Unveiling the role of rational inattention: Tax incentives and participation in commercial pension insurance

Zheng Chen[a], Jia Huang[b,*]

[a]*School of Management, Guangdong University of Technology, Guangzhou 510520, PR China*
[b]*Lingnan College, Sun Yat-sen University, Guangzhou 510275, PR China*

Abstract

This paper examines why tax incentives fail to stimulate participation in China's third-pillar commercial pension insurance, emphasizing the role of rational inattention. Using household survey data from 2014-2022 and a difference-in-differences-in-differences design, we find the tax-deferred pilot policy did not increase enrollment, with rational inattention/limited financial literacy explaining much of its ineffectiveness. We develop a dynamic consumption-portfolio model featuring costly information acquisition, and then resolve limitations of standard models through a dynamic framework with distinct savings channels and policy-focused rational inattention. The models show that rational inattention distorts perceptions of tax benefits and wage growth, raising participation costs, while multiple savings channels dilute incentives. Only households with higher financial literacy substantially respond to the policy. Our results reveal how cognitive frictions undermine pension reform and offer implications for designing behaviorally-informed retirement schemes.

JEL classification: C61; G11; D14

Keywords: Rational inattention; Tax incentives (benefits); Third-pillar pension reforms; Multiple savings channels; Policy complexity costs

1. Introduction

Currently, China is faced with the increasingly severe issues of low fertility rates and aging populations. A reasonable three-pillar pension system is gradually becoming a prerequisite for the sustainable development of the Chinese old-age security system and social stability. On April 2, 2018, the Ministry of Finance of the People's Republic of China and four other departments jointly issued the "Notice on Carrying out Pilot Programs for Individual Tax-Deferred Commercial Pension Insurance". They announced that one-year pilot programs for individual tax-deferred commercial pension insurance would be launched in Shanghai Municipality, Fujian Province and

*Corresponding author.
Email addresses: chenzheng@gdut.edu.cn (Zheng Chen), huangj535@mail.sysu.edu.cn (Jia Huang)

Suzhou Industrial Park from May 1, 2018. Individuals purchasing commercial pension insurance products via commercial pension accounts in the pilot regions will be eligible for tax incentives on their contributions, accumulation and withdrawals of tax-deferred commercial pension insurance, which marked the implementation of the third-pillar pension system of China.[1] Since this policy was introduced, the pilot program for the tax-deferred commercial pension insurance of individuals in China has progressed smoothly, albeit with an overall modest business scale. As of December 2021, pilot regions had cumulatively achieved a premium income of RMB 628.8065 million, with 52,130 insurance policies issued. In this context, a core issue that policy authorities and scholars are widely concerned about is whether tax incentive policies are fulfilling their intended roles. Tax incentive policies are aimed at stimulating the enthusiasm of individuals or families for accumulating pension reserves by reducing their tax burden when they participate in commercial pension insurance. The effectiveness of this policy concerns whether it can attract more people to actively get involved in constructing the third pillar of the pension system, which thereby broadens the sources of pension funds and enhances the sustainability of the pension system. Meanwhile, the effectiveness of this policy has a direct impact on the construction and improvement of the multi-tiered old-age security system of a country. Therefore, whether tax incentive policies are truly effective in promoting the development of the third-pillar pension system has become a pressing practical question that needs to be answered.

Research on the empirical or theoretical directions of the impact of tax incentive policies on behaviors like savings and insurance demand has a long history. In terms of empirical testing, scholars have successively examined the influence of tax incentive policies on individual or household financial decisions. For example, studies have explored their impact on health insurance purchase (Gruber & Poterba, 1994); on life insurance purchase (Sauter et al., 2015); and on income and pension wealth accumulation (Basten et al., 2017; Beshears et al., 2017). In the aspect of theoretical research, scholars also have examined the influence of tax incentive policies on individual or household financial decisions. For example, studies have explored their impact on household investment and financial optimization (Howe & Mistic, 2003; Dammon et al., 2004; Marekwica et

[1]The issuance of the "Notice on Carrying out Pilot Programs for Individual Tax-Deferred Commercial Pension Insurance" in 2018 marked the practical implementation of the third pillar of China's pension system. On December 12, 2024, the Ministry of Human Resources and Social Security of the People's Republic of China, along with other relevant departments, jointly released the "Notice on the Comprehensive Implementation of the Individual Pension System," signifying a new stage of development for China's pension finance. The issue of pension reform in China has long been a focal point of research, with notable contributions from scholars such as Feldstein (1999) and Jia (2017), among others.

al., 2013; Smith & Seay, 2016); on policy heterogeneity and behavioral responses (Zwick & Mahon, 2017; Hines & Park, 2019); and on household welfare (Brown et al., 2017). Recently, in integrating theoretical and empirical analyses, Xiao et al. (2024) developed a theoretical framework and conducted empirical analysis, using firm databases in China to examine the impact of tax incentives on firms' social security contributions. The findings reveal that tax incentives significantly enhance firms' social security payments.

Due to the limited time frame of policy pilot implementation, existing literature has relatively inadequately studied the impact of tax deferral policies on the asset allocation of Chinese households and, more specifically, the participation of individuals in pension plans. In light of the low participation rate in the commercial pension insurance sector of China, no research has incorporated behavioral and psychological factors up to now, and the realistic characteristic that individuals often exhibit rational inattention is overlooked. Simon (1955) proposed that humans possess bounded rationality, which means that "while individuals are consciously rational, this rationality is limited". Sims (2003) argued that bounded rationality entails "pursuing rationality subjectively, but only being able to achieve it to a limited extent objectively". He sought to identify the "constraints limiting decision-makers' information processing capabilities" and believed that "the crucial issue is how much information we can process rather than the availability of information". What kind of information can our knowledge enable us to analyze, and what meaningful parts can we extract from it? Owing to limited information processing capabilities and costs associated with information acquisition, information processing capacity imposes constraints on economic agents, which causes individuals to rationally choose to overlook certain relatively unimportant information, a phenomenon known as rational inattention.

