Interactions** in equation (38) refer to interaction terms of attention and explanatory variables. **Interactions** in equation (39) refer to interaction terms of high uncertainty and measures of rich. $\boldsymbol{X}$ includes the demographic characteristics mentioned above.

6.7 Results

In this section, I will answer the following three questions by using my regression results. (i) Who are more likely to process information by themselves? (ii) Who feel more uncertain about their income next year? (iii) Who desire more precautionary savings for unexpected emergencies?

6.7.1 Information acquisition for investing decisions

Information acquisition (Paying attention) in my paper means that respondents reported that they process information regarding investment decisions by themselves. In order to investigate the heterogeneous precautionary saving behavior, I first show how the information acquisition behavior is correlated with wealth.

Results 1 *People with more wealth are less likely to acquire information regarding investment decisions. Wealthier people who delegate information procession to financial advisers are less likely to acquire information by themselves, whereas wealthier people who have more financial literacy are more likely to acquire information.*

More precisely, column 1 and 3 of Table 1 show regression results of "pay attention" on two different measures of wealth, information delegation, financial literacy. Here I control for log of direct risky asset holding because omitting it would cause bias: the amount of direct risky asset holding affects not only wealth and net worth but also the likelihood of paying attention to investment. Ten percent increase in wealth or net worth leads to 0.25 percentage point and 0.21 percentage point higher likelihood of acquiring information regarding investment decisions. Column 2 and 4 show similar results when including interaction terms of wealth/net worth with information delegation and financial literacy. Here I want to study the conditional effect of wealth/net worth on the likelihood of self information acquisition. Regressions 2 and 4 present that households with more wealth are still less likely to acquire information by themselves, however those who delegate information have even lower likelihood. On the contrary, rich people who have more financial literacy are more likely to pay attention to investment decisions. One reason is that, as predicted by my model in Section 4 and 5, poor households face more limited budget, which motivates them to pay more attention to their future income from investment. This result is also in line with findings of some previous studies. For example, Shah et al. (2018) study the mental lives of households with low income, and find that the poor pay more attention to monetary opportunity cost of economic decisions due to more sever scarcity of resource. Sicherman et al. (2016) investigate attention choice of investors and show that wealthy investors are more likely to engage in financial "ostrich behavior", meaning that they pay less attention to their asset holdings when market declines than poorer investors who cannot afford losses in their future capital income.

Now let us turn to the effects of information delegation and financial literacy on the likelihood of information acquisition. Results in Table 1 also lends supports to some other results in

Section 4, the optimal amount of attention to capital income risk is increasing with the prior volatility of capital return, but decreasing with the marginal cost of attention. From the first and third columns of Table 1, I first observe that households who delegate information to professional advisers are less likely to process information by themselves. One explanation is that, when delegating the information procession to professional advisers, investors have lower prior volatility in investment return, which discourage them to acquire information by themselves. Financial literacy is used to measure the marginal cost of processing information. Then we can also observe that people who have more financial literacy, or equivalently lower cost of processing information, are more likely to process information by themselves. The intuition is that people with higher financial literacy can access and process information about investment more easily and this is also in line with the rational inattention theory. In the second and fourth columns, we should calculate the marginal effects of information delegation and financial literacy because interaction terms of these two variables with wealth/net worth. In this sample, the mean of log (wealth) is 14.16 and the mean of log (net worth) is 13.84, we can easily show that the average marginal effect of information delegation on information acquisition is negative, whereas, the marginal effect of financial literacy on information acquisition is positive.

I then conduct several robustness checks, which are shown in appendix D.2. First, as shown in Fagereng et al. (2016), households with wealth below the median wealth level hold very small share of risky assets in their portfolio. Following this argument, I take the sub-sample of observations with wealth larger than the median level. In addition, my model predictions in Section 4 rely on the uncertainty from future capital return, therefore, I make another restrictions that households in this sub-sample have to hold some risky asset either directly or indirectly. Results in Table 5 also show that

I consider households with wealth/net worth larger than median levels no matter they (directly) hold risky assets or not. Table 5 and Table 6 show that in this case households with more wealth or net worth are less likely to acquire information by themselves. Those who delegate information to someone else and those who have lower financial literacy are less likely to pay attention to investment decisions. Second, I consider a different measure of wealth. In these exercises above, I use total net worth to measure wealth, but here I check the robustness by using financial net worth, which excludes assets and debts from housing and vehicles. From Table 7 and Table 8, we can conclude that baseline results in 1 are very robust with financial net worth.

Table 1: Relationship between wealth and attention devoted to capital return

The table presents coefficients from OLS regressions with the dummy variable of self information-procession as the dependent variable. Controls include direct risky asset holding, number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Wealth (cash-on-hand) is the sum of net worth and current income. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
log (wealth)	-0.0251***	-0.0518**		
	(-2.72)	(-2.15)		
log (net worth)			-0.0213**	-0.0444**
			(-2.43)	(-2.01)
Information delegation	-0.247***	0.194	-0.248***	0.164
	(-14.35)	(1.60)	(-14.41)	(1.46)
log (wealth) * delegation		-0.0299***		
		(-3.80)		
log (net worth) * delegation				-0.0283***
				(-3.83)
Financial literacy	0.0607***	-0.208	0.0606***	-0.180
	(3.25)	(-1.74)	(3.24)	(-1.63)
log (wealth) * financial literacy		0.0187**		
		(2.25)		
log (net worth) * financial literacy				0.0170**
				(2.19)
Constant	0.981***	1.384***	0.931***	1.284***
	(5.65)	(3.87)	(5.47)	(3.90)
N	1742	1742	1742	1742

6.7.2 Attention and uncertainty

This section tests the second hypothesis that the rich households are more likely to feel more uncertain about future income. According to my model, it is due to less attention paid to future capital return. In order to investigate the effect of information acquisition on uncertainty in future income, I run regression of a dummy variable of having no idea about future income on measures of wealth.

Results 2 *People who have more wealth or net wroth are more likely to feel uncertain regarding their future income, where the perceived uncertainty is measured by whether respondents have a good idea about their income next year or not.*

Column 1 and 2 of Table 2 show results of regressing likelihood of high perceived uncertainty on log of wealth and other explanatory variable by controlling direct risky asset holding and total risky asset holding. Column 3 and 4 of Table 2 show results of regressing likelihood of high perceived uncertainty on log of net worth and other explanatory variable by controlling direct risky asset holding and total risky asset holding. First, we can see that respondents with more wealth or net worth has positive effect are more likely to report that they do not have a good idea about their income next year. More precisely, 10 percent increase in wealth or net worth will cause about 0.2 percentage point in the likelihood of uncertainty. It is surprising that delegating information to financial advisers or professional planners has positive effect on the likelihood of uncertainty, however these coefficients are not significantly different from zero. We can also observe that investors with more financial literacy are more likely to know better about their future income although the coefficients are not statistically significant when controlling log of direct risky asset holding. According to the Proposition 4, I can argue that rich households feel more uncertain about their future income because they pay less attention to capital return, which has been supported by the previous section.

Deidda (2013) argues that, from Italian survey data, households who invest into risky assets require higher amount of precautionary savings than those who only invest into safe assets. The main channel provided in this paper is that people who hold risky asset have higher uncertainty and hence save more for precautionary motive. However, by using SCF data we do not find that households with more risky assets have higher perceived uncertainty regarding their future income. Instead, we can see that coefficients of either log (DEQ) or log (EQUITY) are negative, although the former ones are not statistically significant. My results suggest that investing into risky assets does not necessarily increase the perceived uncertainty in investors' minds, and paying less attention to capital return is a potential explanation why do the rich feel more uncertain in their future income.

Table 2: Relationship between wealth and high uncertainty regarding capital income

The table presents coefficients from OLS regressions with the dummy variable of high perceived uncertainty as the dependent variable. Controls include direct risky asset holdings, number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Wealth (cash-on-hand) is the sum of net worth and current income. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
log (wealth)	0.0220**	0.0224***		
	(2.39)	(3.21)		
log (net worth)			0.0197**	0.0200***
			(2.27)	(3.18)
log (DEQ)	-0.00619		-0.00549	
	(-0.96)		(-0.86)	
log (EQUITY)		-0.0110**		-0.0108**
		(-2.07)		(-2.04)
Information delegation	0.00209	0.000788	0.00256	0.000965
	(0.08)	(0.05)	(0.10)	(0.06)
Financial literacy	-0.0275	-0.0267**	-0.0275	-0.0267**
	(-1.43)	(-2.31)	(-1.43)	(-2.31)
Constant	0.338*	0.141*	0.379	0.174
	(1.73)	(0.60)	(1.95)	(0.76)
N	1738	3450	1738	3450

Similar to the previous section, I also conduct the following robustness checks for this empirical question. When I take the sub-sample of households that hold risky assets and wealth/net worth larger than median level, Table 9 presents that rich households are more likely to feel uncertain about their future income and acquiring information helps to reduce the likelihood of being uncertain. In some of these regressions we can see that the coefficient of financial literacy is negative and statistically significant. This indicates that households with more financially literacy can understand the economy better, and as a result they are more likely to have good idea about their future income. Then I use the financial net worth to measure wealth. Results in Table 10 shows that the households with more financial wealth or financial net worth are more likely to report that they do not have good idea about their income next year, indicating that they are more likely to have higher perceived uncertainty in future income.

