

Yin, Penghui

**Working Paper**

## The Optimal Amount of Attention to Capital Income Risk

CESifo Working Paper, No. 7413

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Yin, Penghui (2018) : The Optimal Amount of Attention to Capital Income Risk, CESifo Working Paper, No. 7413, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/198773>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# The Optimal Amount of Attention to Capital Income Risk

*Penghui Yin*

## **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](mailto:office@cesifo.de)

Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl

[www.cesifo-group.org/wp](http://www.cesifo-group.org/wp)

An electronic version of the paper may be downloaded

- from the SSRN website: [www.SSRN.com](http://www.SSRN.com)
- from the RePEc website: [www.RePEc.org](http://www.RePEc.org)
- from the CESifo website: [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# The Optimal Amount of Attention to Capital Income Risk

## Abstract

This paper studies attention allocation behavior of rationally inattentive consumers who have CRRA preferences, face uninsured capital income risk, and suffer from an information-processing capacity constraint. For given attention devoted to capital income risk, we solve for the optimal consumption-saving choices and show that the expected welfare is increasing with capacity, assuming a relative risk aversion degree larger than unity. Furthermore, we solve for attention choice and find that households would pay more attention to capital income risk if they have (i) lower initial wealth endowment, (ii) lower marginal cost of information, (iii) higher prior volatility of capital return, and (iv) higher degree of patience.

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*

October, 2018

I am very grateful to Michael Evers and Mirko Wiederholt for their advice and support. I also thank Stefania Albanesi, Chris Carroll, Elisabeth Falck, Michael Haliassos, Michael Hasler, Marten Hillebrand, Joe Kaboski, Markus Kontny, Hayk Kamalyan, Kristoffer Nimark, Petr Sedláček, Zexi Sun, Hélény Rey, Johannes Wohlfart, Jiankang Zhang as well as seminar and conference participants at Goethe University Frankfurt, York University, Renmin University of China, Vanderbilt University, University of Hohenheim, and ifo Institute Munich for helpful comments and suggestions. All remaining errors are my own.

# 1 Introduction

Understanding effects of capital income risk on saving behavior is an important task in macroeconomic studies<sup>1</sup>. One reason is that according to empirical studies with surveys data, such as [Bertaut and Starr-McCluer \(2002\)](#), a large share of households invest directly or indirectly in stocks, pension funds, and corporate equities, which have highly nondiversified risk. Furthermore, private businesses account for almost half of aggregate capital and employment in the U.S., and they are subject to dramatic idiosyncratic capital income risk. 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. However, still little is known about what determines investors' information choice. Since understanding information acquisition is crucial to understanding households' consumption-saving behavior, this paper applies rational inattention theory of [Sims \(2003\)](#) to jointly investigate information acquisition behavior and consumption-saving choices under capital income risk.

In our model, agents allocate their limited attention to future capital return before making consumption-saving decision subject to a cost of attention, as in [Van Nieuwerburgh and Veldkamp \(2010\)](#) who jointly model information acquisition and static portfolio choice problem. In this situation, agents face a trade-off between paying attention to capital income risk. On the one hand, paying more attention increases expected utility because a more precise signal helps agents to make more efficient consumption allocation plan. On the other hand, paying attention causes an information cost in utility due to the attention capacity constraint. Since agents have scarce consumption resource and limited attention capacity, they need to solve two optimization problems: consumption choice and attention allocation. A two-period framework with stochastic capital return allows us to solve these two problems analytically and make the following predictions.

First, in contrast to previous literature where the capital return cannot be learned <sup>2</sup>, agents in our model have smaller demand for savings assuming relative risk aversion larger than one. This is because paying attention to future capital return reduces the perceived capital income uncertainty, and this reduction leads to less precautionary savings when the substitution effect dominates. Second, expected utility from consumption is increasing attention devoted to capital return. The intuition is that when agents pay more attention to capital return, they can observe

---

<sup>1</sup>See, inter alia, [Angeletos \(2007\)](#) and [Benhabib et al. \(2015\)](#).

<sup>2</sup>See, inter alia, [Levhari and Srinivasan \(1969\)](#), [Sandmo \(1970\)](#), [Rothschild and Stiglitz \(1971\)](#).

a more precise signal, which helps them to make more efficient consumption decisions.

Finally, optimal attention amount is increasing with risk aversion, degree of patience and prior volatility of capital return, but decreasing with the marginal cost of information. The key finding is that poorer agents pay more attention to capital income risk relative to their wealthier counterparts, because more limited budgets lead to paying attention more profitable under CRRA preferences. The intuition is that the poor is more susceptible to capital income risk than the rich, and paying more attention helps them to allocate scarce resource more efficiently. This finding is in line with that of [Shah et al. \(2018\)](#), who study money in the mental lives of the poor with experiments and conclude that, facing more limited budget, the poor are often think more carefully than the rich when it comes to monetary opportunity cost of economic decisions.

Our paper builds on [Sims \(2003, 2006\)](#). Sims argues that consumers cannot attend perfectly to all freely available information. He proposes modeling consumers' limited attention as a constraint on information flow, which is measured by entropy as in [Shannon \(1948\)](#)'s information theory. Following works of Sims, rational inattention is popularly applied to solve consumption-saving problems. For example, [Luo \(2008\)](#) solves excess smoothness puzzle and excess sensitivity puzzle, and [Tutino \(2013\)](#) studies asymmetric consumption responses to wealth shocks. However, these papers all study consumption-saving choices of rationally inattentive agents who face stochastic labor income but constant capital income<sup>3</sup>. Our paper mainly investigates attention allocation and consumption choice problems jointly under capital income risk. Therefore we assume that at the beginning of period 1, lifetime labor income is known with certainty, whereas capital return is stochastic but can be learned.

Another difference between our model and the others in the rational inattention consumption literature is that we model agents with CRRA preferences. For the reason of tractability, linear quadratic utility or log utility are popularly used in these studies<sup>4</sup>. However, when investigating saving behavior with capital income risk neither linear quadratic nor log utility can be employed, because the former utility leads to certainty equivalence and the latter one has a property that income and substitution effects cancel out. [Sims \(2006\)](#) first solves rationally inattentive agent's saving problem with CRRA utility in a two-period framework, and especially the author solves

---

<sup>3</sup>[Maćkowiak and Wiederholt \(2015\)](#) also study a stochastic capital income in their model, but different from our paper, the authors solve the RI model numerically under linear quadratic framework

<sup>4</sup>For example, [Sims \(2003\)](#), [Luo \(2008\)](#) and [Luo and Young \(2016\)](#) directly work with linear quadratic framework. [Luo \(2010\)](#) and [Maćkowiak and Wiederholt \(2015\)](#) work with log-quadratic approximation of CRRA utility.

the model analytically for a log form. [Tutino \(2013\)](#) extends the two-period model of [Sims \(2006\)](#) into a fully dynamic rational inattention consumption model and solves it numerically. On the contrary, in this paper we can obtain analytical solutions to the consumption-saving choice and attention allocation problems with a stochastic interest rate and CRRA preferences in a two-period RI framework<sup>5</sup>.

Our work also contributes to the literature studying consumption choices with non-learnable 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<sup>6</sup>. The discussion about 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 learn about future capital return and solves optimal consumption-saving choice subject to improving information about future capital return, but the amount of information is exogenous to agents. In our paper, we go one step further and allow agents to choose not only how much to consume but also how much information to acquire regarding future capital income. This endogenous information choice is important to understand heterogeneous consumption-saving behavior. For example, our finding on the relationship between wealth inequality and attention choice can explain why the rich save more<sup>7</sup>.

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.

---

<sup>5</sup>[Van Nieuwerburgh and Veldkamp \(2010\)](#) also solves a static portfolio choice and information choice problems analytically with CARA and CRRA preferences respectively. However, different from their study that utility comes from terminal wealth, in our paper agents make optimal intertemporal consumption allocation under limited attention. Furthermore, we not only solve the model analytically under assumption on Gaussian signal as in this paper, but also check the robustness of our results numerically without Gaussian assumption.