Based on the aforementioned analysis, this paper attempts to explore, from the relatively novel perspective of rational inattention, how tax incentives influence participation in commercial pension insurance and focuses on analyzing the mechanism through which rational inattention operates within the framework of pension investment portfolios. The contributions of this study primarily lie in the following three aspects.

Firstly, this study expands the literature on tax-deferred personal pensions and enriches empirical research on the effectiveness of China's tax-deferred policies for pension insurance. In practical research, rational inattention is difficult to observe directly, but financial literacy can be measured through standardized questions. The complexity of financial markets, the diversity of products, and information asymmetry impose high cognitive costs on individual investors to acquire and understand relevant information, such as comprehending the risk-return profiles of different investment products. Financial literacy influences the efficiency and cost of information processing; investors

with higher financial literacy tend to be better at acquiring information, thereby reducing their degree of bounded rationality and influencing asset allocation decisions, including the purchase of commercial pension insurance. Bounded rationality implies that not all economic agents will invest in the same financial instruments, significantly affecting their investment and consumption choices (Calvet et al., 2009; Dabla-Norris et al., 2015; Zhu, 2018). Improving financial literacy increases individuals' willingness to participate in financial markets (Van Rooij et al., 2011; Yang et al., 2023; Wang et al., 2024). Particularly in China, where financial literacy is relatively low, individuals with bounded rationality are influenced by their inherent information processing capabilities when handling financial information, and financial literacy is closely related to these capabilities. Households with higher financial literacy are more likely to understand the risk and return characteristics of financial products, reducing uncertainty arising from information processing (Yin et al., 2014; Feng et al., 2019; Wu et al., 2022). Therefore, does rational inattention affect the policy effectiveness of tax-deferred pilots and, consequently, influence individual participation in commercial pension insurance? To answer these questions, the first research endeavor of this paper addresses the practical paradox of "significant incentives but insufficient participation" in China's tax-deferred policies for the third pillar of pensions. Utilizing data from five waves (2014-2022) of the China Family Panel Studies (CFPS), this study employs a DDD model to evaluate the tax-deferred pilot policies, revealing that the failure of China's tax-deferred policies to achieve expected outcomes may be attributed to rational inattention. This provides new empirical evidence for assessing the effectiveness of China's tax-deferred policies for pension insurance. The reason for using a DDD model is that the effects of tax-deferred policies vary across different pilot regions, and there are significant differences in residents' financial literacy among these regions. Shanghai, with its relatively high financial literacy, has achieved better pilot results, providing a rare natural experiment scenario for evaluating whether tax incentives promote the development of the third pillar of pensions. Intuitively, financial literacy may have a positive promoting effect on tax-deferred commercial pension insurance, while rational inattention largely depends on the inherent level of financial literacy. Individuals with varying degrees of rational inattention differ in their ability to acquire and process financial information, thereby influencing their purchase decisions regarding tax-deferred commercial pension insurance. Furthermore, considering that existing studies primarily use financial knowledge as an indicator of financial literacy to quantify the impact of rational inattention, this approach overlooks the influence of financial asset allocation behavior. This paper uses non-standardized questions about financial behavior to measure financial literacy, better capturing the dynamic impact of rational inattention.

Secondly, this study expands the research in the field of rational inattention. Within the frame-

work of pension investment portfolios, it breaks away from the assumption of complete rationality in traditional life-cycle models and constructs a dynamic decision-making model that incorporates wage growth uncertainty and tax complexity. By obtaining analytical solutions, it finds that a reduction in the degree of rational inattention helps consumers improve the accuracy of their wage growth rate predictions and tax rate perceptions, lowers information acquisition costs, and thereby promotes the purchase of commercial pension insurance. This provides a valuable addition to the existing literature on pension investment portfolio issues. Furthermore, as a fundamental assumption in behavioral finance, rational inattention often influences financial decisions through uncertainties in wealth, stock prices, or wages (Sims, 2003; Luo, 2017; Luo & Young, 2010; Miao et al., 2022; Hansen et al., 2025). Therefore, the ability to make reasonable predictions about future wage growth rates and possess accurate knowledge of the tax-deferred policies for China's third-pillar commercial pension insurance will affect how rational inattention influences participation in tax-deferred commercial pension insurance. The second research endeavor of this paper is to construct a pension investment-consumption-contribution model based on the assumption of rational inattention and derive the optimal strategy, further revealing the mechanism through which rational inattention operates. The basic results in closed-form indicate that attention allocation is a crucial mechanism through which rational inattention exerts an influence on the effectiveness of tax deferral policies. Specifically, when the level of rational inattention decreases, it enhances the precision of predictions regarding wage growth rates and perceptions of tax rates. This reduction also substantially cuts down on the costs associated with acquiring information, ultimately encouraging the purchase of commercial pension insurance.

Thirdly, this study further advances to embed rational inattention within a multi-channel savings model (Dammon et al., 2004; Poterba et al., 1995), which allows us to quantify the resulting welfare losses and distributional effects of pension policy complexity. Specifically, we develop a dynamic stochastic life-cycle framework that departs from full-information assumptions by incorporating dual cognitive frictions: endogenous inattention to policy details, which distorts the perceived effective tax benefit τ_{eff}, and an explicit cost of processing complex rules χ_{eff}. This setup explains the persistent puzzle of low pension uptake despite generous subsidies: when the perceived cost of understanding the policy outweighs its perceived benefit, households rationallythough suboptimallydivert savings toward more accessible taxable accounts, incurring deadweight efficiency losses. The model further reveals how complexity initiates a self-reinforcing cycle of inattention and misunderstanding, disproportionately discouraging contributions from less financially literate households and concentrating tax subsidies among the more attentive (Chetty et al., 2014). These results advance the theoretical literature on pension portfolios by formalizing how information fric-

tions in multi-pillar systems reshape savings trade-offs, and they provide a novel micro-founded framework for designing policies that mitigate complexity-induced participation barriers (Chetty, 2015; Beshears et al., 2018).