6.7.3 Heterogeneous precautionary saving behavior

In the last two subsections I presented empirical evidence that respondents with more wealth are less likely to pay attention to investment decisions feel more uncertain about their future income. This provides us the first stage to investigate the relationship between wealth inequality and precautionary saving rate. Therefore, in this section I examine households' precautionary saving behavior with their wealth status and information acquisition. In order to show people with high uncertainty in their future income save more for precautionary usage, I interact this variable with several explanatory variables such as wealth/net worth, information delegation, financial literacy.

Results 3 *Precautionary saving rate is positively correlated with wealth/net worth and perceived uncertainty of future income.*

In the first two regressions that are shown in Column 1 and 2 of Table 3, I run regressions of the ratio of desired precautionary savings over permanent income on measures of wealth, and control variables. What we can learn from these results are that households with more wealth/net worth save more for precautionary motive out of their permanent income. Similarly, in the last two regressions, I use logarithm of the desired amount of precautionary savings as the dependent variable. We can observe that one percent increase in wealth/net worth will lead to a 0.65 and 0.59 percent increase in the desired amount of precautionary savings. This is in line with previous studies on precautionary saving behavior, such as Kennickell and Lusardi (2005), who find that people with more uncertainty in their future income save more for precautionary motive. Together with results in previous sections, this suggestive evidence can be interpreted as follows: households with more wealth are more uncertain about their future income, and they have larger motive to save for unexpected emergencies. However, I do not find significant effect of information delegation and financial literacy on either the ratio or the log of desired precautionary savings.

Table 3: Relationship between wealth and precautionary saving rate

The table presents coefficients from OLS regressions with the ratio of desired amount of precautionary savings over permanent income and the logarithm of desired amount of precautionary savings as the dependent variables respectively. Controls include direct risky asset holding, number of kids, age, male, marriage, occupation, year of education, posterior uncertainty, as well as a dummy variable of self-employment. Wealth (cash-on-hand) is the sum of net worth and current income. Dependent variable "Ratio" is the desired amount of savings divided by permanent income, which is measured by the income in "normal years". Dependent variable "log (pre_savings)" refers to the logarithm of the desired amount of precautionary savings. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1) Ratio	(2) Ratio	(3) log (pre_savings)	(4) log (pre_savings)
log (wealth)	0.0280*		0.651***	
	(1.82)		(21.71)	
log (net worth)		0.0351**		0.591***
		(2.40)		(20.55)
Information delegation	-0.0384	-0.0392	-0.0991	-0.0863
	(-0.67)	(-0.69)	(-1.37)	(-1.17)
Financial literacy	0.0242	0.0228	0.00143	-0.00152
	(0.71)	(0.67)	(0.03)	(-0.03)
Constant	0.0196	0.0439	0.929*	2.115***
	(0.05)	(0.10)	(1.83)	(4.15)
N	1734	1734	1734	1734

I also conduct following robustness checks for this part. First, I again take a sub-sample of households who hold risky assets either directly or indirectly and whose wealth is larger than the median level.

Here I again check robustness of the correlation between precautionary saving rate and uncertainty with two different samples. First, as shown in Table 14, among households whose wealth/net worth is above median level, a 10 percent increase in wealth/net worth leads to precautionary saving rate increases by about 0.25 percentage point. In addition, households desire even a higher precautionary saving rate if they feel uncertain about their future income. When considering all households with positive holding of risky assets, I still find more wealth/net worth has significantly positive effect on desired precautionary saving rate, but the effect of being uncertain is not significant.

Finally, I want to mention that in order to check the robustness of results from 2016 SCF data I also pooled data from another two survey waves in 2010 and 2013. The results are shown in appendix D.5. However, in surveys before 2016 there was no information about the financial literacy, which is used to measure the marginal cost of paying attention. In the exercises of D.5 I use education level as a proxy of financial literacy. After controlling the survey year, we can observe the stories that rich people are less likely to pay attention to investment decisions, more likely to feel uncertain about future income, save more for precautionary saving motivation are still supported by these pooled dataset.

7 Conclusion

This paper incorporates information processing constraint into a consumption-saving model of capital income risk. I find that when agents face capital income risk that affects their future income, their consumption-saving choices depend largely on how much attention devoted to capital return.

Then I study the effect of changing attention allocated to capital return on ex-ante expected utility and show that paying more attention benefits agents by allowing them to make more efficient consumption-saving decisions. However, agents cannot choose infinite capacity level due to the positive marginal cost of attention, and therefore, they need also solve for optimal information choice problem. In order to study determinants of attention allocation, I first apply binomial theorem on indirect utility and then apply implicit function theorem on the first order condition. Analytical results suggest that the optimal attention allocated to capital income is positively correlated with prior volatility and negatively correlated with initial endowment and information processing cost. Relations between attention allocation and information cost and prior volatility are the common themes in the RI literature, and I can easily understand that agents pay less attention to learning capital return when the opportunity cost of paying attention is higher or when agents learn from their prior knowledge that capital return tomorrow will not be very volatile.

Main contribution of this paper is to show the relationship between attention allocation and inequality, as well as its feedback effect on saving behavior leads to the fact that rich people save at a higher rate. Due to the constant marginal cost and decreasing marginal utility of paying attention, rich households benefit less from devoting attention to capital return. Moreover, CRRA utility implies increasing absolute risk tolerance, meaning that rich households would like bear more risk in their future capital income. This finding can also help to explain the puzzling empirical fact that rich households save at a higher rate (see, e.g. Dynan et al. (2004)). Due to rich households pay less attention to future capital return, they have higher uncertainty in their capital income, and the precautionary motivation makes them to save more. In order to show the robustness of the relationship between initial endowment and attention in a non-Gaussian framework, I relax the assumption on noisy Gaussian signal. Provided reasonable parameter values, the numerical result also shows a negative effect of initial endowment on the optimal attention choice. The numerical exercise also shows discrete choice of saving rate due to costly attention, which is similar to findings in Sims (2006) and Matějka (2016).

I also provide empirical evidence for the model predictions above. By using 2016 SCF data, I show that the likelihood of paying attention to investment decisions is negatively correlated with measures of wealth. Similar empirical findings can also be found in recent works, such as Sicherman et al. (2016) and Shah et al. (2018). But people who did not delegate information or those who have more financial literacy are more likely to acquire information regarding investment by themselves. The secodn empirical exercise shows that people with more wealth are more likely to have higher perceived uncertainty in their future income, which is in line with my model result that perceived variance in future capital return is higher for those who pay less attention to capital return. The last empirical exercise studies who save more for precautionary motive and it turns out that households with more wealth require higher amount of savings for unexpected emergencies. This result is robust by using either ratio of desired precautionary savings over permanent income or logarithm of desired amount of precautionary savings as dependent variable.

Finally, I want to point that there are still problems that cannot be solved in this paper. First, in such a two-period consumption saving model it is not possible to study the change in the amount of information on current consumption-saving choices, because in this paper all decisions are ex post decisions, i.e. after signal observation. Therefore, in future research I will extend this model into three or more periods in order to study such effect, as this is crucial to understand the consumption-investment behavior in a incomplete information model. Second, in the model households have only one asset to invest, however, in reality there are many different assets that households can invest in the financial market. Therefore, in future research, it would be interesting to disentangle the total savings into different assets, for example home asset and foreign asset, in order to study the international capital flow under information-processing constraint and how portfolio rebalance affects consumption-saving decisions. Third, the paper can be extended into a infinite-time horizon model with both capital income risk and labor income risk. However, such extensions require demanding numerical computation. Due to the main purpose of this paper is to show analytically the mechanism that leads to heterogeneous attention allocation behavior, I leave the extensions for the future research.

References

Angeletos, G.-M. (2007). Uninsured idiosyncratic investment risk and aggregate saving. *Review of Economic Dynamics*, 10:1–30.

Bertaut, C. and Starr-McCluer, M. (2002). Household portfolios in the united states. *In L. Guiso, M. Haliassos, and T. Jappelli, Editor, Household Portfolios MIT Press, Cambridge, MA.*

Carroll, C. (2000). Why do the rich save so much? *Does Altlas Shrug? The Economic Consequences of Taxing the Rich, Harvard University Press.*

Chew, S. H. and Epstein, L. (1989). The structure of preferences and attitudes towards the timing of the resolution of uncertainty. *International Economic Review*, 30(1):103–117.

Cover, T. and Thomas, J. (2006). Elements of information theory. *Wiley, Hoboken, NJ.*

Deidda, M. (2013). Precautionary saving, financial risk, and portfolio choice. *The review of income and wealth*, 59:133–156.

Dréze, J. and Modigliani, F. (1972). Consumption decisions under uncertainty. *Journal of Economic Theory*, 5:308–335.

Dynan, K. E., Skinner, J., and Zeldes, S. (2004). Do the rich save more. *Journal of Political Economy*, 112(2):397–444.

Eeckhoudt, L. and Schlesinger, H. (2008). Changes in risk and the demand for saving. *Journal of Monetary Economics*, 55:1329–1336.