<sup>6</sup>The prior belief is usually the true distribution of capital return. This assumption is also made in the literature of precautionary savings such as in [Skinner \(1988\)](#) and [Kimball \(1990\)](#) who discuss both labor income risk and capital income risk.

<sup>7</sup>See [Carroll \(2000\)](#) and [Quadrini \(2000\)](#).

## 2 Two-period consumption-saving model

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

**Agents.** In our 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,

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

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  $Y_0$  be the initial wealth endowment<sup>8</sup> 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,

$$C_0 + K_1 = Y_0 \quad (2)$$

$$C_1 = Y_1 \quad (3)$$

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:

$$Y_t = A_t K_t \quad (4)$$

where  $A_t$  is productivity or return to capital, and  $K_t$  is the capital used for producing consumption goods  $Y_t$ . Productivity follows the process:

$$A_t = \exp(\epsilon_t) \quad (5)$$

where  $\epsilon_t$  is the exogenous productivity shock.

Agents are endowed with a prior belief about the distribution from which the productivity disturbance is drawn:  $\epsilon_t \sim N(-\frac{1}{2}\sigma^2, \sigma^2)$ . However, the realization of  $\epsilon_t$  cannot be observed at

---

<sup>8</sup>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.



period  $t - 1$ . In our two-period model the capital return in period 1 is

$$A_1 = \exp(\epsilon_1) \quad (6)$$

**Signal structure.** We assume that agents learn the exogenous productivity shock by observing a signal<sup>9</sup>:

$$s_0 = \epsilon_1 + \psi_0 \quad (7)$$

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  $\epsilon_1$ , the signal precision is only determined by the noise variance  $\sigma_\psi^2$ . According to Section 2, this signal structure is equivalent to say that our agent learns the productivity shock through a noisy Gaussian channel.

**Bayes' Law.** Agents use Bayes' Law to combine their prior belief and the observed signal in (7) such that  $\epsilon_1 | s_0 \sim N(\hat{\epsilon}_1, \hat{\sigma}^2)$ , where

$$\hat{\epsilon}_1 \equiv E[\epsilon_1 | s_0] = \frac{(-\frac{1}{2}\sigma^2)\sigma_\psi^2 + \sigma^2 s_0}{\sigma^2 + \sigma_\psi^2} \quad (8)$$

$$\hat{\sigma}^2 \equiv Var[\epsilon_1 | s_0] = \frac{\sigma^2 \sigma_\psi^2}{\sigma^2 + \sigma_\psi^2} \quad (9)$$

where  $\hat{\epsilon}_1$  is the posterior mean. Notice that the posterior variance  $\hat{\sigma}^2$  is determined by the prior variance  $\sigma^2$  and noise variance  $\sigma_\psi^2$ . Given prior belief, (9) 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  $\sigma_\psi^2$ .

**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.

$$\mathfrak{I}^1 = \left\{ Y_0, \epsilon_1 \sim N\left(-\frac{\sigma^2}{2}, \sigma^2\right), \hat{\epsilon}_1 \sim N\left(-\frac{\sigma^2}{2}, \sigma^2 - \hat{\sigma}^2\right) \right\}$$

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

where the ex-ante distribution of the posterior mean  $\hat{\epsilon}_1$  is derived in the appendix **A.1**.

---

<sup>9</sup>As the productivity in the AK model is also the marginal return to capital, which is determined by the exogenous shock  $\epsilon_1$  only, in the rest of this paper we interchangeably refer to paying attention to productivity shock as paying attention to capital return or return on savings.

**Information constraint.** According to the RI literature, information constraint is defined as the maximum reduction in uncertainty, which is measured by entropy<sup>10</sup>. 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  $\epsilon_1$  is contained in  $s_0$ ; the equality implies that choosing the optimal signal precision is equivalent to allocating the optimal amount of attention  $\kappa$  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:

$$V = \max_{\{\hat{\sigma}^2\}} E_{\mathcal{J}^1} [u(C_0^*) + \beta u(C_1^*)] - \lambda \kappa \quad (10)$$

Subject to

$$C_0^* = \arg \max_{C_0} E_{\mathcal{J}^2} [u(C_0) + \beta u(C_1)] \quad (11)$$

$$C_1^* = A_1(Y_0 - C_0^*) \quad (12)$$

$$\frac{1}{2} \log \left( \frac{\sigma^2}{\hat{\sigma}^2} \right) = \kappa \quad (13)$$

where equation (10) is the objective function for the consumer;  $E_{\mathcal{J}^2}[\cdot]$  is the expectation conditional on the information set  $\mathcal{J}^2$  and  $E_{\mathcal{J}^1}[\cdot]$  is the expectation over all possible signals<sup>11</sup>;  $\lambda$  is the marginal information cost, which can be interpreted as the opportunity cost of devoting some

---

<sup>10</sup>In information theory, entropy is defined as the negative expected log of the density function of the random variable [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”.

<sup>11</sup>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 our model due to limited information processing capacity stochastic capital return is only partially revealed at the initial period.

of the scarce resource attention to the stochastic capital return or equivalently to  $\epsilon_1$ ; budget constraints are incorporated into equations (11) and (12); equation (13) is the information constraint. Later we 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.

### 3 Analytical solution with Gaussian signal

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

#### 3.1 Solution approach and equilibrium

Given a noisy Gaussian signal, we 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 3), 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.

The two-period model with capital income risk allows us to solve the model backward. For a given signal realization and signal precision, we can solve for the optimal consumption-saving choice, which depends on the posterior mean and variance of the productivity shock  $\epsilon_1$ . 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.

$$\begin{aligned} \max_{C_0, K_1, C_1} U(C_0) + \beta E_{\mathcal{I}^2} U(C_1) & \quad (14) \\ \text{s.t. budget constraints} & \end{aligned}$$

Solving the maximization problem (14) we will obtain the optimal consumption in each period denoted as  $C_0^*$  and  $C_1^*$ . In Section 4.2 below, we will show that the optimal consumption

is a function of the realized signal and posterior variance of the capital return, which are contained in the second stage information set  $\mathcal{J}^2$ . Therefore, when plugging  $C_0^*$  and  $C_1^*$  back into the utility function, we 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  $\kappa$ . Formally agents solve:

$$\max_{\hat{\sigma}^2} E_{\mathcal{J}^1} [U(C_0^*) + \beta E_{\mathcal{J}^2} U(C_1^*)] - \lambda \kappa \quad (15)$$

subject to  $C_0^*(\mathcal{J}^2)$ ,  $C_1^*(\mathcal{J}^2)$  and

$$\frac{1}{2} \log \left( \frac{\sigma^2}{\hat{\sigma}^2} \right) = \kappa \quad (16)$$

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  $C_0^*(\mathcal{J}^2)$ ,  $C_1^*(\mathcal{J}^2)$ , and demand for savings  $K_1^*(\mathcal{J}^2)$  such that:*

1. *given  $Y_0$ ,  $f(\epsilon_1)$  and signal observation  $s_0$ , agents choose optimal consumption and savings according to the maximization problem defined in (14)*
2. *and the optimal posterior variance according to the maximization problem defined in (15).*

### 3.2 Stage 2 solution: optimal consumption-saving choices

According to the two-step approach, our first step is to solve for the second optimization problem where the signal is already realized, and thus we can rewrite the optimization problem in (14) as:

$$\begin{aligned} U &= u(C_0) + \beta E[u(C_1)|s_0] \\ &= \frac{(Y_0 - K_1)^{1-\gamma}}{1-\gamma} + \beta E \left[ \frac{(A_1 K_1)^{1-\gamma}}{1-\gamma} \middle| s_0 \right] \end{aligned} \quad (17)$$

F.O.C:

$$\frac{\partial U}{\partial K_1} = -(Y_0 - K_1)^{-\gamma} + \beta E[(A_1 K_1)^{-\gamma} A_1 | s_0] = 0 \quad (18)$$

Solving equation (18), we get optimal consumption and savings conditional on the observed signal<sup>12</sup>:

$$K_1^* = \frac{1}{[\beta E[A_1^{1-\gamma}|s_0]]^{-\frac{1}{\gamma}} + 1} Y_0 \quad (19)$$

$$C_0^* = \frac{1}{[\beta E[A_1^{1-\gamma}|s_0]]^{\frac{1}{\gamma}} + 1} Y_0 \quad (20)$$

$$C_1^* = A_1 K_1^* \quad (21)$$

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 we can easily obtain that

$$E[A_1^{1-\gamma}|s_0] = E[\exp((1-\gamma)\epsilon_1)|s_0] \quad (22)$$

follows a log-normal distribution.