The remainder of the paper is organized as follows. Section 2 provides empirical evidence from China on the effects of tax-deferred pension pilots and the role of rational inattention. Section 3 develops a theoretical model to analyze how rational inattention shapes optimal contribution and investment decisions under wage growth and tax complexity. Section 4 presents numerical simulations to illustrate the theoretical results. Section 5 extends the model to incorporate multiple saving channels and policy complexity as a core informational friction. Section 6 concludes. Technical proofs, additional numerical tests, and extended model setups are provided in the appendices.

2. Empirical evidence on rational inattention and the pilot program of tax-deferred commercial pension insurance

2.1. Stylized facts

As shown in Table 1, the purchase of tax-deferred commercial pension insurance varies across different pilot regions. Except for Shanghai, the proportion of premium income in Fujian and Suzhou Industrial Park has declined compared to the initial pilot stage in October 2018, reflecting an overall low willingness to purchase tax-deferred commercial pension insurance. According to the "2022 China Household Financial Literacy Report" jointly released by the Shanghai Advanced Institute of Finance (SAIF) at Shanghai Jiao Tong University and Charles Schwab, there are also regional disparities in financial literacy levels across China's provinces. Shanghai ranks highest with an average score of 72.5, and the tax-deferred pilot policy in this region has also achieved relatively favorable results.

Table 1: Status of the tax-deferred commercial pension insurance pilot program as of the end of 2021.

Pilot Regions	Premium Income (in 10,000 yuan)	Premium Proportion	Initial Pilot Proportion
Shanghai	49754.56	79.13%	74.15%
Fujian	9786.30	15.56%	19.74%
Suzhou Industrial Park	3339.79	5.31%	6.11%

Data source: Compiled based on data from the former China Insurance Regulatory Commission.

Furthermore, we employ binned scatter plots to visualize the relationship between financial literacy and commercial pension insurance purchase in the pilot areas. Our preliminary investigation reveals that higher levels of financial literacy among residents in the pilot areas are associated with

increased purchases of commercial pension insurance. The aforementioned analysis of typical facts suggests a possible correlation between rational inattention and the effectiveness of the tax-deferred pilot policy, but it does not strongly establish a causal relationship between the two. In the subsequent econometric analysis, we will employ a more rigorous triple difference model and conduct a series of robustness tests to identify the impact of rational inattention on the effectiveness of the tax-deferred pilot policy.

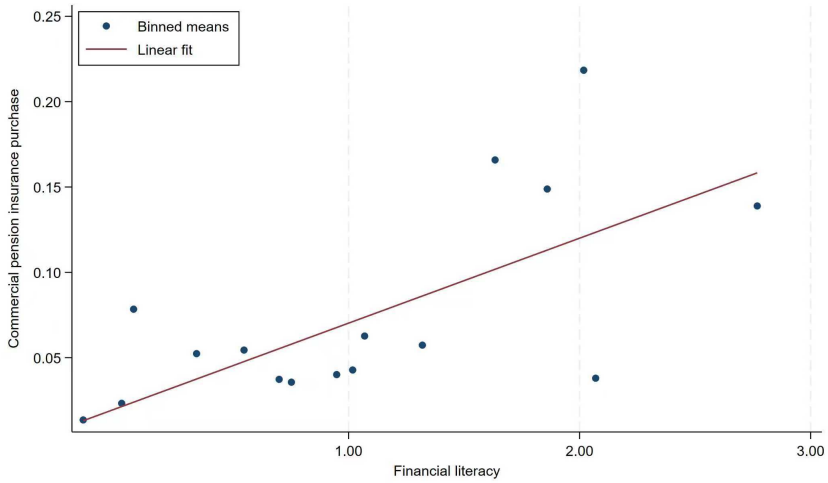


Figure 1: Commercial Pension Insurance Purchase by Financial Literacy Level. Note: The analysis employs data drawn from CFPS.The independent variable (financial literacy) is divided into 20 equal-sized bins based on quantiles (vigintiles). The mean of the dependent variable (pension insurance purchase) is computed for each bin and plotted against the bin's midpoint. The linear fit is plotted after absorbing year and province fixed effects.

2.2. Econometric model

When examining the policy effectiveness of tax-deferred pilot programs on commercial pension insurance, it is essential to account for both regional policy heterogeneity at the pilot area level and the degree of rational inattention at the individual level. Prior to conducting the DDD analysis, this study first employs a simple difference-in-differences (DID) model to provide a preliminary estimate of the policy impact of tax-deferred pilots on household participation in commercial pension insurance. The model specification is as follows:

$$Y_{ipt} = \alpha_0 + \beta_1 Treat_p \times Post_t + \mu_i + \delta_t + \alpha_1 X_{ipt} + \varepsilon_{ipt}. \tag{1}$$

Based on the evaluation of the effectiveness of the tax-deferred pilot program, this paper focuses on analyzing whether rational inattention plays a part in the impact of the tax-deferred policy on household participation in commercial pension insurance. Under bounded rationality, individuals

processing of financial information is influenced by their inherent information processing capabilities. Financial literacy is closely related to these capabilities. Households with a higher level of financial literacy are more likely to understand the risks and returns of financial products, which reduces the uncertainty arising from information processing (Yin et al., 2014; Wu et al., 2022; Feng et al., 2019). Thus, based on Equation (1), this paper used financial literacy as a measure of rational inattention to construct a DDD model to estimate the heterogeneous impact of rational inattention on the effectiveness of the tax-deferred pilot policy. The specific model was set as follows:

$$Y_{ipt} = \alpha_0 + \beta_1 Treat_p \times Post_t \times Finit_i^h + \mu_i + \lambda_{pt} + \alpha_1 X_{ipt} + \varepsilon_{ipt}, \tag{2}$$