Ellison, M. and Macaulay, A. (2019). Rational inattention unemployment trap. *CEPR Discussion Paper No.13761.*

Epstein, L. (1980). Decision making and the temporal resolution of uncertainty. *International Economic Review*, 21(2):269–283.

Fagereng, A., Guiso, L., Malacrino, D., and Pistaferri, L. (2016). Heterogeneity and persistence in returns to wealth. *NBER Working paper 22822.*

Fuster, A., Perez-Truglia, R., Wiederholt, M., and Zafar, B. (2018). Expectations with endogenous information acquisition: an experimental investigation. *Manuscript.*

Gentry, W. and Hubbard, R. (2000). Entrepreneurship and household saving. *NBER Working Paper No.7894*.

Gollier, C. (2001). The economics of risk and time. *The MIT Press*.

Grossman, S. and Stiglitz, J. (1980). On the impossibility of informationally efficient markets. *The American Economic Review*, 70(3):393–48.

Hubbard, R., Skinner, J., and Zeldes, S. (2000). The importance of precautionary motives in explaining individual and aggregate saving. *Carnegie-Rochester Conference on Public Policy*, 40.

Hurst, E., Lusardi, A., Kennickell, A., and Torralba, F. (2010). The importance of business owner in assessing the size of precautionary savings. *The Review of Economics and Statistics*, 92(1).

Jung, J., Kim, J., Matějka, F., and Sims, C. (2019). Discrete actions in information-constrained problems. *Review of Economic Studies*, 0:1–25.

Kennickell, A. and Lusardi, A. (2005). Disentangling the importance of the precautionary saving motive. *CFS Working paper No. 2006/15*.

Kimball, M. (1990a). Precautionary saving and the marginal propensity to consume. *NBER Working Paper 3403*.

Kimball, M. (1990b). Precautionary saving in the small and in the large. *Econometrica*, 58:53–73.

Levhari, D. and Srinivasan, T. (1969). Optimal savings under uncertainty. *Review of Economic Studies*, 36(2):153–163.

Luo, Y. (2008). Consumption dynamics under information processing constraint. *Review of Economic Dynamics*, 11(2):366–385.

Luo, Y. (2010). Rational inattention, long-run consumption risk, and portfolio choice. *Review of Economic Dynamics*, 13:843–860.

Luo, Y. (2017). Robustly strategic consumption-portfolio rules with informational frictions. *Management Science*, 63(12):4158–4174.

Luo, Y., Nie, J., Wang, G., and Young, E. (2017). Rational inattention and the dynamics of consumption and wealth in general equilibrium. *Journal of Economic Theory*, 172:55–87.

Luo, Y. and Young, E. (2010). Risk-sensitivity consumption and savings under rational inattention. *American Economic Journal: Macroeconomics*, 2(4):281–325.

Luo, Y. and Young, E. (2016). Induced uncertainty, market price of risk, and the dynamics of consumption and wealth. *Journal of Economic Theory*, 163:1–41.

Maćkowiak, B., Matějka, F., and Wiederholt, M. (2018). Dynamic rational inattention: Analytical results. *Journal of Economic Theory*.

Maćkowiak, B. and Wiederholt, M. (2009). Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803.

Maćkowiak, B. and Wiederholt, M. (2015). Business cycle dynamics under rational inattention. *Review of Economic Studies*, 82(4):1502–1532.

Matějka, F. (2010). Rationally inattention: prices and information in economics. *Ph.D. dissertation*.

Matějka, F. (2016). Rationally inattentive seller: Sales and discrete pricing. *Review of Economic Studies*, 86(3):1125–1155.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–298.

Miao, J. (2004). A note on consumption and savings under knightian uncertainty. *Annals of Economics and Finance*, 5:299–311.

Mondria, J., Wu, T., and Zhang, Y. (2010). The determinants of international investment and attention allocation: using internet search query data. *Journal of International Economics*, 81(1):85–95.

Peng, L. and Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, 80:563–602.

Peress, J. (2004). Wealth, information acquisition and portfolio choice. *The Review of Financial Studies*, 17(3):879–914.

Phelps, E. (1962). The accumulation of risky capital: a sequential utility analysis. *Econometrica*, 30(4):729–743.

Quadrini, V. (2000). Entrepreneurship, savings and social mobility. *Review of Economic Dynamics*, 3:1–40.

Roth, C. and Wohlfart, J. (2019). How do expectations about the macroeconomy affect personal expectations and behavior? *The Review of Economics and Statistics, forthcoming.*

Rothschild, M. and Stiglitz, J. (1971). Increasing risk ii: Its economic consequences. *Journal of Economic Theory*, 3:66–84.

Sandmo, A. (1970). The effect of uncertainty on saving decisions. *The Review of Economic Studies*, 37(3):353–360.

Shah, A., Zhao, J., Mullainathan, S., and Shafir, E. (2018). Money in the mental lives of the poor. *Social Cognition*, 36(2):4–19.

Shannon, C. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423.

Sicherman, N., Loewnstein, G., Seppi, D., and Utkus, S. (2016). Financial attention. *The Review of Financial Studies*, 29(4):863–897.

Sims, C. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50:665–690.

Sims, C. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review*, 96(2):158–163.

Sims, C. (2010). Rational inattention and monetary economics. *Benjamin J. Friedman, Michael Woodford (Eds.), Handbook of Monetary Economics*, pages 155–181.

Sims, C. (2015). Rational inattention and monetary economics. *Handbook of Monetary Policy.*

Tutino, A. (2013). Rational inattentive consumption choices. *Review of Economic Dynamics*, 16:421–439.

Van Nieuwerburgh, S. and Veldkamp, L. (2010). Information acquisition and under-diversification. *Review of Economic Studies*, 77:779–805.

Weil, P. (1990). Nonexpected utility in macroeconomics. *The Quarterly Journal of Economics*, 105(1):29–42.

Wiederholt, M. (2010). Rational inattention. *Prepared for the New Palgrave Dictionary of Economics.*

Appendix

Appendix A

A.1. Derivation of unconditional mean and variance of posterior expectation $\hat{\epsilon}_1$

Distribution of posterior mean $\hat{\epsilon}_1$

First, according to the Bayesian updating rule I have

$$\hat{\epsilon}_1 \equiv E[\epsilon_1|s_0] = \frac{-\frac{1}{2}\sigma^2\sigma_\psi^2 + \sigma^2 s_0}{\sigma^2 + \sigma_\psi^2}$$
$$\hat{\sigma}^2 \equiv Var[\epsilon_1|s_0] = \frac{\sigma^2\sigma_\psi^2}{\sigma^2 + \sigma_\psi^2}$$

Second, according to the prior belief $\epsilon_1 \sim N(-\frac{1}{2}\sigma^2, \sigma^2)$ and signal structure $s_0 = \epsilon_1 + \psi_0$, with $\psi_0 \sim N(0, \sigma_\psi^2)$, I can obtain $s_0 \sim N(-\frac{1}{2}\sigma^2, \sigma^2 + \sigma_\psi^2)$. Therefore,

$$\begin{aligned} E[\hat{\epsilon}_1] &= \frac{-\frac{1}{2}\sigma^2\sigma_\psi^2 + \sigma^2 E[s_0]}{\sigma^2 + \sigma_\psi^2} \\ &= \frac{-\frac{1}{2}\sigma^2\sigma_\psi^2 + \sigma^2\left(-\frac{1}{2}\sigma^2\right)}{\sigma^2 + \sigma_\psi^2} \\ &= -\frac{1}{2}\sigma^2 \end{aligned}$$

and,

$$\begin{aligned} Var[\hat{\epsilon}_1] &= \frac{(\sigma^2)^2 Var[s_0]}{\left(\sigma^2 + \sigma_\psi^2\right)^2} \\ &= \frac{(\sigma^2)^2\left(\sigma^2 + \sigma_\psi^2\right)}{\left(\sigma^2 + \sigma_\psi^2\right)^2} \\ &= \frac{(\sigma^2)^2}{\sigma^2 + \sigma_\psi^2} \\ &= \sigma^2 - \hat{\sigma}^2 \end{aligned}$$