Given the results in (8), (9), (19), (20) and (22) we can rewrite the consumption rule as :

$$C_0^{\kappa < \infty} = \frac{1}{\left[ \beta \exp \left( (1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2 \right) \right]^{\frac{1}{\gamma}} + 1} Y_0 \quad (23)$$

and the savings decision is given by

$$K_1^{\kappa < \infty} = \frac{1}{\left[ \beta \exp \left( (1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2 \right) \right]^{-\frac{1}{\gamma}} + 1} Y_0 \quad (24)$$

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:

$$C_0^{\kappa=0} = \frac{1}{\left[ \beta \exp \left( -\frac{\gamma(1-\gamma)}{2}\sigma^2 \right) \right]^{\frac{1}{\gamma}} + 1} Y_0 \quad (25)$$

And accordingly, the savings is

$$K_1^{\kappa=0} = \frac{1}{\left[ \beta \exp \left( -\frac{\gamma(1-\gamma)}{2}\sigma^2 \right) \right]^{-\frac{1}{\gamma}} + 1} Y_0 \quad (26)$$

---

<sup>12</sup>These results are also obtained by Epstein (1980), Gollier (2001) and Miao (2004).

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 are not influenced by the information acquisition:  $C_0 = Y_0/(1+\beta)$ . This is because the income effect and substitution effect of changing the perceived riskiness will off set with  $\gamma = 1$ . Furthermore, we 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<sup>13</sup>. This result directly shows the role of information constraint in consumer's precautionary saving behavior. Equations (25) and (26) are results that can also be found in the previous literature without learning capital return, such as Sandmo (1970). Comparing with this literature, our results shown equations (23) and (24) imply that paying attention to learning capital return reduces perceived future income uncertainty and demand for savings. This is because an increase in capital income risk leads to two opposite effects on saving demand, which are positive income risk and negative substitution effect, and CRRA preferences with a degree of relative risk aversion larger than unity leads to income effect dominance<sup>14</sup>.

### 3.3 Welfare implications of information processing constraints

In order to examine the welfare effects of capital income fluctuations under information processing constraint, we present how an increase in perceived uncertainty leads to changes in the ex-ante expected utility. Given equations (19)-(21) we can reorganize the expected utility function as:

$$E[U] = \frac{Y_0^{1-\gamma}}{1-\gamma} E \left[ \left( \beta \exp \left( (1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2 \right) \right)^{\frac{1}{\gamma}} + 1 \right]^\gamma \quad (27)$$

A complete derivation of indirect utility function (27) 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 we go one step further and study the effect of changing capital income risk on expected utility. However, calculating the expected utility in equation (27) analytically is not easy, and in the current paper we provide an alternative approach to study the welfare effect of information constraint by applying the binomial theorem<sup>15</sup>.

<sup>13</sup>We show the proof for this result in appendix A.3.

<sup>14</sup>Many other papers also discuss this point. See Phelps (1962), Levhari and Srinivasan (1969), Rothschild and Stiglitz (1971), Miao (2004), Eeckhoudt and Schlesinger (2008).

<sup>15</sup> Appendix A.5 shows the details of this approach.

**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 we show an analytical 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 our 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 we will show how to solve the information choice problem and explain the determinants of attention allocation.

### 3.4 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<sup>16</sup>.

$$V = \max_{\hat{\sigma}^2} \frac{Y_0^{1-\gamma}}{1-\gamma} E \left[ \left( \beta \exp \left( (1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2}\hat{\sigma}^2 \right) \right)^{\frac{1}{\gamma}} + 1 \right]^\gamma - \lambda\kappa \quad (28)$$

$$\frac{1}{2} \log \left( \frac{\sigma^2}{\hat{\sigma}^2} \right) = \kappa \quad (29)$$

This optimization problem is a well-posed mathematical problem with a concave objective function (28) and a convex constraint set (29). 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) we apply binomial theorem on indirect utility with integer  $\gamma > 1$ , and (b) we conduct comparative statics by applying implicit function theorem on the first order condition.

First, for integer relative risk aversion degree we can obtain the first order condition for the

---

<sup>16</sup>An endogenous information choice is sometimes missing in the RI literature. For example, [Luo \(2008, 2010\)](#) uses a different approach solves the dynamics of consumption with a similar result as equation (23), however, the author finds interesting results of consumption behavior based on some fixed exogenous attention capacity.

above maximization problem<sup>17</sup>:

$$\begin{aligned} & \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp \left( -\frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \right] \\ & \times \exp \left( \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \hat{\sigma}^2 \left( 1 - \frac{q}{\gamma} \right) \right) \left( \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \left( 1 - \frac{q}{\gamma} \right) \right) + \frac{1}{2} \lambda \frac{1}{\hat{\sigma}^2} = 0 \end{aligned} \quad (30)$$

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 we show how optimal signal precision characterized by the first order condition above is affected by factors such as prior volatility, information cost, initial wealth endowment, and time discounting factor<sup>18</sup>. Let us start with two common themes in the RI literature regarding the effects of information processing cost and prior volatility on information choice<sup>19</sup>.

**Proposition 2** *Assuming integer  $\gamma > 1$ , the optimal amount of attention allocated to capital income is decreasing with the information processing cost  $\lambda$ .*

Proof can be found in appendix **A.8**. 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 3** *Assuming integer  $\gamma > 1$ , the optimal amount of attention allocated to capital income is increasing with the volatility of the exogenous disturbance  $\sigma^2$ .*

Proof can be found in **A.9**. 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

---

<sup>17</sup>Details can be found in appendix A.5 and A.7.

<sup>18</sup>In the literature such as [Sims \(2006\)](#), [Maćkowiak and Wiederholt \(2009\)](#) and [Tutino \(2013\)](#), authors tend to discuss the optimal capacity allocation instead of the signal precision, and therefore, in the following text we use attention allocated to capital income risk to substitute the choice of signal precision.

<sup>19</sup>[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.



rationally inattentive agents would pay more attention to the variable with higher volatility. The intuition in our model is that for given marginal cost of attention, the higher the prior volatility, the higher the marginal benefit of paying attention.

The next two propositions on the effects of patience and initial wealth on the optimal amount attention are main findings in our paper.

**Proposition 4** *Assuming integer  $\gamma > 1$ , the optimal amount of attention allocated to capital income is decreasing with the initial endowment  $Y_0$ .*

Proof for this proposition can be found in **A.10**. The key mechanism through which initial wealth operates on attention allocation comes from our assumptions on the utility function and marginal cost of paying attention. Agents in our model have CRRA-type preferences, which implies that absolute risk aversion is smaller when initial wealth is higher. A decreasing absolute risk aversion is equivalent to increasing absolute risk tolerance, meaning that agents with very large initial wealth 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  $\lambda$  is a positive constant number and identical to all agents. However, from the first order condition in equation (30) we can observe that the marginal benefit of paying attention in terms of consumption utility is lower for wealthier agents, and this decreasing benefit of paying attention over initial wealth can also explain the negative relationship between initial wealth and the optimal amount of attention devoted to capital income risk<sup>20</sup>.

**Proposition 5** *Assuming integer  $\gamma > 1$ , the optimal amount of attention allocated to capital income is increasing with the patience level denoted by  $\beta$ .*

Proof can be found in **A.11**. In the literature it is rare to see the effect of patience or the discount factor on the attention allocation. However in our two-period setup we think it is worth discussing the role of patience in the information choice, because attention is paid to future capital income, which contributes to all second period consumption and thus the second period utility. Agents with larger  $\beta$  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.