where i, p and t represent household, province, and time, respectively; the dependent variable Y_{ipt} represents the participation of household i in province p in commercial pension insurance in year t; $Post_t$ is a treatment variable indicating whether the policy was implemented in year t, with $Post_t = 1$ indicating the year of implementation and subsequent years, and 0 otherwise; $Treat_p$ is a grouping variable indicating whether province p, where household i is located, is a policy pilot region, with $Treat_p = 1$ for pilot regions and 0 otherwise; $Finit_i^h = 1$ indicates higher financial literacy of the household, and 0 otherwise; μ_i, δ_t, and λ_{pt} represent household-level individual fixed effects, joint time fixed effects by year, and joint fixed effects by province and year, respectively. In Equation (1), they are used to control for the level terms of $Treat_p$ and $Post_t$, while in Equation (2), they are employed to control for the level terms of $Treat_p$, $Post_t$ and $Finit_i^h$, as well as their pairwise interaction terms; X_{ipt} only includes control variables reflecting household- and household-head-level characteristics, with features that vary across provinces and time being absorbed by δ_t in Equation (1) and λ_{pt} in Equation (2), respectively; ϵ_{ipt} is a random disturbance term. Since the policy implementation occurs at the province level, robust standard errors clustered at the province level are reported. In Equations (1) and (2), the coefficient of the double difference term $Treat_p \times Post_t$ measures the policy effect of the tax-deferred pilot on household participation in commercial pension insurance, while the coefficient of the triple difference term $Treat_p \times Post_t \times Finit_i^h$ reflects the heterogeneous impact on household participation in commercial pension insurance between high and low rational inattention groups when subjected to the shock of the tax-deferred policy.

Although the DDD model relaxed the parallel trends assumption for both treatment and control groups, a parallel trends test was conducive to further ensuring the validity of the estimation results. This paper adopted an event study approach to test the parallel trends assumption in double and triple models before the policy shock.

For the DID model, we replace $Post_t$ in Equation (1) with year-specific indicator variables

$year(t = t_0 + j)$, with the base period set to 2016, which is one period before the tax-deferred pilot (i.e., $t_0 = -1$ and $j = 0$). The rest of the setup remains consistent with Equation (1). If the parallel trends assumption for the double difference holds, the estimated coefficient β_{-1} for the pre-shock dummy variable should be insignificant.

$$Y_{ipt} = \alpha_0 + \sum_{j=-2}^{2} \beta_j \times Year(t = t_0 + j) \times Treat_p + \mu_i + \delta_t + \alpha_1 X_{ipt} + \varepsilon_{ipt}. \tag{3}$$

For the DDD model, similarly, we replace $Post_t$ in Equation (2) with year-specific indicator variables. Specifically, the variable representing two periods before the pilot implementation, $year(t = -2)$, takes a value of 1 in 2014; the variable representing one period before the pilot implementation, $year(t = -1)$, takes a value of 1 in 2016; the variable representing one period after the pilot implementation, $year(t = 1)$, takes a value of 1 in 2020; the variable representing two periods after the pilot implementation, $year(t = 2)$, takes a value of 1 in 2020; the base period is set as the period when the pilot is implemented, $year(t = 0)$ (i.e., 2018). If the parallel trends assumption for the DDD holds, the estimated coefficient β_{-1} for the pre-shock dummy variable should be insignificant.

$$Y_{ipt} = \alpha_0 + \sum_{j=-1}^{3} \beta_j \times Year(t = t_0 + j) \times Treat_p \times Finit_i^h + \mu_i + \lambda_{pt} + \alpha_1 X_{ipt} + \varepsilon_{ipt}. \tag{4}$$

2.3. Data Sources and Variable Selection

2.3.1. Data Sources

The data were derived from the CFPS conducted by the Institute of Social Science Survey (ISSS) at Peking University. The CFPS sample covered 25 provinces (including municipalities directly under the central government and autonomous regions). All the households and their members identified through the baseline survey in 2010 served as the survey respondents. We used the households interviewed in the CFPS databases for the years 2014, 2016, 2018, 2020 and 2022 as the research sample. The data were processed as follows: (1) Basic information from individual and household questionnaires was merged, and duplicate observations were eliminated; (2) Samples with severe missing values were deleted; (3) Samples only existing during the policy period, before implementation or after implementation were excluded. Additionally, we applied 1% winsorization to all continuous variables to mitigate the influence of extreme values.

2.3.2. Variable Selection

(1) Dependent Variable. Participation in Household Commercial Pension Insurance Y_{ipt}. Based on the question in the CFPS survey: "Which types of pension insurance programs are you enrolled

in?", if a household member participates in commercial pension insurance, Y_{ipt} is 1; otherwise, it is 0.

(2) Main Explanatory Variable. Tax Deferral Pilot $Treat_p \times Post_t$. Taking 2018 as an exogenous policy shock for tax-deferred commercial pension insurance, $Post_t$ is 1 for the years 2018, 2020, and 2022, and 0 for the years 2014 and 2016. $Treat_p$ indicates whether province p is a pilot region for implementing tax-deferred commercial pension insurance. $Treat_p$ is 1 for pilot regions and 0 for non-pilot regions. In the process of selecting the treatment and control groups, the following approach was taken: Since the pilot regions for implementing individual tax-deferred commercial pension insurance include Shanghai Municipality, Fujian Province, and Jiangsu Province, and Jiangsu Province only designated the Suzhou Industrial Park as a pilot area, including the entire province as a pilot region would underestimate the policy effect of tax deferral. Therefore, Jiangsu Province was excluded from the sample.

Financial Literacy $Finit$. Existing literature measures financial literacy through standardized questions about financial knowledge, such as interest rate calculation, understanding of inflation, risk perception, and so on. However, the CFPS questionnaire only included some standardized questions related to financial knowledge in its 2014 and 2018 editions, and the specific content of these questions was not entirely consistent across years. Notably, only the 2014 questionnaire featured a complete module on financial knowledge, encompassing 26 aspects including interest rate calculation, currency purchasing power, investment risk, financial products, financial planning, borrowing behavior, investment choices, risk preference, etc. Nevertheless, if the financial knowledge data from 2014 were used to fill in for other years, the resulting measure of financial literacy would represent a static stock, which clearly fails to meet the parallel trend test requirements for the DDD model.