A.2. Timeline of agents' optimization problem

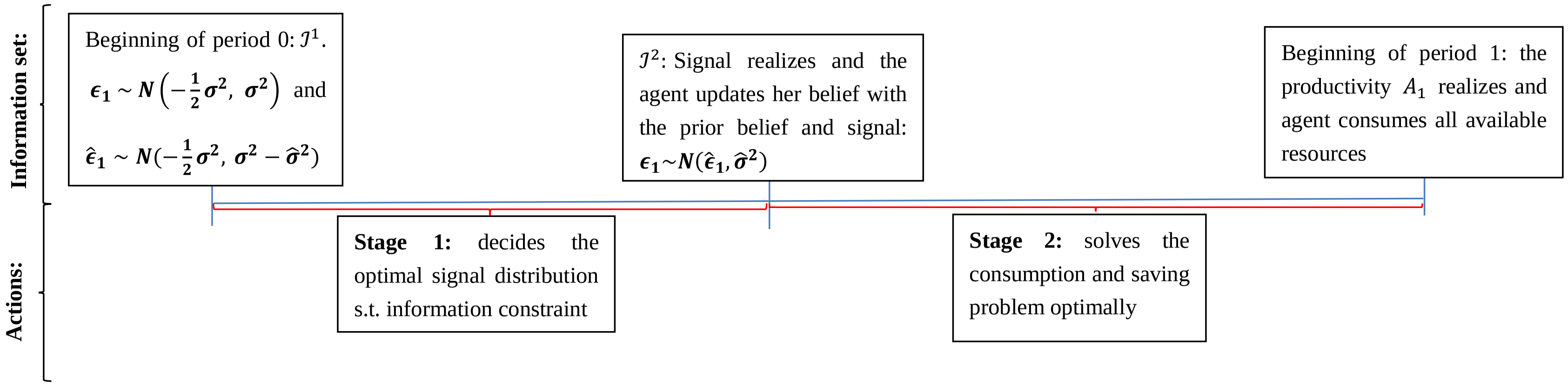


Figure 10: timeline for a two-period consumption-saving problem

A.3. Proof of the relation between perceived risk and unconditional expectation of savings for $\gamma > 1$:

Proof. If denote $\mathfrak{Y} = \left[\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right]^{\frac{1}{\gamma}} + 1$ then $\frac{1}{\mathfrak{Y}}$ is a convex function on the range of $\mathfrak{Y} \in (0, \infty)$. Therefore, Jensen's inequality can show

$$\begin{aligned}
E[C_0^{\kappa<\infty}] &= E\left[\frac{1}{\left[\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right]^{\frac{1}{\gamma}} + 1} W_0\right] \\
&\geq \frac{1}{E\left[\left[\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right]^{\frac{1}{\gamma}} + 1\right]} W_0 \\
&= \frac{1}{\beta^{\frac{1}{\gamma}} \exp\left(-\frac{\sigma^2}{2}\frac{1-\gamma}{\gamma} + \frac{1}{2}\frac{(1-\gamma)^2}{\gamma}\hat{\sigma}^2 + \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)^2(\sigma^2 - \hat{\sigma}^2)\right) + 1} W_0 \qquad (40) \\
&> \frac{1}{\beta^{\frac{1}{\gamma}} \exp\left(\frac{\gamma-1}{2}\sigma^2\right) + 1} W_0 = C_0^{\kappa=0} \\
\implies E[C_0^{\kappa<\infty}] &> C_0^{\kappa=0} \qquad (41)
\end{aligned}$$

where equation (40) is derived from equations (7) and (8) and $\hat{\epsilon} \sim N\left(-\frac{1}{2}\sigma^2, \sigma^2 - \hat{\sigma}^2\right)$.

To prove the inequality in (41) holds, I only need to show

$$-\frac{\sigma^2}{2}\frac{1-\gamma}{\gamma} + \frac{1}{2}\frac{(1-\gamma)^2}{\gamma}\hat{\sigma}^2 + \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)^2(\sigma^2 - \hat{\sigma}^2) < \frac{\gamma-1}{2}\sigma^2$$

And it is true because when $\kappa \in (0, \infty)$ I always have $\hat{\sigma}^2 < \sigma^2$ and with assumption $\gamma > 1$ then

$$\begin{aligned}
&-\frac{\sigma^2}{2}\frac{1-\gamma}{\gamma} + \frac{1}{2}\frac{(1-\gamma)^2}{\gamma}\hat{\sigma}^2 + \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)^2(\sigma^2 - \hat{\sigma}^2) + \frac{1-\gamma}{2}\sigma^2 \\
&= -\frac{\sigma^2}{2}\frac{1-\gamma}{\gamma}(1-\gamma) + \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)^2(\sigma^2 - \hat{\sigma}^2) + \frac{1}{2}\frac{(1-\gamma)^2}{\gamma}\hat{\sigma}^2 \\
&= -\frac{1}{2}\frac{(1-\gamma)^2}{\gamma}\sigma^2 + \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)^2(\sigma^2 - \hat{\sigma}^2) + \frac{1}{2}\frac{(1-\gamma)^2}{\gamma}\hat{\sigma}^2 \\
&= -\frac{(1-\gamma)^2}{2\gamma}(\sigma^2 - \hat{\sigma}^2) + \frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)^2(\sigma^2 - \hat{\sigma}^2) \\
&= -\frac{1}{2}\left(\frac{1-\gamma}{\gamma}\right)^2(\sigma^2 - \hat{\sigma}^2)(\gamma - 1) < 0
\end{aligned}$$

■

A.4. Derivation of the indirect utility function in stage 1:

First, denote $\phi = \frac{1}{\left[\beta E[A_1^{1-\gamma}|s_0]\right]^{\frac{1}{\gamma}}+1}$.

Then, given the equations (18)-(20) I can rewrite the expected utility function as:

$$\begin{aligned} E[U] &= E\left\{u(C_0^*) + \beta E[u(C_1^*)|s_0]\right\} \\ &= E\left\{\frac{(\phi W_0)^{1-\gamma}}{1-\gamma} + \beta E\left[\frac{(A_1(1-\phi)W_0)^{1-\gamma}}{1-\gamma}|s_0\right]\right\} \\ &= E\left\{\frac{(\phi W_0)^{1-\gamma}}{1-\gamma} + \beta\frac{((1-\phi)W_0)^{1-\gamma}}{1-\gamma}E[A_1^{1-\gamma}|s_0]\right\} \end{aligned} \tag{42}$$

If I denote $X = \beta E[A_1^{1-\gamma}|s_0]$,

$$\begin{aligned} \phi &= \frac{1}{\left[\beta E[A_1^{1-\gamma}|s_0]\right]^{\frac{1}{\gamma}}+1} \\ &= \frac{1}{X^{\frac{1}{\gamma}}+1} \end{aligned}$$

and

$$\begin{aligned} \phi^{1-\gamma} &= \left(\frac{1}{X^{\frac{1}{\gamma}}+1}\right)^{1-\gamma} \\ &= \left(X^{\frac{1}{\gamma}}+1\right)^{\gamma-1} \end{aligned}$$

$$\begin{aligned} (1-\phi)^{1-\gamma} &= \left(\frac{X^{\frac{1}{\gamma}}}{X^{\frac{1}{\gamma}}+1}\right)^{1-\gamma} \\ &= \left(X^{\frac{1}{\gamma}}+1\right)^{\gamma-1}X^{\frac{1-\gamma}{\gamma}} \end{aligned}$$

therefore, I can now write the indirect utility function

$$\begin{aligned} E[U] &= E\left\{\frac{(W_0)^{1-\gamma}}{1-\gamma}\left[\phi^{1-\gamma} + \beta(1-\phi)^{1-\gamma}E[A_1^{1-\gamma}|s_0]\right]\right\} \\ &= E\left\{\frac{W_0^{1-\gamma}}{1-\gamma}\left[\left(X^{\frac{1}{\gamma}}+1\right)^{\gamma-1} + \left(X^{\frac{1}{\gamma}}+1\right)^{\gamma-1}X^{\frac{1-\gamma}{\gamma}}X\right]\right\} \\ &= E\left\{\frac{W_0^{1-\gamma}}{1-\gamma}\left[\left(X^{\frac{1}{\gamma}}+1\right)^{\gamma-1} + \left(X^{\frac{1}{\gamma}}+1\right)^{\gamma-1}X^{\frac{1}{\gamma}}\right]\right\} \\ &= E\left\{\frac{W_0^{1-\gamma}}{1-\gamma}\left(X^{\frac{1}{\gamma}}+1\right)^{\gamma}\right\} \\ &= \frac{W_0^{1-\gamma}}{1-\gamma}E\left\{\left(X^{\frac{1}{\gamma}}+1\right)^{\gamma}\right\} \end{aligned} \tag{43}$$

The result above is similar to that in Rothschild and Stiglitz (1971). However, here I can also calculate the expected utility under constrained information i.e. when $\kappa \in (0, \infty)$,

$$\begin{aligned} X &= \beta E[A_1^{1-\gamma}|s_0] \\ &= \beta E[\exp((1-\gamma)\epsilon_1)|s_0] \\ &= \beta \exp\left((1-\gamma)E[\epsilon_1|s_0] + \frac{(1-\gamma)^2}{2} var[\epsilon_1|s_0]\right) \\ &= \beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right) \end{aligned} \tag{44}$$

Combine (43) and (44),

$$E[U] = \frac{W_0^{1-\gamma}}{1-\gamma} E\left[\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}} + 1\right]^{\gamma}$$

A.5. Solving the expected indirect utility

Before presenting the proof, I need to know the following characteristics of the indirect expected utility function:

1. Ex ante, the posterior mean $\hat{\epsilon}_1$ is a random variable with $\hat{\epsilon}_1 \sim N(-\frac{1}{2}\sigma^2, \sigma^2 - \hat{\sigma}^2)$.

2. $\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}} + 1$ is then a shifted lognormal distributed variable. Then $E\left[\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}} + 1\right]^{\gamma}$ is equivalent to solving the γth moment of a shifted lognormal variable.