---

<sup>20</sup>Maćkowiak and Wiederholt (2015) also study the attention allocation of heterogeneous households. However different from ours, 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.

## 3.5 Additional results

### 3.5.1 Attention choice with $\gamma = 2$

Now, let us show how to solve the optimal attention allocation problem analytically for  $\gamma = 2$ . Here we call this simplified optimization problem for attention choice the baseline model.

$$V = \max_{\hat{\sigma}^2} -\frac{1}{Y_0} \left[ \beta \exp(\sigma^2) + 2\beta^{\frac{1}{2}} \exp\left(\frac{3}{8}\sigma^2 + \frac{1}{8}\hat{\sigma}^2\right) + 1 \right] - \lambda\kappa \quad (31)$$

subject to

$$\frac{1}{2} \log\left(\frac{\sigma^2}{\hat{\sigma}^2}\right) = \kappa \quad (32)$$

where Equation (32) implies that choosing the optimal posterior variance is equivalent to choosing the optimal attention capacity  $\kappa$ .

And the optimal attention choice is then given by the first order condition:

$$-2\frac{1}{Y_0}\beta^{\frac{1}{2}} \exp\left(\frac{3}{8}\sigma^2 + \frac{1}{8}\hat{\sigma}^2\right) \frac{1}{8} + \frac{1}{2}\lambda \frac{1}{\hat{\sigma}^2} = 0 \quad (33)$$

where the posterior variance or equivalently the attention allocated to learning the capital return is determined by this first-order condition.

Replacing the posterior variance  $\hat{\sigma}^2$  by the attention amount  $\kappa$  according to the information constraint (32), the first-order condition is re-organized as:

$$\log\left(\frac{\beta^{\frac{1}{2}}}{2Y_0}\right) + \frac{3}{8}\sigma^2 = \log(\lambda) - \log(\sigma^2) + 2\kappa - \frac{\sigma^2}{8} \exp(-2\kappa) \quad (34)$$

First notice that it is not possible to solve for  $\kappa^*$  with closed form directly from the first order condition (34). However, it is also not difficult to approximate the last term on the right hand side of the first order condition up to order 1 around  $\kappa = 0$ . An approximately linearized first order condition delivers the optimal attention<sup>21</sup> choice as:

$$\kappa^* = \begin{cases} \frac{1}{2 + \frac{\sigma^2}{4}} \log\left(\frac{\beta^{\frac{1}{2}} \sigma^2 \exp(\frac{1}{2}\sigma^2)}{2\lambda Y_0}\right), & \text{if } \lambda < \frac{\beta^{\frac{1}{2}} \sigma^2 \exp(\frac{1}{2}\sigma^2)}{2Y_0} \\ 0, & \text{otherwise} \end{cases} \quad (35)$$

where the ratio  $\frac{\beta^{\frac{1}{2}} \sigma^2 \exp(\frac{1}{2}\sigma^2)}{2Y_0}$  is the marginal benefit of paying attention to capital return at  $\kappa = 0$ . A full derivation for  $\kappa^*$  can be found in appendix **A.12**.

This equation shows that agents pay some amount of attention if the marginal cost is smaller than this marginal benefit, otherwise they would pay no attention to capital return and make

<sup>21</sup>A similar result can be found in [Maćkowiak and Wiederholt \(2009\)](#) and [Wiederholt \(2010\)](#).

consumption-saving decisions only based on their prior beliefs. More importantly, this results suggests that paying attention to capital income risk is not always optimal. At  $\kappa = 0$ , agent's marginal benefit maybe smaller than the marginal cost due to large wealth holding, or small prior volatility. Under this situation, this consumer prefers to pay no attention to capital income risk, and it is equivalent to that this agent prefers late resolution of uncertainty.

### 3.5.2 The effect of risk aversion on attention allocation

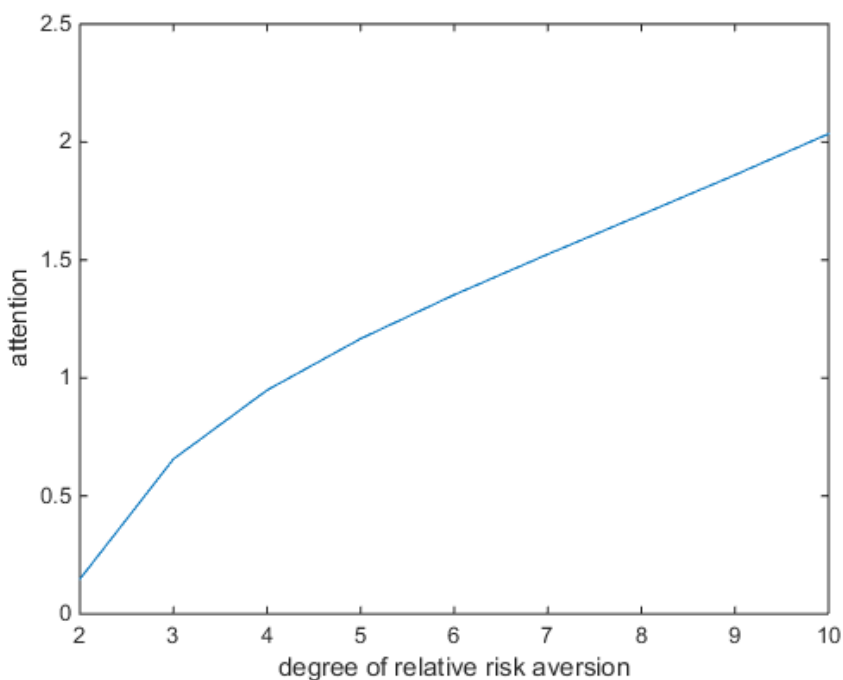


Figure 1: Relative risk aversion and attention

In the comparative statics studies, we cannot analytically show how relative risk aversion affects attention choice. Here we present a simulation result on the relation between attention allocated to capital income risk and relative risk aversion in Figure 1 (parameter values can be found in appendix B.1.). The intuition is that for higher relative risk aversion, agents dislike the uncertainty in future capital income more, and therefore, they would like to pay more attention in order to reduce the uncertainty, or equivalently agents would like to pay more attention in order to obtain a more precise signal on future capital return<sup>22</sup>.

<sup>22</sup>This finding can also be found in Tutino (2013). For three values of relative risk aversion coefficient (0.5, 1 and 5), the author also numerically shows that the attention allocation is increasing over relative risk aversion degree.

## 4 Moving beyond the Gaussian framework

To solve the model analytically, we assume that the signal on productivity disturbance follow a Gaussian distribution in the previous sections. Here we check whether our main finding on the correlation between wealth and attention choice under Gaussian assumption also holds true in a non-Gaussian setup. However, solving the RI model analytically without a specific distribution form for the unobservable state variable is very difficult. Therefore, in this situation we focus on numerical solutions by following [Sims \(2006\)](#), [Maćkowiak and Wiederholt \(2009\)](#) and [Tutino \(2013\)](#) to solve for the joint distribution of control variable and the unobserved state variable, which are  $K_1$  and  $A_1$  in the model.

### 4.1 Model

Following the model specification in [Sims \(2006\)](#), we form the optimization problem as the following.

$$\max_f \int_{0 < K_1 < Y_0} \left( \frac{C_0^{1-\gamma}}{1-\gamma} + \beta \frac{C_1^{1-\gamma}}{1-\gamma} \right) f(K_1, A_1) dK_1 dA_1 \quad (36)$$

subject to

$$C_0 + K_1 = Y_0 \quad (37)$$

$$C_1 = A_1 K_1 \quad (38)$$

$$1 \geq f(K_1, A_1) \geq 0 \quad (39)$$

$$\int_{0 < K_1 < Y_0} f(K_1, A_1) dK_1 = g(A_1) \quad (40)$$

$$\begin{aligned} & (\lambda : ) \\ & \int_{0 < K_1 < Y_0} \log(f(K_1, A_1)) f(K_1, A_1) dK_1 dA_1 \\ & - \int_0^\infty \left( \log \left( \int_{K_2}^\infty f(K_1, A_1) dA_1 \right) \int_{K_1}^\infty f(K_1, A_1) dA_1 \right) dK_1 \\ & - \int_0^\infty \log(g(A_1)) g(A_1) dA_1 = \kappa \end{aligned} \quad (41)$$

where equation (36) implies that we are still maximizing expected utility, which is the sum of the expected utility in the initial period's consumption,  $\frac{Y_0 - K_1}{1-\gamma}$  and the discounted expected utility of next period's consumption,  $\beta \frac{(A_1 K_1)^{1-\gamma}}{1-\gamma}$ . However, in this case the control variable is not consumption or attention as in Section 4, but the joint distribution  $f(K_1, A_1)$ . Equation (39) shows the boundaries of feasible values for probability densities, which are  $f = 0$  and  $f = 1$ .