In fact, financial literacy is a multidimensional and comprehensive indicator that reflects an individual's understanding of financial concepts, products, and risks, as well as their ability to effectively manage financial resources and enhance financial well-being by utilizing this knowledge and skills (PACFL, 2008). It encompasses fundamental economic and financial knowledge, along with the capacity to apply this knowledge and skills to effectively manage financial resources throughout one's lifetime for financial well-being (PACFL, 2008). Its core essence includes the combination of awareness, knowledge, behaviors, skills, attitudes, and the alignment of attitudes with behaviors necessary for making appropriate financial decisions to achieve personal financial well-being (OECD INFE, 2011). Due to data limitations, we are unable to obtain all standardized or non-standardized questions covering financial knowledge, financial behaviors, financial skills, and financial attitudes. However, financial knowledge, skills, and attitudes are, to some extent, reflected in financial be-

haviors. For instance, households with extensive financial knowledge, a healthy financial attitude, and good financial skills in handling basic financial affairs are often more willing to engage in a wider variety of financial behaviors. Therefore, based on qualitative and quantitative analyses of the survey data from the CFPS questionnaire, we will characterize financial literacy through financial behaviors, primarily evaluating it from aspects such as past actual financial activities, liabilities, and the holding of financial assets. This approach aims to compensate for the lack of continuous financial knowledge data in the CFPS while better reflecting the dynamic process of rational inattention. We selected four non-standardized questions from the CFPS questionnaire that are closely related to financial behaviors. These include: 1) "Does your household have any outstanding bank loans for purchasing, constructing, or renovating a home?"; 2) "Apart from loans for purchasing, constructing, or renovating a home, does your household currently have any other outstanding bank loans?"; 3) "Does your household currently hold any financial products, such as stocks, funds, government bonds, trust products, foreign exchange products, or internet financial products like Yu'ebao?"; and 4) "Does your household currently hold any fixed-term deposits?".[2] Based on the responses from interviewed households to these questions, we generated a continuous variable ranging from 0 to 4. The more questions answered with "yes," the more diversified the household's use of financial instruments and asset allocation strategies, indicating a higher level of financial literacy.[3] Among these questions, the first and second questions reflect respondents' awareness of interest rate levels by examining whether they hold loans and fixed-term deposits. The third question demonstrates respondents' awareness of financial risks through their holding of financial products. The fourth question reflects respondents' understanding of interest rates and inflation through their holding of fixed-term deposits. Households are grouped based on the median financial literacy level ($Finit$). If the financial literacy of household i exceeds the median financial literacy level of all sample households, then $Finit_i^h$ is assigned a value of 1, indicating that household i exhibits more diversified financial behaviors, has relatively higher financial literacy, and demonstrates a lower degree of rational inattention. Otherwise, $Finit_i^h$ is assigned a value of 0.

(3) Control Variables. At the household level, control variables include: *Size*, which is equal to the number of household members; *Income*, which is the logarithm of total household income;

[2]For the fourth question related to financial literacy, "Does your household currently hold any time deposits?", the response is based on the amount of time deposits reported in the CFPS questionnaire, i.e., whether the household's time deposits are greater than 0.

[3]Since the non-standardized issues related to financial literacy yield a binary (0-1) variable, factor analysis and entropy weight method cannot be employed for index construction, as these methods are applicable to continuous or ordinal categorical variables.

urban status *Urban*, which is 1 for non-agricultural households and 0 otherwise; other housing rental income *Houserent*, which is 1 if the household rents out housing and 0 otherwise;[4] dependency ratio *Depratio*, which is the sum of the child dependency ratio and the elderly dependency ratio, with the child dependency ratio calculated as the number of individuals aged 0-14 divided by the total household population, and the elderly dependency ratio calculated as the number of individuals aged 65 and over divided by the total household population.

At the household head level, the CFPS questionnaire defines the "household head" based on concepts such as the main decision-maker, financial respondent, property owner, etc. In this paper, the household head is selected based on whether the individual is the financial respondent. The control variables specifically include: *Age*, which is the logarithm of the respondent's age; *Age2*, which is calculated as the square of the variable "age" divided by 100; marital status *Marriage*, which is 1 if the respondent is married and 0 otherwise; and education level *Edu*, which is 1 if the respondent's highest educational attainment is bachelor's degree or above and 0 otherwise.

2.3.3. Descriptive statistics

The descriptive statistics for the variables utilized in this paper are illustrated in Table 2. The results show that approximately 6.104% of the households in the sample excluding Suzhou Industrial Park belong to tax-deferred pilot regions. The mean and median participation rates of households in commercial pension insurance are 0.037 and 0.000, respectively, with a standard deviation of 0.188, which indicates a generally low willingness to purchase commercial pension insurance among households. The mean and median scores of financial literacy as a measure of rational inattention are 0.546 and 0.000, respectively, with a standard deviation of 0.641, which reflects a relatively low average level of financial literacy and significant variations in financial literacy among households.

2.4. Empirical results

2.4.1. Baseline regression

The baseline regression results for the DID and DDD models are reported in Table 3. Firstly, this paper initially estimated the policy effects of the tax-deferred pilot program on household participation in commercial pension insurance. The DID regression results without including control variables and controlling for household and household head characteristics are presented in

[4]If a household owns more than one property and rents out the extra property, obtaining significant rental income to support retirement life, it may reduce the household's willingness to participate in commercial pension insurance. Based on the related question in the CFPS questionnaire, "Do you rent out any of your properties for compensation?", we define other housing rental income *Houserent*, where *Houserent* is 1 if the household rents out housing and 0 otherwise.

Table 2: Descriptive statistics.

Variable	Mean	Standard Deviation	Median	Minimum	Maximum	Observations
Y_{ipt}	0.037	0.188	0.000	0.000	1.000	35190
$Treat_p \times Post_t$	0.033	0.178	0.000	0.000	1.000	35190
$Finit$	0.546	0.641	0.000	0.000	2.000	35190
$Favpay$	7.901	1.032	8.006	5.298	10.309	35190
$Size$	3.168	1.528	3.000	1.000	8.000	35190
$Asset$	12.584	1.328	12.588	8.700	15.846	35190
$Income$	11.626	0.816	11.644	9.462	13.647	35190
$Housrent$	0.064	0.244	0.000	0.000	1.000	35190
$Urban$	0.486	0.500	0.000	0.000	1.000	35190
$Depratio$	0.257	0.308	0.200	0.000	1.000	35190
Age	50.502	12.784	50.000	23.000	79.000	35190
$Age2$	27.139	13.039	25.000	5.290	62.410	35190
$Marriage$	0.883	0.321	1.000	0.000	1.000	35190
Edu	0.015	0.120	0.000	0.000	1.000	35190
$Inform$	12.758	6.869	13.000	0.000	28.000	35190
$WageSD$	0.674	0.640	0.851	0.009	3.196	7542
$Taxcomplex$	0.498	0.577	0.000	0.000	2.000	35190

columns (1) and (2). The estimated coefficient of the double difference term Treat $Treat_p \times Post_t$ is insignificant. This may be due to some differences between treatment and control groups.