3. I can apply the binomial theorem to expand the expression (25) above, i.e. when the relative risk aversion coefficient γ is integer I have the indirect utility function as

$$\begin{aligned} E[U] &= \frac{W_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \begin{pmatrix} \gamma \\ q \end{pmatrix} E\left[\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}}\right]^{q} \\ &= \frac{W_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \begin{pmatrix} \gamma \\ q \end{pmatrix} E\beta^{\frac{q}{\gamma}} \exp\left(\frac{1-\gamma}{\gamma} q\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\right) \end{aligned} \tag{45}$$

where q are integers from 0 to γ.

Equation (45) can be written as the sum of expectations of lognormal variables. Since the distribution of $\hat{\epsilon}_1$ is known, I can also obtain the distribution of $\frac{1-\gamma}{\gamma}q\hat{\epsilon}_1$,

$$\frac{1-\gamma}{\gamma} q\hat{\epsilon}_1 \sim N\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2, \left(\frac{1-\gamma}{\gamma}q\right)^2(\sigma^2 - \hat{\sigma}^2)\right)$$

Inserting the expected value of lognormal variables delivers

$$
\begin{aligned}
E[U] &= \frac{W_0^{1-\gamma}}{1-\gamma}\sum_{q=0}^{\gamma}\binom{\gamma}{q}\beta^{\frac{q}{\gamma}}\exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2+\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2(\sigma^2-\hat{\sigma}^2)+\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\right)\\
&= \frac{W_0^{1-\gamma}}{1-\gamma}\sum_{q=0}^{\gamma}\binom{\gamma}{q}\beta^{\frac{q}{\gamma}}\exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2+\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2\sigma^2\right)\exp\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\left(1-\frac{q}{\gamma}\right)\right)
\end{aligned} \tag{46}
$$

where the expected indirect utility $E[U]$ is a function of posterior variance $\hat{\sigma}^2$ of productivity shock ϵ_1.

A.6. Proof of Proposition 1

Proof.

$$
\begin{aligned}
\frac{\partial E[U]}{\partial \hat{\sigma}^2} &= \frac{W_0^{1-\gamma}}{1-\gamma}\sum_{q=0}^{\gamma}\left[\binom{\gamma}{q}\beta^{\frac{q}{\gamma}}\exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2+\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2\sigma^2\right)\right]\\
&\quad\times\exp\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\left(1-\frac{q}{\gamma}\right)\right)\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\left(1-\frac{q}{\gamma}\right)\right)\\
&\quad< 0
\end{aligned} \tag{47}
$$

where $\gamma > 1$.

■

A.7. Concave objective function and convex constraint set

Proof. First, maximizing a concave objective function can be replaced by minimizing a convex function. I rewrite the objective function (14) as

$$
\max_{\hat{\sigma}^2} V(\hat{\sigma}^2) = E[U] - \lambda\kappa \tag{48}
$$

s.t.

$$
\kappa = \frac{1}{2}\log\left(\frac{\sigma^2}{\hat{\sigma}^2}\right) \tag{49}
$$

Then the first-order condition is

$$
\begin{aligned}
\frac{\partial V(\hat{\sigma}^2)}{\partial \hat{\sigma}^2} &= \frac{W_0^{1-\gamma}}{1-\gamma}\sum_{q=0}^{\gamma}\left[\binom{\gamma}{q}\beta^{\frac{q}{\gamma}}\exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2+\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2\sigma^2\right)\right]\\
&\quad\times\exp\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\left(1-\frac{q}{\gamma}\right)\right)\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\left(1-\frac{q}{\gamma}\right)\right)+\frac{1}{2}\lambda\frac{1}{\hat{\sigma}^2}
\end{aligned} \tag{50}
$$

And the second-order condition is

$$\frac{\partial^2 V(\hat{\sigma}^2)}{\partial(\hat{\sigma}^2)^2} = \frac{W_0^{1-\gamma}}{1-\gamma}\sum_{q=0}^{\gamma}\left[\begin{pmatrix}\gamma\\ q\end{pmatrix}\beta^{\frac{q}{\gamma}}\exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2+\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2\sigma^2\right)\right]$$
$$\times\exp\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\left(1-\frac{q}{\gamma}\right)\right)\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\left(1-\frac{q}{\gamma}\right)\right)^2-\frac{1}{2}\lambda\frac{1}{(\hat{\sigma}^2)^2}<0 \tag{51}$$

(51) shows that the objective function is concave with respect to $\hat{\sigma}^2$ assuming $\gamma > 1$.

Let us denote

$$g(\hat{\sigma}^2) = \frac{1}{2}\log(\hat{\sigma}^2) - \frac{1}{2}\log(\sigma^2) < 0 \tag{52}$$

I can also show concave function $g(\cdot)$ defines a convex constraint set. Then the optimization is defined by a concave objective function and a convex set, which is a well-posed mathematical problem with local maximum being also global maximum.

■

A.8. Proof of Proposition 2

Proof.

I first rewrite the first order condition (**??**) as

$$F = \frac{W_0^{1-\gamma}}{1-\gamma}\sum_{q=0}^{\gamma}\left[\begin{pmatrix}\gamma\\ q\end{pmatrix}\beta^{\frac{q}{\gamma}}\exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2+\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2\sigma^2\right)\right]$$
$$\times\exp\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\left(1-\frac{q}{\gamma}\right)\right)\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\left(1-\frac{q}{\gamma}\right)\right)+\frac{1}{2}\lambda\frac{1}{\hat{\sigma}^2} \tag{53}$$

From the equation above I can apply implicit function theorem and obtain:

$$\frac{\partial F}{\partial\hat{\sigma}^2} = \frac{W_0^{1-\gamma}}{1-\gamma}\sum_{q=0}^{\gamma}\left[\begin{pmatrix}\gamma\\ q\end{pmatrix}\beta^{\frac{q}{\gamma}}\exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2+\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2\sigma^2\right)\right]$$
$$\times\exp\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\left(1-\frac{q}{\gamma}\right)\right)\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\left(1-\frac{q}{\gamma}\right)\right)^2-\frac{1}{2}\lambda\frac{1}{(\hat{\sigma}^2)^2}<0 \tag{54}$$

$$\frac{\partial F}{\partial W_0} = W_0^{-\gamma}\sum_{q=0}^{\gamma}\left[\begin{pmatrix}\gamma\\ q\end{pmatrix}\beta^{\frac{q}{\gamma}}\exp\left(-\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)\sigma^2+\frac{1}{2}\left(\frac{1-\gamma}{\gamma}q\right)^2\sigma^2\right)\right]$$
$$\times\exp\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\hat{\sigma}^2\left(1-\frac{q}{\gamma}\right)\right)\left(\frac{(1-\gamma)^2}{2}\frac{q}{\gamma}\left(1-\frac{q}{\gamma}\right)\right)>0 \tag{55}$$

$$\frac{\partial\hat{\sigma}^2}{\partial W_0} = -\frac{\frac{\partial F}{\partial W_0}}{\frac{\partial F}{\partial\hat{\sigma}^2}} > 0 \tag{56}$$

■

A.9. Proof of Proposition 3

Proof.

$$
\begin{aligned}
\frac{\partial F}{\partial \beta} &= \frac{W_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[\begin{pmatrix} \gamma \\ q \end{pmatrix} \frac{q}{\gamma} \beta^{\frac{q}{\gamma}-1} \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left(\frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \right] \\
&\quad \times \exp\left(\frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \hat{\sigma}^2 \left(1 - \frac{q}{\gamma} \right) \right) \left(\frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \left(1 - \frac{q}{\gamma} \right) \right) < 0 \qquad (57) \\
\frac{\partial \hat{\sigma}^2}{\partial \beta} &= -\frac{\frac{\partial F}{\partial \beta}}{\frac{\partial F}{\partial \hat{\sigma}^2}} < 0 \qquad (58)
\end{aligned}
$$

■

A.10. Proof of Proposition 4

Proof.