Equation (40) 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 (41), and  $\lambda > 0$  is the marginal cost of using the capacity to process information.

We follow the solution methodology by Tutino (2013) to discretize the space of saving demand  $K_1$  and capital return  $A_1$  and their marginal and joint distributions. Then by inserting the budget constraints (37) and (38) into the objective function (36), the optimization above becomes:

$$V = \max_{f(K_1, A_1)} \left[ \sum_{A_1 \in \Omega_{A_1}} \left( \sum_{K_1 \in \Omega_{K_1}} \frac{(Y_0 - K_1)^{1-\gamma}}{1-\gamma} f(K_1, A_1) \right) + \beta \sum_{A_1 \in \Omega_{A_1}} \left( \sum_{K_1 \in \Omega_{K_1}} \frac{(A_1 K_1)^{1-\gamma}}{1-\gamma} f(K_1, A_1) \right) \right] \quad (42)$$

subject to:

$$1 \geq f(K_1, A_1) \geq 0 \quad \forall (K_1, A_1) \in B \quad (43)$$

$$\sum_{K_1 \in \Omega_{K_1}} f(K_1, A_1) = g(A_1) \quad (44)$$

( $\lambda$  :)

$$\kappa = \sum_{A_1 \in \Omega_{A_1}} \sum_{K_1 \in \Omega_{K_1}} f(K_1, A_1) \left( \log \frac{f(K_1, A_1)}{\sum_{K_1 \in \Omega_{K_1}} f(K_1, A_1) * g(A_1)} \right) \quad (45)$$

where  $\Omega_{A_1}$  is the space for productivity, and  $\Omega_{K_1}$  is the space for savings choice, and  $B \equiv \{(K_1, A_1) : K_1 < Y_0\}$ .

## 4.2 Numerical technique

To solve the model above we apply the gradient-based search method, which is also used in Sims (2006) and Tutino (2013). First, we choose  $n_{A_1} = 25$  equi-spaced grid points ranging from 1.02 to 1.8. Let  $g(A_1)$  be the probability of state  $A_1 \in \Omega_{A_1}$ , then points in the simplex  $\Delta(g(A_1))$  are  $n_{A_1}$  distinct values for the marginal probability mass function  $g(A_1)$  in the interval  $[0, 1]$ . 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. We choose  $n_{K_1} = 25$  equi-spaced grid points ranging from 10% to 90% of initial wealth  $Y_0$ , and those values constitute the behavior space  $\Omega_{K_1}$ .

Second, for each simplex point in  $\Delta(g(A_1))$  we 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  $g(A_1)$ . 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, we apply value function iteration to find the optimal joint distribution<sup>23</sup>. This optimal  $f(K_1, A_1)$  allows us to calculate the optimal attention choice according to equation (45). The parameter values for conducting this numerical analysis can be found in the appendix B.2.

### 4.3 Main results

As shown in figure 2, without making any assumption on the signal distribution, we find that the optimal attention choice is decreasing with wealth, which is in line with our findings in the Gaussian case. We 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 wealthier households would like to bear more risk in their future capital income.

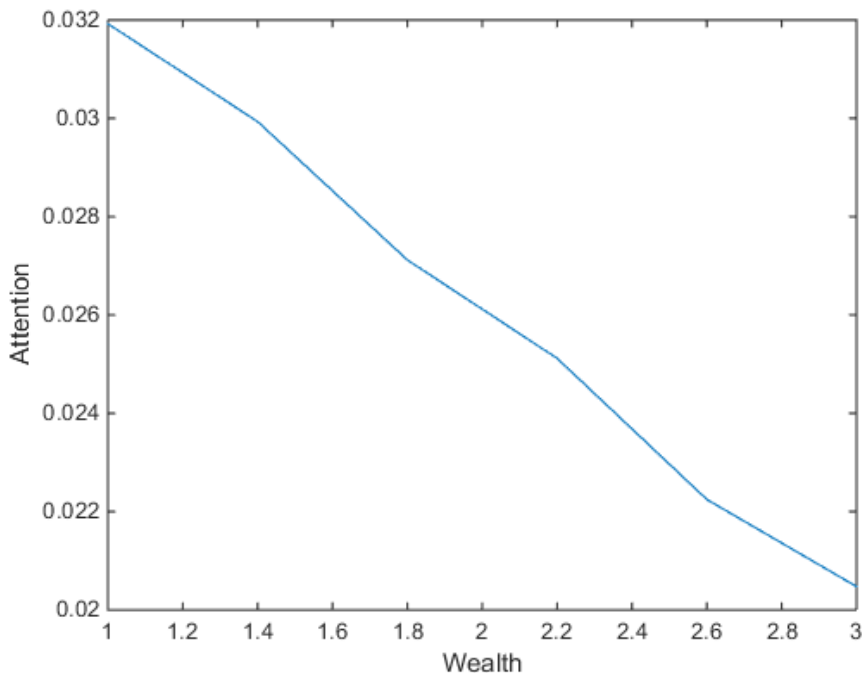


Figure 2: Attention choice and wealth in a non-Gaussian RI model

This result is in line with findings in surveys and experiments. For example, [Shah et al. \(2018\)](#) study the mental lives of the poor with with very limited budgets, and argue that the

<sup>23</sup>The pseudo codes can be found in the appendix **A.13**.

poor think more carefully about opportunity cost of economic decisions. This suggests that the poor must pay more attention to everyday economic dimensions (financial, temporal, and etc.) in order to be less susceptible to potential loss in future income. [Sicherman et al. \(2016\)](#) study financial attention using daily investor online logins as a proxy. The authors find that during market declines, wealthy investors were more likely to engage in financial “ostrich behavior”, allowing them to pay less attention to their financial holdings relative to poorer investors who found it harder to suppress monetary concerns.

## 5 Conclusion

This paper incorporates information processing constraint into a consumption-saving model of capital income risk. We find that when agents face capital income risk that affects their future income, their consumption-saving choices depend largely on the degree of risk aversion and how much attention they will pay to learning about the return on savings. These results challenges the empirical literature on precautionary saving behavior, such as [Hurst et al. \(2010\)](#) who directly used the exogenous variance but not taking the reality that households learn future capital return into account. According to [Epstein \(1980\)](#), the different informativeness in signal about future capital return affects current consumption-saving choices.

Then we study the effect of changing perceived riskiness of capital return on expected utility and we find that the welfare loss is decreasing with attention capacity. However, agents cannot choose infinite capacity level due to the positive marginal cost of attention, and therefore, agents need also solve for optimal information choice problem. In order to study determinants of attention allocation, we 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 wealth and information processing cost. Relations between attention allocation and information cost and prior volatility are the common themes in the RI literature, and we 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. Our main contributions are showing relations between attention allocation and wealth inequality and degree of patience. Due to the constant marginal cost and decreasing marginal utility of paying attention, wealthier households benefit less from devoting attention to capital return. Moreover, CRRA utility implies increasing absolute risk tolerance, meaning that wealthier households would like bear more risk

in their future capital income. In addition, we solve for a closed-form (with approximation) attention choice for agents with a degree of relative risk aversion 2. This solution intuitively show under which condition will agents pay attention to capital return, and under which condition agents do not pay any attention at all. To show the role of relative risk aversion playing in attention allocation, we simulate our model and show a positive effect of higher degree of relative risk aversion on the optimal attention allocation. This is because more risk-averse agents dislike future uncertainty more, and this motivates them to allocate more attention to future return on savings.