Based on the analysis of characteristic facts, it is evident that variations exist in financial literacy levels across different pilot regions. For households, higher financial literacy is beneficial to reducing the uncertainty arising from the processing of financial information, which leads to significant differences in rational inattention. Hence, this paper further controlled for these differences in financial literacy using a DDD model to more accurately estimate the policy effects of the tax-deferred pilot program. The DDD regression results including control variables are demonstrated in column (3). It can be observed that the estimated coefficient of Treat $Treat_p \times Post_t \times Finit_i^h$ is significantly positive at the level of 5%. The findings indicate that when faced with the impact of tax-deferred policies, households with high financial literacy increase their purchases of tax-deferred

commercial pension insurance. Under the assumption of bounded rationality, a household's ability to process its own financial information primarily depends on its financial literacy. Households with high financial literacy are better positioned to develop an internal capacity for processing information about tax-deferred policies, enhancing their financial understanding of tax-deferred commercial pension insurance. This, in turn, reduces the adverse impact of rational inattention on their willingness to purchase commercial pension insurance. These results validate the influence of rational inattention. That is, the tax-deferred pilot program significantly stimulates the purchase of commercial pension insurance among households with higher financial literacy.

In addition, we also made an intriguing discovery: the coefficient for *Income* is significantly negative, indicating that households with higher incomes are less likely to purchase commercial pension insurance. This result rules out the income effect as a driving factor behind households' decisions to buy commercial pension insurance and further supports, from a different perspective, the influence of rational inattention. Further analysis reveals a possible explanation: the estimated coefficients for *Age* and *Age2* are both significantly negative, suggesting an inverted U-shaped relationship between the head of household's age and the family's decision to purchase commercial pension insurance. Generally, households with older heads of household tend to have higher income levels. This implies that families with young to middle-aged heads of household are the primary target group for commercial pension insurance, and an increase in income level does not necessarily equate to an increased likelihood of purchasing commercial pension insurance.

Table 3: Rational inattention and the policy effectiveness of tax-deferred pilot programs: Baseline regression.

Dependent variable: Y_{ipt}	(1)	(2)	(3)
$Treat_p \times Post_t \times Finit_i^h$			0.016**
			(0.007)
$Treat_p \times Post_t$	-0.001	-0.002	
	(0.003)	(0.003)	
$Size$		0.003*	0.003*
		(0.001)	(0.001)
$Income$		−0.005*	−0.005*
		(0.003)	(0.003)
$Housrent$		−0.003	−0.003
		(0.010)	(0.010)
$Urban$		0.004	0.003
		(0.008)	(0.008)
$Depratio$		−0.001	−0.001
		(0.004)	(0.004)
Age		0.002**	0.002**
		(0.001)	(0.001)
$Age2$		−0.002**	−0.002**
		(0.001)	(0.001)
$Marriage$		0.004	0.004
		(0.007)	(0.007)
Edu		0.005	0.008
		(0.008)	(0.007)
$Constant$		0.046	0.043
		(0.041)	(0.042)
Family fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	No
Province-year fixed effects	No	No	Yes
Observations	35176	35176	35168
Adjusted R^2	0.265	0.285	0.266

Note: Robust clustered standard errors at the province (or city) level are reported in parentheses; *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively. The same notation applies to subsequent tables.

2.4.2. Parallel trends test

There are systematic differences between households in the pilot regions and those in non-pilot regions, raising the possibility that even without the implementation of tax-deferred policies, significant disparities in commercial pension insurance participation would still persist between households in these two types of regions. To mitigate potential biases in estimation results caused by this non-randomness and heterogeneity, this paper employs two methods for parallel trend testing: First, it directly uses Equations (3) and (4) to conduct parallel trend tests; second, given that some characteristics of the treatment and control groups may change differently following policy implementation (Duflo, 2001; Moser & Voena, 2009; Fang et al., 2023), this paper further incorporates interaction terms between all control variables and the indicator for whether the tax-deferred pilot has commenced (post) into the aforementioned parallel trend tests. This allows these control variables to have differential impacts on household commercial pension insurance participation before and after the treatment.

The results of the parallel trend tests are presented in Table 4. Columns (1) and (2) demonstrate that the estimated coefficients for the pre-treatment dummy variables are not statistically significant, indicating that the DID model also satisfies the parallel trend assumption. Moreover, after the implementation of the tax-deferred pilot, the estimated coefficient for the post-treatment dummy variable $year(t = 1) \times Treat_p$ is significantly negative, while the coefficient for $year(t = 2) \times Treat_p$ is significantly positive. This, to a certain extent, reflects the time-lagged effect of the tax-deferred pilot policy. Similarly, Columns (3) and (4) show that the estimated coefficients for the pre-treatment dummy variables in the DDD model ($year(t = -1) \times Treat_p \times Finit_i^h$ and $year(t = -2) \times Treat_p \times Finit_i^h$) are not statistically significant, suggesting that prior to the tax-deferred pilot, there were no significant differences in commercial pension insurance participation between households with high and low financial literacy. This indicates that differences among households with varying levels of rational inattention within both the treatment and control groups followed a common trend, meeting the fundamental requirement of parallel trends. Additionally, after the tax-deferred pilot, the estimated coefficients for the post-treatment dummy variables ($year(t = 1) \times Treat_p \times Finit_i^h$ and $year(t = 2) \times Treat_p \times Finit_i^h$) are both statistically significant. Furthermore, after accounting for the pre-treatment characteristics of the control variables, the results of the parallel trend tests remain robust.