First, taking first-order derivative of equation (53) delivers:

$$
\frac{\partial F}{\partial \lambda} = \frac{1}{\hat{\sigma}^2} > 0 \qquad (59)
$$

Together with the result derived in equation (54), I can easily show

$$
\frac{\partial \hat{\sigma}^2}{\partial \lambda} = -\frac{\frac{\partial F}{\partial \lambda}}{\frac{\partial F}{\partial \hat{\sigma}^2}} > 0 \qquad (60)
$$

■

A.11. Proof of Proposition 5

Proof. Similarly to the proof in A.8, I can also obtain that

$$
\begin{aligned}
\frac{\partial F}{\partial \sigma^2} &= \frac{W_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[\begin{pmatrix} \gamma \\ q \end{pmatrix} \beta^{\frac{q}{\gamma}} \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left(\frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \left(-\frac{q}{2\gamma} (1-\gamma) \left(1 - \frac{q}{\gamma} + q \right) \right) \right] \\
&\quad \times \exp\left(\frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \hat{\sigma}^2 \left(1 - \frac{q}{\gamma} \right) \right) \left(\frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \left(1 - \frac{q}{\gamma} \right) \right) < 0 \qquad (61) \\
\frac{\partial \hat{\sigma}^2}{\partial \sigma^2} &= -\frac{\frac{\partial F}{\partial \sigma^2}}{\frac{\partial F}{\partial \hat{\sigma}^2}} < 0 \qquad (62)
\end{aligned}
$$

■

A.12. Approximated solution to information choice on the first stage

First, equation (**??**) can be rewritten as

$$\frac{1}{2}\frac{1}{W_0}\beta^{\frac{1}{2}}\exp\left(\frac{3}{8}\sigma^2\right)=\lambda\frac{1}{\hat{\sigma}^2\exp\left(\frac{1}{8}\hat{\sigma}^2\right)} \tag{63}$$

Then taking the logarithm on both sides of (63)

$$\log\left(\frac{\beta^{\frac{1}{2}}}{2W_0}\right)+\frac{3}{8}\sigma^2=\log(\lambda)-\log(\hat{\sigma}^2)-\frac{1}{8}\hat{\sigma}^2 \tag{64}$$

From the information constraint in equation (15), I have $2\kappa=\log(\sigma^2)-\log(\hat{\sigma}^2)$ and $\hat{\sigma}^2=\exp(-2\kappa)\sigma^2$. Therefore, (64) becomes

$$\begin{aligned}\log\left(\frac{\beta^{\frac{1}{2}}}{2W_0}\right)+\frac{3}{8}\sigma^2&=\log(\lambda)-\log(\sigma^2)+2\kappa-\frac{\sigma^2}{8}\exp(-2\kappa)\\&\approx\log(\lambda)-\log(\sigma^2)+2\kappa-\frac{\sigma^2}{8}(-2\kappa+1)\\&\approx\log(\lambda)+\left(2+\frac{\sigma^2}{4}\right)\kappa-\log(\sigma^2)-\frac{\sigma^2}{8}\\\implies\kappa&\approx\frac{1}{2+\frac{\sigma^2}{4}}\log\left(\frac{\beta^{\frac{1}{2}}\sigma^2\exp\left(\frac{1}{2}\sigma^2\right)}{2\lambda W_0}\right)\end{aligned} \tag{65}$$

When the information cost is large enough such that $\log\left(\frac{\beta^{\frac{1}{2}}\sigma^2\exp\left(\frac{1}{2}\sigma^2\right)}{2\lambda}\right)\leq 0$, by the definition of attention capacity, I set $\kappa=0$,

$$\kappa^*=\begin{cases}\frac{1}{2+\frac{\sigma^2}{4}}\log\left(\frac{\beta^{\frac{1}{2}}\sigma^2\exp\left(\frac{1}{2}\sigma^2\right)}{2\lambda W_0}\right), & \text{if } \lambda<\frac{\beta^{\frac{1}{2}}\sigma^2\exp\left(\frac{1}{2}\sigma^2\right)}{2}\\0, & \text{otherwise}\end{cases}$$

A.13. Joint distribution of productivity shock and savings behavior

Similar to the price setting problem of a rationally inattentive firm in Matějka (2016), the rationally inattentive consumers in my model also choose: (1). how to process information through a channel of limited information capacity, and (2). how to respond to the realized posterior knowledge (or signal).

More specifically, there exists a sequence of elementary pieces of signals for any $f(s_0|\epsilon_1)$ that consumers choose to achieve, as long as it satisfies the information constraint. Once a particular signal on the productivity shock is realized, consumers then choose an optimal action,

$$K_1^*(s_0)=\arg\max_{K_1}\int U(K_1,\epsilon_1)f(\epsilon_1|s_0)ds_0$$

where the posterior belief $f(\epsilon_1|s_0)$ is given by Bayes law as shown in Section 2.

With a noisy Gaussian channel, I can separate these two decisions. However, in general without the noisy Gaussian channel, these two decisions can be merged into one choice. When choosing the optimal way to extract signal, i.e. $f(s_0|\epsilon_1)$, agents are aware of their policy function, $K_1^*(s_0)$.

When deciding how to extract signal, agents actually solve:

$$f(s_0|\epsilon_1) = \arg\max_{\tilde{f}(s_0|\epsilon_1)} \int U(K_1^*(s_0), \epsilon_1)\tilde{f}(s_0|\epsilon_1)g(\epsilon_1)d\epsilon_1 ds_0$$

$$\text{s.t. information constraint and} \quad K_1^*(s_0)$$

Again, very similar to the story in Matějka (2016), in my model the utility function is a single-peaked function of K_1, meaning that there is only one utility-maximizing saving decision. In addition, since the information constraint is a concave function of the joint distribution, there is never optimal to choose a collection with two posteriors of different forms resulting in the same optimal saving decision. An averaged posterior distribution of these two posteriors would necessarily lead to the same expected utility but with a lower utilization of the information capacity.

Therefore, I can equivalently describe the signal extraction problem in terms of the resulting posterior distribution of productivity shock that the observed signal lead to, instead of the distribution of signals. I can substitute $f(s_0|\epsilon_1)$ by $f(K_1^*|\epsilon_1)$, where $K_1^* = K_1^*(s_0)$. Due to the productivity A_1 is only determined by ϵ_1

Appendix B: solving the joint density with gradient-based search method

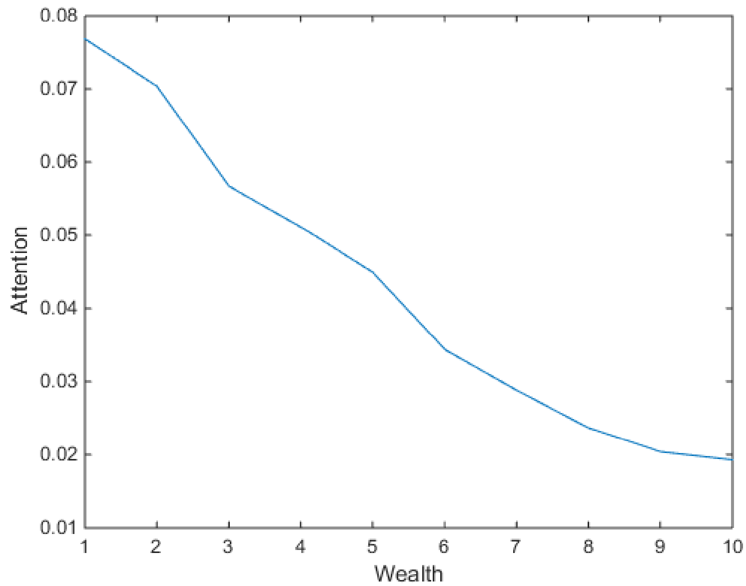
To solve the model above I apply the gradient-based search method, which is developed by Chris Sims. First, in order to apply this method, I discretize the model with $n_{A_1} = 10$ equi-spaced grid points ranging from 1.02 to 1.25, and $n_{K_1} = 10$ equi-spaced grid points ranging from 15% to 75% of the initial endowment W_0. For a given W_0, I build a prior belief simplex for the state variable A_1. Let $f(A_1)$ be the probability of state $A_1 \in \Omega_{A_1}$. The simplex is constructed using uniform random samples from the unit simplex, whose sum per row is one and each column contains also n_{A_1} random values in the $[0, 1]$ interval.

Then, for this prior belief simplex I initialize the corresponding joint distribution of savings K_1 and capital return A_1. As $n_{A_1} = n_{K_1}$, the joint distribution of savings and capital return for a given multidimensional grid point is a square matrix with rows corresponding to levels of savings and columns corresponding to capital return. Summing the matrix per row returns the marginal distribution of savings $p(K_1)$ and summing the matrix per column returns the marginal distribution of capital return $f(A_1)$. I restrict the joint distribution $f(K_1, \tilde{A}_1) = 0$ for some values $\tilde{A}_1$ with $g(\tilde{A}_1) = 0$.

Finally, I solve for the optimal joint density by using gradient-based search method by Chris Sims (CSMINWEL), and iterating on the value function. The numerical solution is easy to compute as I have a two-period model and agents choose the optimal joint distribution only once for a given prior belief. Convergence is achieved in about 100 iterations, with each iteration taking around 5.5 minutes[31]. Once I have the optimal $f^*(K_1, A_1)$, I calculate the marginal distributions of K_1 and A_1 respectively. Then I can compute the optimal attention choice according to equation (32) for each simplex point. I use the same prior belief simplex and simulate the model for 10 heterogeneous agents with different initial endowments. The parameter values for conducting this numerical analysis can be found in the appendix **C.4**.

[31] I also use these optimal joint distributions as the starting belief simplex and see if the value function converges to a different optimal value of the joint distribution of productivity shock and investment. In these exercises, it does not converge to different optimal joint distributions.