Our analytical solutions to consumption-saving choice and information choice are based on the assumption of Gaussian signal. In order to show our main findings are valid in a non-Gaussian world, we extend our baseline model to a non-Gaussian framework. Provided reasonable parameter values, our numerical results also show a negative effect of initial wealth on the optimal attention choice.

Finally, we want to point that there are still problems unsolved 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 our model all decisions are ex post decisions, i.e. after signal observation. Therefore, in future research we will extend this model into three or more periods in order study such effect. This is crucial to understand the consumption-investment behavior in a incomplete information model, such as the puzzling empirical fact that households from different wealth group having different savings rates. Second, in our 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, we also plan 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 re-balance affects consumption-saving decisions.



# Appendix

## Appendix A

### A.1. Derivation of the distribution of the pre-posterior mean $\hat{\epsilon}_1$

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

First, according to the Bayesian updating rule we have

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

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  $\psi_0 \sim N(0, \sigma_\psi^2)$ , we can obtain  $s_0 \sim N(-\frac{1}{2}\sigma^2, \sigma^2 + \sigma_\psi^2)$ . Therefore,

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

and,

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

### A.2. Timeline of agents' optimization problem

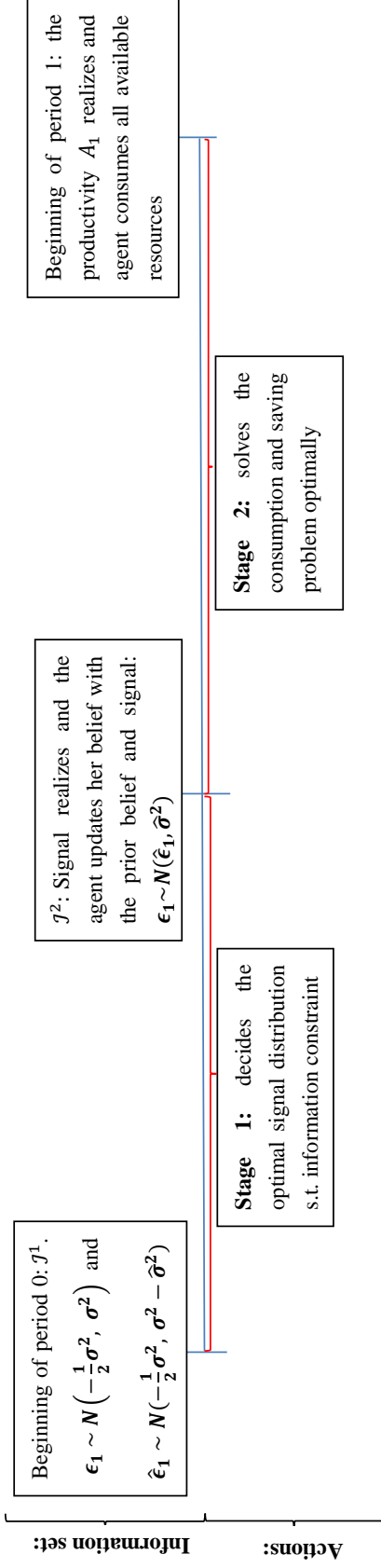


Figure 3: timeline for a two-period consumption-saving problem

### 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

$$\begin{aligned}
 E[C_0^{\kappa < \infty}] &= E \left[ \frac{1}{\left[ \beta \exp \left( (1 - \gamma) \hat{\epsilon}_1 + \frac{(1 - \gamma)^2}{2} \hat{\sigma}^2 \right) \right]^{\frac{1}{\gamma}} + 1} Y_0 \right] \\
 &\geq \frac{1}{E \left[ \left[ \beta \exp \left( (1 - \gamma) \hat{\epsilon}_1 + \frac{(1 - \gamma)^2}{2} \hat{\sigma}^2 \right) \right]^{\frac{1}{\gamma}} + 1 \right]} Y_0 \\
 &= \frac{1}{\beta^{\frac{1}{\gamma}} \exp \left( -\frac{\sigma^2}{2} \frac{1 - \gamma}{\gamma} + \frac{1}{2} \frac{(1 - \gamma)^2}{\gamma} \hat{\sigma}^2 + \frac{1}{2} \left( \frac{1 - \gamma}{\gamma} \right)^2 (\sigma^2 - \hat{\sigma}^2) \right) + 1} Y_0 \quad (46) \\
 &> \frac{1}{\beta^{\frac{1}{\gamma}} \exp \left( \frac{\gamma - 1}{2} \sigma^2 \right) + 1} Y_0 = C_0^{\kappa = 0}
 \end{aligned}$$

$$\implies E[C_0^{\kappa < \infty}] > C_0^{\kappa = 0} \quad (47)$$

where equation (46) is derived from equations (8) and (9) and  $\hat{\epsilon} \sim N \left( -\frac{1}{2} \sigma^2, \sigma^2 - \hat{\sigma}^2 \right)$ .

To prove the inequality in (47) holds, we 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  $\kappa \in (0, \infty)$  we 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  $\phi = \frac{1}{[\beta E[A_1^{1-\gamma}|s_0]]^{\frac{1}{\gamma}} + 1}$ .

Then, given the equations (19)-(21) we can rewrite the expected utility function as:

$$\begin{aligned}
E[U] &= E \{u(C_0^*) + \beta E[u(C_1^*)|s_0]\} \\
&= E \left\{ \frac{(\phi Y_0)^{1-\gamma}}{1-\gamma} + \beta E \left[ \frac{(A_1(1-\phi)Y_0)^{1-\gamma}}{1-\gamma} |s_0 \right] \right\} \\
&= E \left\{ \frac{(\phi Y_0)^{1-\gamma}}{1-\gamma} + \beta \frac{((1-\phi)Y_0)^{1-\gamma}}{1-\gamma} E[A_1^{1-\gamma}|s_0] \right\}
\end{aligned} \tag{48}$$

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

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

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, we can now write the indirect utility function

$$\begin{aligned}
E[U] &= E \left\{ \frac{(Y_0)^{1-\gamma}}{1-\gamma} [\phi^{1-\gamma} + \beta(1-\phi)^{1-\gamma} E[A_1^{1-\gamma}|s_0]] \right\} \\
&= E \left\{ \frac{Y_0^{1-\gamma}}{1-\gamma} \left[ \left( X^{\frac{1}{\gamma}} + 1 \right)^{\gamma-1} + \left( X^{\frac{1}{\gamma}} + 1 \right)^{\gamma-1} X^{\frac{1-\gamma}{\gamma}} X \right] \right\} \\
&= E \left\{ \frac{Y_0^{1-\gamma}}{1-\gamma} \left[ \left( X^{\frac{1}{\gamma}} + 1 \right)^{\gamma-1} + \left( X^{\frac{1}{\gamma}} + 1 \right)^{\gamma-1} X^{\frac{1}{\gamma}} \right] \right\} \\
&= E \left\{ \frac{Y_0^{1-\gamma}}{1-\gamma} \left( X^{\frac{1}{\gamma}} + 1 \right)^{\gamma} \right\} \\
&= \frac{Y_0^{1-\gamma}}{1-\gamma} E \left\{ \left( X^{\frac{1}{\gamma}} + 1 \right)^{\gamma} \right\}
\end{aligned} \tag{49}$$

Then we can calculate the expected utility under constrained information i.e. when  $\kappa \in$

$(0, \infty)$ ,

$$\begin{aligned}
X &= \beta E[A_1^{1-\gamma} | s_0] \\
&= \beta E[\exp((1-\gamma)\epsilon_1) | s_0] \\
&= \beta \exp\left((1-\gamma)E[\epsilon_1 | s_0] + \frac{(1-\gamma)^2}{2} \text{var}[\epsilon_1 | s_0]\right) \\
&= \beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2} \hat{\sigma}^2\right)
\end{aligned} \tag{50}$$