Table 4: Parallel trends test results.

Dependent variable: Y_{ipt}	(1)	(2)	(3)	(4)
$yeat(t=-2) \times Treat_p$	0.007	0.008		
	(0.005)	(0.007)		
$yeat(t=-1) \times Treat_p$	0.0004	0.002		
	(0.005)	(0.008)		
$yeat(t=1) \times Treat_p$	-0.012^{*}	-0.012^{*}		
	(0.007)	(0.007)		
$yeat(t=2) \times Treat_p$	0.022^{**}	0.023^{**}		
	(0.009)	(0.009)		
$yeat(t=-2) \times Treat_p \times Finit_i^h$			0.003	0.003
			(0.016)	(0.016)
$yeat(t=-1) \times Treat_p \times Finit_i^h$			0.003	0.003
			(0.010)	(0.010)
$yeat(t=1) \times Treat_p \times Finit_i^h$			0.029^{*}	0.028^{*}
			(0.014)	(0.014)
$yeat(t=2) \times Treat_p \times Finit_i^h$			0.039^{***}	0.039^{***}
			(0.012)	(0.013)
Controls $\times$ $Post_t$	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes
Family fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	No	No
Province-year fixed effects	No	No	Yes	Yes
Observations	35176	35176	35168	35168
Adjusted R^2	0.265	0.265	0.266	0.266

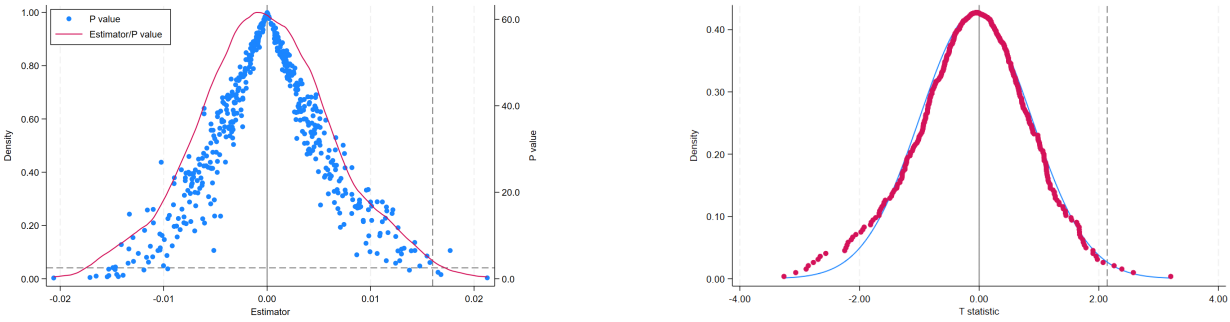
2.4.3. Robustness test

(1) Placebo test

In real-world scenarios, the implementation of tax-deferred policies is subject to various influencing factors, making it impossible to exhaustively account for all control variables. To eliminate interference from unobservable random factors and omitted variable bias, this paper further employs a placebo test to analyze whether the triple-difference regression results genuinely reflect the policy effects of the tax-deferred pilot.

Specifically, this paper randomly selects samples from the treatment group, financial literacy subgroups, and the timing of policy introduction to reconstruct a policy dummy variable. Subsequently, this reconstructed policy dummy variable is substituted into Equation (1) for regression. The aforementioned process is repeated 500 times, and kernel density plots are generated for the spurious estimated coefficients, along with their corresponding t-statistics and p-values, across these 500 regression iterations. As illustrated in Figure 2, the spurious estimated coefficients for $Treat_p \times Post_t \times Finit_i^h$ are concentrated around zero and mostly insignificant. Only a minuscule fraction of spurious estimated coefficients exhibit t-statistics exceeding those of the baseline regression results, which aligns with the expectations of the placebo test. This indicates that spurious policy variables have no significant influence on the tax deferral pilot program. As a consequence, it suggests that

the policy effects of the tax deferral pilot on commercial pension insurance participation are indeed influenced by the rational inattention levels of households.



(a) (b)

Figure 2: Placebo test, where (a): Spurious estimated coefficients and p-values; (b): t-statistics.

(2) Instrumental variable regression

In addition to the aforementioned omitted variable bias, endogeneity issues may also arise from reverse causality and measurement errors. First, the degree of rational inattention among households may be influenced by their participation in financial markets. Specifically, households with experience investing in financial markets are better positioned to enhance their understanding of tax-deferred policies, thereby affecting the impact of rational inattention on their participation in commercial pension insurance. Second, due to the inability to obtain continuous data on standardized questionssuch as interest rate calculations, inflation comprehension, and risk perceptionfrom survey questionnaires, the measurement variables for rational inattention inevitably exhibit some degree of bias. This may lead to imprecise responses from respondents regarding financial literacy-related questions, potentially biasing the estimation results.

To address these potential endogeneity issues, this paper further employs the instrumental variable (IV) method. First, we select instrumental variables for rational inattention: (1) financial literacy lagged by one period; (2) following Yang et al. (2023), we use the average financial literacy *FinitIV* of households within the same province, excluding household i, as an instrumental variable; (3) we use the educational attainment of the household head *Edu* as an instrumental variable, given that the head's educational level is closely related to the financial literacy that determines the household's degree of rational inattention; (4) we use the average information channel availability *InformIV* of household heads within the same province, excluding household i, as an instrumental variable, since households' financial information processing is also influenced by external information acquisition. That is, investors with greater information-seeking behavior may obtain valuable financial information externally, thereby optimizing their intrinsic information-processing capabil-

ities during financial activities.[5] Subsequently, following Tian and Zhang (2022), we construct two instrumental variables ($Treat_p \times Post_t \times FinitIV_i$ and $Treat_p \times Post_t \times Edu_i$) for the triple-difference term by interacting the aforementioned instrumental variables with the tax-deferred pilot indicator. We then perform two-stage least squares (2SLS) estimation.