Figure 11: Attention choice and initial endowment in a non-Gaussian RI model



Appendix C: parameter values

C.1. Parameter values for Figure 2

Parameter	Value
Initial endowment W_0	[1 2 3 4 5]
Prior Variance σ^2	0.08
Discounting factor β	0.98
Information cost λ	0.015
Range for posterior variance $\hat{\sigma}^2$	$(0.001, \sigma^2 - 0.0001)$
Risk aversion degree γ	4

C.2. Parameter values for Figure 3

Parameter	Value
Initial endowment W_0	2
Prior Variance σ^2	0.08
Discounting factor β	0.98
Information cost λ	0.015
Range for posterior variance $\hat{\sigma}^2$	$(0.001, \sigma^2 - 0.0001)$
Risk aversion degree γ	2, 3, 4, 5, 6, 7, 8, 9, 10

C.3. Parameter values for Figure 5, ??, and 6

Parameter	Value
Initial endowment W_0	[1,...,2]
Capital return space A_1	[0.8,...,1.2]
Savings space K_1	$[0.1 * W_0$,...,$0.75 * W_0]$
Discounting factor β	0.98
Joint distribution per simplex point, $f(K_1, A_1)$	50×50
Marginal distribution of K_1,	50×1
Marginal distribution of A_1,	50×1
Information cost λ	0.001
Risk aversion degree γ	4

C.4. Parameter values for Figure 11

Parameter	Value
Initial endowment W_0	[1,...,10]
Capital return space A_1	[1.02,...,1.25]
Savings space K_1	$[0.05 * W_0$,...,$0.7 * W_0]$
Discounting factor β	0.98
Joint distribution per simplex point, $f(K_1, A_1)$	8×8
Marginal distribution of K_1,	8×1
Marginal distribution of A_1,	8×1
Information cost λ	0.015
Risk aversion degree γ	4

Appendix D: empirical results

D.1. Summary statistics

Table 4: Summary table

Data are from 2016 SCF dataset. Attention is calculated by the the information sources that are used to make investment decisions. Permanent income is measured by the reported income.

	N	Mean	StdDev	Max	Min
Attention	5148	0.72	0.45	1	0
High uncertainty	5137	0.28	0.45	1	0
log of net worth	5148	12.70	2.96	21.19	0
log of wealth	5148	13.29	2.33	21.21	7.16
log of permanent income	5148	11.65	1.60	19.48	6.92
log of desired amount of precautionary savings	5148	9.63	1.98	18.71	3.91
Information delegation	5132	0.67	0.47	1	0
Financial literacy	5148	2.35	0.82	3	0
Low risk aversion	5148	0.05	0.21	1	0
Education	5148	10.11	2.87	14	-1
Male	5148	0.79	0.41	1	0
Age	5148	56.30	14.23	95	30
Number of kids	5148	0.78	1.12	7	0
Married	5148	0.66	0.47	1	0

D.2. Robustness check: who pays more attention

Table 5: Relationship between wealth and attention devoted to capital return (robustness check 1)

The table presents coefficients from OLS regressions with the dummy variable of self information-procession as the dependent variable. Controls include number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (wealth)	-0.0216***		-0.0239***	
	(-3.83)		(-3.58)	
log (net worth)		-0.0205***		-0.0218***
		(-3.68)		(-3.29)
Information delegation	-0.244***	-0.253***	-0.257***	-0.263***
	(-15.44)	(-16.01)	(-13.67)	(-13.85)
Financial literacy	0.0455***	0.0472***	0.0638***	0.0695***
	(2.76)	(2.90)	(2.74)	(3.08)
Constant	1.133***	1.118***	1.026***	0.997***
	(6.33)	(6.19)	(4.35)	(4.25)
N	2349	2334	1500	1496

Table 6: Relationship between wealth and attention devoted to capital return (robustness check 2)

The table presents coefficients from OLS regressions with the dummy variable of self information-procession as the dependent variable. Controls include number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (wealth)	-0.0520*		-0.0753**	
	(-1.94)		(-2.10)	
log (net worth)		-0.0506**		-0.0634*
		(-1.97)		(-1.92)
Information delegation	0.256**	0.185	0.285*	0.232
	(2.10)	(1.57)	(1.86)	(1.54)
log (wealth) * delegation	-0.0340***		-0.0356***	
	(-4.29)		(-3.65)	
log (net worth) * delegation		-0.0298***		-0.0326***
		(-3.87)		(-3.39)
Financial literacy	-0.265***	-0.242*	-0.389***	-0.311*
	(-2.00)	(-1.92)	(-2.15)	(-1.87)
log (wealth) * financial literacy	0.0211**		0.0296**	
	(2.33)		(2.52)	
log (net worth) * financial literacy		0.0199**		0.0252**
		(2.30)		(2.30)
Constant	1.600***	1.563***	1.827***	1.611***
	(3.76)	(3.88)	(3.10)	(2.97)
N	2334	2318	1495	1489

Table 7: Relationship between wealth and attention devoted to capital return (robustness check 3)

The table presents coefficients from OLS regressions with the dummy variable of self information-procession as the dependent variable. Controls include number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Financial wealth (cash-on-hand) is the sum of financial net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (financial wealth)	-0.0227***		-0.0219***	
	(-3.68)		(-3.00)	
log (financial net worth)		-0.0176***		-0.0203***
		(-2.94)		(-2.87)
Information delegation	-0.254***	-0.264***	-0.263***	-0.267***
	(-16.43)	(-17.09)	(-14.17)	(-14.33)
Financial literacy	0.0535**	0.0565***	0.0770***	0.0779***
	(3.13)	(3.35)	(3.37)	(3.41)
Constant	1.081***	1.078***	0.939***	0.877***
	(6.11)	(5.97)	(4.15)	(3.80)
N	2242	2245	1474	1472

Table 8: Relationship between wealth and attention devoted to capital return (robustness check 4)

The table presents coefficients from OLS regressions with the dummy variable of self information-procession as the dependent variable. Controls include number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Financial wealth (cash-on-hand) is the sum of financial net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (financial wealth)	-0.0460		-0.0604 *	
	(-1.56)		(-1.65)	
log (financial net worth)		-0.0482*		-0.0698**
		(-1.65)		(-1.98)
Information delegation	0.184	0.0460	0.182	0.0715
	(1.50)	(0.40)	(1.14)	(0.48)
log (financial wealth) * delegation	-0.0310***		-0.0306***	
	(-3.67)		(-2.88)	
log (financial net worth) * delegation		-0.0224***		-0.0237**
		(-2.79)		(-2.38)
Financial literacy	-0.193	-0.183	-0.263	-0.279
	(-1.36)	(-1.33)	(-1.48)	(-1.64)
log (financial wealth) * financial literacy	0.0176*		0.0234 *	
	(1.75)		(1.93)	
log (financial net worth) * financial literacy		0.0175*		0.0251**
		(1.75)		(2.11)
Constant	1.431***	1.512***	1.521***	1.602***
	(3.20)	(3.51)	(2.65)	(2.93)
N	2242	2245	1474	1472

D.3. Robustness check: who is more uncertain about future income

Table 9: Relationship between wealth and high uncertainty regarding capital income (robustness check 1)

The table presents coefficients from OLS regressions with the dummy variable of high perceived uncertainty as the dependent variable. Controls include number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (wealth)	0.0176***		0.0164**	
	(3.28)		(2.53)	
log (net worth)		0.0179***		0.0169***
		(3.41)		(2.67)
Information delegation	0.0275	0.0232	0.0297	0.0279
	(1.07)	(0.89)	(0.95)	(0.88)
Financial literacy	-0.0280*	-0.0308*	-0.0394*	-0.0400*
	(-1.76)	(-1.95)	(-1.74)	(-1.80)
Constant	0.107	0.134	0.299	0.288
	(0.58)	(0.71)	(1.20)	(1.16)
N	2346	2331	1498	1494

Table 10: Relationship between wealth and high uncertainty regarding capital income (robustness check 2)

The table presents coefficients from OLS regressions with the dummy variable of high perceived uncertainty as the dependent variable. Controls include number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Financial wealth (cash-on-hand) is the sum of financial net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (financial wealth)	0.0183***		0.0180**	
	(3.03)		(2.49)	
log (financial net worth)		0.0186***		0.0190***
		(3.21)		(2.71)
Information delegation	0.0238	0.0240	0.0197	0.0218
	(1.13)	(1.14)	(0.69)	(0.77)
Financial literacy	-0.0202	-0.0244	-0.0394*	-0.0332
	(-1.24)	(-1.54)	(-1.74)	(-1.49)
Constant	0.107	0.134	0.299	0.288
	(0.58)	(0.71)	(1.20)	(1.16)
N	2346	2331	1498	1494

D.4. Robustness check: who saves more due to precautionary motive

Table 11: Relationship between wealth and precautionary saving rate (robustness check 1)

The table presents coefficients from OLS regressions with the ratio of desired amount of precautionary savings over permanent income as the dependent variable. Controls include number of kids, age, male, marriage, occupation, year of education, posterior uncertainty, as well as a dummy variable of self-employment. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (wealth)	0.0287***		0.0269**	
	(2.73)		(2.16)	
log (net worth)		0.0290***		0.0289**
		(2.80)		(2.39)
Information delegation	-0.0394	-0.0372	-0.0727	-0.0746
	(-0.82)	(-0.76)	(-1.04)	(-1.05)
Financial literacy	-0.00833	-0.00928	0.0139	0.0138
	(-0.30)	(-0.34)	(0.31)	(0.32)
Constant	0.187	0.141	0.254	0.190
	(0.41)	(0.32)	(0.37)	(0.29)
N	2340	2325	1492	1488

Table 12: Relationship between wealth and precautionary saving rate (robustness check 2)