Combine (49) and (50),

$$E[U] = \frac{Y_0^{1-\gamma}}{1-\gamma} E\left[\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2} \hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}} + 1\right]^\gamma$$

### A.5. Solving the expected indirect utility

Before presenting the proof, we need to know the following regarding this indirect utility function:

1. Ex ante, the posterior mean  $\hat{\epsilon}_1$  is a random variable with  $\hat{\epsilon}_1 \sim N(-\frac{1}{2}\sigma^2, \sigma^2 - \hat{\sigma}^2)$ <sup>24</sup>.
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  $\gamma$ th moment of a shifted lognormal variable.
3. We can apply the binomial theorem to expand the expression (27) above, i.e. when the relative risk aversion coefficient  $\gamma$  is integer we have the indirect utility function as

$$\begin{aligned}
E[U] &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \binom{\gamma}{q} E\left[\left(\beta \exp\left((1-\gamma)\hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2} \hat{\sigma}^2\right)\right)^{\frac{1}{\gamma}}\right]^q \\
&= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \binom{\gamma}{q} E\beta^{\frac{q}{\gamma}} \exp\left(\frac{1-\gamma}{\gamma} q \hat{\epsilon}_1 + \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \hat{\sigma}^2\right)
\end{aligned} \tag{51}$$

where  $q$  are integers from 0 to  $\gamma$ .

Equation (51) can be written as the sum of expectations of lognormal variables. Since the distribution of  $\hat{\epsilon}_1$  is known, we can also obtain the distribution of  $\frac{1-\gamma}{\gamma} q \hat{\epsilon}_1$ ,

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

---

<sup>24</sup>Details can be found in appendix A.3.2

Inserting the expected value of lognormal variables delivers

$$\begin{aligned}
E[U] &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma} q\right) \sigma^2 + \frac{1}{2} \left(\frac{1-\gamma}{\gamma} q\right)^2 (\sigma^2 - \hat{\sigma}^2) + \frac{(1-\gamma)^2 q}{2} \frac{\hat{\sigma}^2}{\gamma}\right) \\
&= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma} q\right) \sigma^2 + \frac{1}{2} \left(\frac{1-\gamma}{\gamma} q\right)^2 \sigma^2\right) \exp\left(\frac{(1-\gamma)^2 q}{2} \frac{\hat{\sigma}^2}{\gamma} \left(1 - \frac{q}{\gamma}\right)\right)
\end{aligned} \tag{52}$$

where the expected indirect utility  $E[U]$  is a function of posterior variance  $\hat{\sigma}^2$  of productivity shock  $\epsilon_1$ .

### A.6. Proof of Proposition 1

**Proof.**

$$\begin{aligned}
\frac{\partial E[U]}{\partial \hat{\sigma}^2} &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma} q\right) \sigma^2 + \frac{1}{2} \left(\frac{1-\gamma}{\gamma} q\right)^2 \sigma^2\right) \right] \\
&\quad \times \exp\left(\frac{(1-\gamma)^2 q}{2} \frac{\hat{\sigma}^2}{\gamma} \left(1 - \frac{q}{\gamma}\right)\right) \left(\frac{(1-\gamma)^2 q}{2} \frac{1}{\gamma} \left(1 - \frac{q}{\gamma}\right)\right) \\
&< 0
\end{aligned} \tag{53}$$

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. We rewrite the objective function (15) as

$$\max_{\hat{\sigma}^2} V(\hat{\sigma}^2) = E[U] - \lambda \kappa \tag{54}$$

s.t.

$$\kappa = \frac{1}{2} \log\left(\frac{\sigma^2}{\hat{\sigma}^2}\right) \tag{55}$$

Then the first-order condition is

$$\begin{aligned}
\frac{\partial V(\hat{\sigma}^2)}{\partial \hat{\sigma}^2} &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma} q\right) \sigma^2 + \frac{1}{2} \left(\frac{1-\gamma}{\gamma} q\right)^2 \sigma^2\right) \right] \\
&\quad \times \exp\left(\frac{(1-\gamma)^2 q}{2} \frac{\hat{\sigma}^2}{\gamma} \left(1 - \frac{q}{\gamma}\right)\right) \left(\frac{(1-\gamma)^2 q}{2} \frac{1}{\gamma} \left(1 - \frac{q}{\gamma}\right)\right) + \frac{1}{2} \lambda \frac{1}{\hat{\sigma}^2}
\end{aligned} \tag{56}$$

And the second-order condition is

$$\begin{aligned} \frac{\partial^2 V(\hat{\sigma}^2)}{\partial(\hat{\sigma}^2)^2} &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp \left( -\frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \right] \\ &\times \exp \left( \frac{(1-\gamma)^2 q}{2} \hat{\sigma}^2 \left( 1 - \frac{q}{\gamma} \right) \right) \left( \frac{(1-\gamma)^2 q}{2} \left( 1 - \frac{q}{\gamma} \right) \right)^2 - \frac{1}{2} \lambda \frac{1}{(\hat{\sigma}^2)^2} < 0 \end{aligned} \quad (57)$$

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

Let us denote

$$g(\hat{\sigma}^2) = \frac{1}{2} \log(\hat{\sigma}^2) - \frac{1}{2} \log(\sigma^2) < 0 \quad (58)$$

we 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.**

We first rewrite the first order condition (34) as

$$\begin{aligned} F &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp \left( -\frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \right] \\ &\times \exp \left( \frac{(1-\gamma)^2 q}{2} \hat{\sigma}^2 \left( 1 - \frac{q}{\gamma} \right) \right) \left( \frac{(1-\gamma)^2 q}{2} \left( 1 - \frac{q}{\gamma} \right) \right) + \frac{1}{2} \lambda \frac{1}{\hat{\sigma}^2} \end{aligned} \quad (59)$$

From the equation above we can obtain:

$$\frac{\partial F}{\partial \lambda} = \frac{1}{\hat{\sigma}^2} > 0 \quad (60)$$

$$\begin{aligned} \frac{\partial F}{\partial \hat{\sigma}^2} &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp \left( -\frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \right] \\ &\times \exp \left( \frac{(1-\gamma)^2 q}{2} \hat{\sigma}^2 \left( 1 - \frac{q}{\gamma} \right) \right) \left( \frac{(1-\gamma)^2 q}{2} \left( 1 - \frac{q}{\gamma} \right) \right)^2 - \frac{1}{2} \lambda \frac{1}{(\hat{\sigma}^2)^2} < 0 \end{aligned} \quad (61)$$

According to the implicit function theorem, we can easily show

$$\frac{\partial \hat{\sigma}^2}{\partial \lambda} = -\frac{\frac{\partial F}{\partial \lambda}}{\frac{\partial F}{\partial \hat{\sigma}^2}} > 0 \quad (62)$$

■

### A.9. Proof of Proposition 3

**Proof.** Similarly to the proof in A.8, we can show that

$$\begin{aligned} \frac{\partial F}{\partial \sigma^2} &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp \left( -\frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \left( -\frac{q}{2\gamma} (1-\gamma) \left( 1 - \frac{q}{\gamma} + \gamma \right) \right) \right] \\ &\quad \times \exp \left( \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \hat{\sigma}^2 \left( 1 - \frac{q}{\gamma} \right) \right) \left( \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \left( 1 - \frac{q}{\gamma} \right) \right) < 0 \end{aligned} \quad (63)$$

$$\frac{\partial \hat{\sigma}^2}{\partial \sigma^2} = -\frac{\frac{\partial F}{\partial \sigma^2}}{\frac{\partial F}{\partial \hat{\sigma}^2}} < 0 \quad (64)$$

■

### A.10. Proof of Proposition 4

**Proof.**

$$\begin{aligned} \frac{\partial F}{\partial \sigma^2} &= Y_0^{-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \beta^{\frac{q}{\gamma}} \exp \left( -\frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \right] \\ &\quad \times \exp \left( \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \hat{\sigma}^2 \left( 1 - \frac{q}{\gamma} \right) \right) \left( \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \left( 1 - \frac{q}{\gamma} \right) \right) > 0 \end{aligned} \quad (65)$$

$$\frac{\partial \hat{\sigma}^2}{\partial Y_0} = -\frac{\frac{\partial F}{\partial Y_0}}{\frac{\partial F}{\partial \hat{\sigma}^2}} > 0 \quad (66)$$