The table presents coefficients from OLS regressions with the log of desired amount of as the dependent variable. Controls include number of kids, age, male, marriage, occupation, year of education, posterior uncertainty, as well as a dummy variable of self-employment. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (wealth)	0.699***		0.714***	
	(35.92)		(31.20)	
log (net worth)		0.680***		0.697***
		(35.46)		(30.86)
Information delegation	-0.0189	-0.00869	-0.134	-0.126
	(-0.30)	(-0.13)	(-1.57)	(-1.46)
Financial literacy	-0.00608	-0.0210	0.0199	0.00818
	(-0.13)	(-0.47)	(0.31)	(0.13)
Constant	0.266	0.664	0.0536	0.592
	(0.49)	(1.23)	(0.08)	(0.85)
N	2340	2325	1492	1488

Table 13: Relationship between wealth and precautionary saving rate (robustness check 3)

The table presents coefficients from OLS regressions with the ratio of desired amount of precautionary savings over permanent income as the dependent variable. Controls include number of kids, age, male, marriage, occupation, year of education, posterior uncertainty, as well as a dummy variable of self-employment. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Financial wealth (cash-on-hand) is the sum of financial net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (financial wealth)	0.0246**		0.0190	
	(1.98)		(1.28)	
log (financial net worth)		0.0300**		0.0261*
		(2.48)		(1.82)
Information delegation	-0.0255	-0.0227	-0.0520	-0.0513
	(-0.52)	(-0.45)	(-0.72)	(-0.71)
Financial literacy	-0.00419	-0.00464	0.0187	0.0148
	(-0.14)	(-0.16)	(0.42)	(0.33)
Constant	0.305	0.253	0.313	0.311
	(0.68)	(0.56)	(0.49)	(0.49)
N	2234	2237	1467	1465

Table 14: Relationship between wealth and precautionary saving rate (robustness check 4)

The table presents coefficients from OLS regressions with the log of desired amount of as the dependent variable. Controls include number of kids, age, male, marriage, occupation, year of education, posterior uncertainty, as well as a dummy variable of self-employment. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Financial wealth (cash-on-hand) is the sum of financial net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (financial wealth)	0.745***		0.749***	
	(34.34)		(29.59)	
log (financial net worth)		0.693***		0.694***
		(32.73)		(27.86)
Information delegation	0.0116	0.0283	-0.124	-0.113
	(0.18)	(0.41)	(-1.45)	(-1.26)
Financial literacy	-0.0153	-0.0194	0.0183	0.000248
	(-0.31)	(-0.36)	(0.28)	(0.00)
Constant	0.113	0.945*	-0.159	1.041
	(0.21)	(1.68)	(-0.23)	(1.45)
N	2234	2237	1467	1465

D.5. Pooling data from 2010, 2013, and 2016

Table 15: Relationship between wealth and attention devoted to capital return (robustness check 5): 2010, 2013 and 2016 data

The table presents coefficients from OLS regressions with the dummy variable of self information-procession as the dependent variable. Controls include number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. EQUITY is the total value of equity in stocks, stock mutual funds, and combination of mutual funds held either directly or indirectly by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
log (wealth)	-0.0144***		-0.0105**	
	(-2.62)		(-2.50)	
log (net worth)			0.00726**	0.00552*
			(2.36)	(1.80)
Information delegation	-0.245***	-0.245***	-0.211***	-0.212***
	(-25.51)	(-25.53)	(-26.57)	(-26.65)
Education level	0.0219***	0.0219***	0.0257***	0.0256***
	(6.53)	(6.53)	(11.25)	(11.21)
log (DEQ)	0.00126	0.000568		
	(0.35)	(0.16)		
log (EQUITY)			0.00726**	0.00552*
			(2.36)	(1.80)
Constant	0.505	0.510	-5.941*	-5.964*
	(0.11)	(0.11)	(-1.72)	(-1.73)
N	4973	4973	9871	9871

Table 16: Relationship between wealth and high uncertainty regarding capital income (robustness check 5): 2010, 2013 and 2016 data

The table presents coefficients from OLS regressions with the dummy variable of high perceived uncertainty as the dependent variable. Controls include number of kids, age, male, marriage, race, occupation, risk aversion, and year of education. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Wealth (cash-on-hand) is the sum of net worth and current income. EQUITY is the total value of equity in stocks, stock mutual funds, and combination of mutual funds held either directly or indirectly by households. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
log (wealth) log_wealth	0.0172***		0.0295***	
	(3.03)		(6.92)	
log (net worth)		0.0146***		0.0238***
		(2.75)		(6.25)
Information delegation	-0.0112	-0.0113	-0.00506	-0.00514
	(-0.75)	(-0.76)	(-0.50)	(-0.51)
Education level	-0.00241	-0.00241	-0.00857***	-0.00845***
	(-0.72)	(-0.72)	(-3.79)	(-3.73)
log (DEQ)	0.00141	0.00235		
	(0.37)	(0.63)		
log (EQUITY)			-0.0154***	-0.0136***
			(-4.78)	(-4.27)
Constant	12.47**	12.47**	13.61***	13.63***
	(2.51)	(2.51)	(3.88)	(3.89)
N	4957	4957	9846	9846

Table 17: Relationship between wealth and precautionary saving rate (robustness check 5): 2010, 2013 and 2016 data

The table presents coefficients from OLS regressions with the ratio of desired amount of precautionary savings over permanent income as the dependent variable. Controls include number of kids, age, male, marriage, occupation, year of education, posterior uncertainty, as well as a dummy variable of self-employment. Equity is the total value of financial assets held by households that are invested in stock. DEQ is the total value of equity in directly held stocks, stock mutual funds, and combination of mutual funds held by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (wealth)	0.00867		0.620***	
	(0.87)		(33.98)	
log (net worth)		0.0154*		0.554***
		(1.70)		(31.67)
Information delegation	0.0168	0.0155	-0.0314	-0.0399
	(0.65)	(0.60)	(-0.75)	(-0.94)
Education level	0.0151**	0.0149**	0.0555***	0.0547***
	(2.35)	(2.32)	(5.65)	(5.48)
log (DEQ)	0.0164**	0.0126*	0.0599***	0.0785***
	(2.49)	(1.93)	(5.52)	(7.13)
Constant	-3.356	-3.441	-0.527	-0.971
	(-0.35)	(-0.35)	(-0.04)	(-0.07)
N	4957	4957	9846	9846

Table 18: Relationship between wealth and precautionary saving rate (robustness check 5): 2010, 2013 and 2016 data

The table presents coefficients from OLS regressions with the ratio of desired amount of precautionary savings over permanent income as the dependent variable. Controls include number of kids, age, male, marriage, occupation, year of education, posterior uncertainty, as well as a dummy variable of self-employment. Equity is the total value of financial assets held by households that are invested in stock. EQUITY is the total value of equity in stocks, stock mutual funds, and combination of mutual funds held either directly or indirectly by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (wealth)	0.00867		0.620***	
	(0.87)		(33.98)	
log (net worth)		0.0154*		0.554***
		(1.70)		(31.67)
Information delegation	0.0168	0.0155	-0.0314	-0.0399
	(0.65)	(0.60)	(-0.75)	(-0.94)
Education level	0.0151**	0.0149**	0.0555***	0.0547***
	(2.35)	(2.32)	(5.65)	(5.48)
log (DEQ)	0.0164**	0.0126*	0.0599***	0.0785***
	(2.49)	(1.93)	(5.52)	(7.13)
Constant	-3.356	-3.441	-0.527	-0.971
	(-0.35)	(-0.35)	(-0.04)	(-0.07)
N	4946	4946	4946	4946

Table 19: Relationship between wealth and precautionary saving rate (robustness check 5): 2010, 2013 and 2016 data

The table presents coefficients from OLS regressions with the ratio of desired amount of precautionary savings over permanent income as the dependent variable. Controls include number of kids, age, male, marriage, occupation, year of education, posterior uncertainty, as well as a dummy variable of self-employment. Equity is the total value of financial assets held by households that are invested in stock. EQUITY is the total value of equity in stocks, stock mutual funds, and combination of mutual funds held either directly or indirectly by households. Wealth (cash-on-hand) is the sum of net worth and current income. I only consider households with wealth/net worth above median level. t statistics are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	(1)	(2)	(3)	(4)
	$Equity > 0$	$Equity > 0$	$DEQ > 0$	$DEQ > 0$
log (wealth)	0.0344***		0.649***	
	(5.21)		(48.57)	
log (net worth)		0.0370***		0.523***
		(6.71)		(42.48)
Information delegation	0.00302	0.00135	-0.00415	-0.00591
	(0.20)	(0.09)	(-0.15)	(-0.21)
Education level	0.00895**	0.00878**	0.0501***	0.0526***
	(2.47)	(2.43)	(7.87)	(8.06)
log (EQUITY)	0.000470	-0.00330	0.0227**	0.0598***
	(0.10)	(-0.67)	(2.50)	(6.35)
Constant	-6.060	-6.107	-25.55***	-25.03**
	(-1.04)	(-1.05)	(-2.65)	(-2.51)
N	9834	9834	9834	9834