■

### A.11. Proof of Proposition 5

**Proof.**

$$\begin{aligned} \frac{\partial F}{\partial \beta} &= \frac{Y_0^{1-\gamma}}{1-\gamma} \sum_{q=0}^{\gamma} \left[ \binom{\gamma}{q} \frac{q}{\gamma} \beta^{\frac{q}{\gamma}-1} \exp \left( -\frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right) \sigma^2 + \frac{1}{2} \left( \frac{1-\gamma}{\gamma} q \right)^2 \sigma^2 \right) \right] \\ &\quad \times \exp \left( \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \hat{\sigma}^2 \left( 1 - \frac{q}{\gamma} \right) \right) \left( \frac{(1-\gamma)^2}{2} \frac{q}{\gamma} \left( 1 - \frac{q}{\gamma} \right) \right) < 0 \end{aligned} \quad (67)$$

$$\frac{\partial \hat{\sigma}^2}{\partial \beta} = -\frac{\frac{\partial F}{\partial \beta}}{\frac{\partial F}{\partial \hat{\sigma}^2}} < 0 \quad (68)$$

■

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

First, equation (33) can be rewritten as

$$\frac{1}{2} \frac{1}{Y_0} \beta^{\frac{1}{2}} \exp \left( \frac{3}{8} \sigma^2 \right) = \lambda \frac{1}{\hat{\sigma}^2 \exp \left( \frac{1}{8} \hat{\sigma}^2 \right)} \quad (69)$$



Then taking the logarithm on both sides of (69)

$$\log\left(\frac{\beta^{\frac{1}{2}}}{2Y_0}\right) + \frac{3}{8}\sigma^2 = \log(\lambda) - \log(\hat{\sigma}^2) - \frac{1}{8}\hat{\sigma}^2 \quad (70)$$

From the information constraint in equation (16), we have  $2\kappa = \log(\sigma^2) - \log(\hat{\sigma}^2)$  and  $\hat{\sigma}^2 = \exp(-2\kappa)\sigma^2$ . Therefore, (70) becomes

$$\begin{aligned} \log\left(\frac{\beta^{\frac{1}{2}}}{2Y_0}\right) + \frac{3}{8}\sigma^2 &= \log(\lambda) - \log(\sigma^2) + 2\kappa - \frac{\sigma^2}{8}\exp(-2\kappa) \\ &\approx \log(\lambda) - \log(\sigma^2) + 2\kappa - \frac{\sigma^2}{8}(-2\kappa + 1) \\ &\approx \log(\lambda) + \left(2 + \frac{\sigma^2}{4}\right)\kappa - \log(\sigma^2) - \frac{\sigma^2}{8} \\ \implies \kappa &\approx \frac{1}{2 + \frac{\sigma^2}{4}} \log\left(\frac{\beta^{\frac{1}{2}}\sigma^2 \exp(\frac{1}{2}\sigma^2)}{2\lambda Y_0}\right) \end{aligned} \quad (71)$$

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

$$\kappa^* = \begin{cases} \frac{1}{2 + \frac{\sigma^2}{4}} \log\left(\frac{\beta^{\frac{1}{2}}\sigma^2 \exp(\frac{1}{2}\sigma^2)}{2\lambda Y_0}\right), & \text{if } \lambda < \frac{\beta^{\frac{1}{2}}\sigma^2 \exp(\frac{1}{2}\sigma^2)}{2} \\ 0, & \text{otherwise} \end{cases}$$

### A.13. Pseudocode

- Step 1: Build the simplex with equi-spaced grid for capital return  $A_1$
- Step 2: For each simplex point, define  $f(K_1, A_1)$  and initialize the value function with 0.
- Step 3: For each simplex point, find the  $f^*(K_1, A_1)$  such that the objective function (42) is maximized.
- Step 4: Optimization using `csminwel` and iterate on the value function up to convergence.
- Step 5: Obtain the optimal attention choice from the binding information constraint and  $f^*(K_1, A_1)$

## Appendix B

### B.1. Parameter values for Figure 1

Parameter	Value
Initial endowment $Y_0$	2
Prior Variance $\sigma^2$	0.08
Discounting factor $\beta$	0.98
Information cost $\lambda$	0.015
Range for posterior variance $\hat{\sigma}^2$	(0.001, $\sigma^2 - 0.0001$ )
Risk aversion degree $\gamma$	2, 3, 4, 5, 6, 7, 8, 9, 10

### B.2. Parameter values for Figure 2

Parameter	Value
Initial endowment $Y_0$	[2,...,2.5]
Capital return space $A_1$	[0.9,...,1.1]
Savings space $K_1$	[0.1 * $Y_0$ ,...,0.9 * $Y_0$ ]
Discounting factor $\beta$	0.98
Joint distribution per simplex point, $f(K_1, A_1)$	60 × 60
Marginal distribution of $K_1$ ,	60 × 1
Marginal distribution of $A_1$ ,	60 × 1
Information cost $\lambda$	0.015
Risk aversion degree $\gamma$	2

## References

- Angeletos, G.-M. (2007). Uninsured idiosyncratic investment risk and aggregate saving. *Review of Economic Dynamics*, 10:1–30.
- Barlevy, G. and Veronesi, P. (2000). Information acquisition in financial markets. *Review of Economic Studies*, 67:79–90.
- Batchuluun, A., Luo, Y., and Young, E. (2018). Portfolio choice with information-processing limits. *Annals of Economics and Finance*.
- Benhabib, J., Bisin, A., and Zhu, S. (2015). The wealth distribution in bewley economies with capital income risk. *Journal of Economic Theory*, 159:489–515.
- 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 Atlas Shrug? The Economic Consequences of Taxing the Rich, Harvard University Press*.
- Cover, T. and Thomas, J. (2006). Elements of information theory. *Wiley, Hoboken, NJ*.
- Dréze, J. and Modigliani, F. (1972). Consumption decisions under uncertainty. *Journal of Economic Theory*, 5:308–335.
- Eeckhoudt, L., Gollier, C., and Treich, N. (2005). Optimal consumption and the timing of the resolution of uncertainty. *European Economic Review*, 49:761–773.
- Eeckhoudt, L. and Schlesinger, H. (2008). Changes in risk and the demand for saving. *Journal of Monetary Economics*, 55:1329–1336.
- Epstein, L. (1980). Decision making and the temporal resolution of uncertainty. *International Economic Review*, 21(2):269–283.
- 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):393–408.

- 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).
- Kimball, M. (1990). Precautionary savings 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., 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. (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. and Wu, T. (2010). The puzzling evolution of the home bias, information processing and financial openness. *Journal of Economic Dynamics and Control*, 34:875–896.
- Paciello, L. (2012). Monetary policy and price responsiveness to aggregate shocks under rational inattention. *Journal of Money, Credit and Banking*, 44(7):1375–1399.
- 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.
- 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., Loewenstein, 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. (2015). Rational inattention and monetary economics. *Handbook of Monetary Policy*.
- Skinner, J. (1988). Risky income, life cycle consumption, and precautionary savings. *Journal of Monetary Economics*, 22:237–255.
- Tutino, A. (2013). Rational inattentive consumption choices. *Review of Economic Dynamics*, 16:421–439.
- Van Nieuwerburgh, S. and Veldkamp, L. (2009). Information immobility and the home bias puzzle. *Journal of Finance*, 56(5):678–695.
- Van Nieuwerburgh, S. and Veldkamp, L. (2010). Information acquisition and under-diversification. *Review of Economic Studies*, 77:779–805.
- Wiederholt, M. (2010). Rational inattention. *Prepared for the New Palgrave Dictionary of Economics*.
- Woodford, M. (2009). Information-constrained state-dependent pricing. *Journal of Monetary Economics*, 56:100–124.