The results, presented in Table 5, show that in the first-stage regression, the selected instrumental variables are significantly correlated with $Treat_p \times Post_t \times Finit_i^h$. The IV regression results further validate the impact of rational inattention on the tax-deferred pilot, demonstrating that financial literacy significantly promotes households' purchase of commercial pension insurance eligible for tax-deferred benefits. Additionally, the Kleibergen-Paap rk LM statistic yields a p-value of 0.031, rejecting the null hypothesis of unidentifiable instrumental variables. The Kleibergen-Paap rk Wald F statistic is 1622.075, exceeding the 5% maximal IV size critical value of 16.85 from the Stock-Yogo weak identification test, thereby rejecting the null hypothesis of weak instrumental variables. Furthermore, the Hansen J test yields a p-value of 0.564, accepting the null hypothesis of exogenous instrumental variables.[6]

(3) Propensity score matching

When selecting pilot regions for tax deferral, there may exist an issue of sample selection bias, wherein the initial conditions of pilot and non-pilot regions are not entirely identical, leading to differing characteristics between the treatment and control groups and subsequently altering household participation in commercial pension insurance. To address this, this paper employs the Propensity Score Matching (PSM) method, selecting household size $Size$, income $Income$, urban-rural status $Urban$, as well as the age Age and education level Edu of the household head

[5]In the CFPS questionnaire, questions related to information channels are as follows: Based on your actual situation, please rate the importance of the following information channels on a scale of 1 to 5, where 1 indicates "very unimportant" and 5 indicates "very important." (1) Importance of television: How important is television for you in acquiring information? (2) Importance of the internet: How important is the internet for you in acquiring information? (3) Importance of newspapers and periodicals: How important are newspapers and periodicals for you in acquiring information? (4) Importance of radio: How important is radio for you in acquiring information? (5) Importance of mobile text messages: How important are mobile text messages for you in acquiring information? (6) Importance of word-of-mouth: How important is word-of-mouth for you in acquiring information? Based on the responses to these questions, a continuous variable *Inform* ranging from 6 to 30 is generated. A higher number of "yes" responses indicates a richer array of information channels.

[6]To address potential inference issues arising from only 27 clusters in the instrumental variable analysis, we further employ the bootstrap method for hypothesis testing of the instrumental variables. Based on 500 bootstrap iterations, the results similarly reject the null hypotheses of unidentifiable and weak instrumental variables while accepting the null hypothesis of exogenous instrumental variables. Detailed bootstrap test results for the instrumental variables are provided in Appendix A.

Table 5: Instrumental variable regression results.

Dependent variable	First-stage regressions: $Treat_p \times Post_t \times Finit_i^h$	IV (2SLS) estimation: Y_{ipt}
L. $Treat_p \times Post_t \times Finit_i^h$	-0.272^{***}	
	(0.056)	
$Treat_p \times Post_t \times FinitIV_i$	-43.479^{*}	
	(23.070)	
$Treat_p \times Post_t \times Edu_i$	0.162^{**}	
	(0.063)	
$Treat_p \times Post_t \times InformIV_i$	-0.325^{**}	
	(0.152)	
$Treat_p \times Post_t \times Finit_i^h$		0.050^{***}
		(0.008)
Controls	Yes	Yes
Family fixed effects	Yes	Yes
Province-year fixed effects	Yes	Yes
Observations	20141	20141
Centered R^2		0.001
Kleibergen-Paap rk LM statistic' p value		0.031
Kleibergen-Paap Wald rk F statistic		1622.075
Hansen J statistic' p value		0.564

Note: 10% maximal IV size of Stock-Yogo weak ID test critical values is 19.93.

as covariates. A 1:1 caliper nearest neighbor matching approach was applied. Following PSM, a difference-in-differences-in-differences model (i.e., PSM-DDD) was constructed, and samples with non-zero weights were re-used for estimation.[7] The regression results of the PSM-DDD are presented in column (1) of Table 6, where the estimated coefficient of $Treat_p \times Post_t \times Finit_i^h$ is significantly positive. This indicates that, after accounting for sample selection bias, households with higher financial literacy are more inclined to purchase commercial pension insurance when faced with a policy shock of tax deferral, aligning with the empirical findings presented earlier.

(4) Subsample regression

Given that the maximum age limit for purchasing commercial pension insurance typically ranges from 60 to 75 years, and individuals beyond this age are ineligible to purchase, this paper excludes samples of elderly individuals over 75 years old and repeats the aforementioned regression steps. The regression results are presented in column (2) of Table 6. After adjusting the sample range,

[7]For the results of the balance test of the propensity matching scores, please refer to Appendix A.

the empirical findings of this paper remain valid.

(5) Excluding the interference of other policies

Considering that the implementation of other concurrent policies during the sample period may produce a certain overlapping effect, potentially influencing the empirical results presented earlier, we need to account for this. From the perspective of tax incentives, the increase in the personal income tax threshold was a significant policy during the 2018 pilot program for individual tax-deferred pension plans. This increase had an income effect of reducing the tax burden, thereby enhancing demand for tax-deferred commercial pension insurance products among insured individuals. To mitigate the potential impact of this policy on household participation in commercial pension insurance, the following approach is adopted: Firstly, due to the increase in the income tax threshold, household income increased accordingly. We construct an income dummy variable $Income_i^h$, which takes the value of 1 if the income of household i exceeds the median income of all sample households, indicating a relatively higher income level for household i, and 0 otherwise. Then, the interaction term between the income dummy variable and the dummy variable for the implementation year of the income tax threshold increase tax Tax_{2018} (0 before 2018, and 1 thereafter) is used as the policy dummy variable for the income tax threshold increase income $Income_i^h \times Tax_{2018}$. The regression is re-run incorporating this variable. The regression results, including the policy dummy variable for the income tax threshold increase, are shown in column (3) of Table 6. The coefficient of the triple difference term remains significantly positive, and the policy dummy variable for the income tax threshold increase is insignificant. This indicates that, even after accounting for the policy interference of the income tax threshold increase, financial literacy still significantly and positively influences the policy effect of the tax-deferred pilot program. This demonstrates the robustness of the research conclusions presented earlier.

2.4.4. Exclude alternative explanations

From the aforementioned empirical results, it can be seen that households with high financial literacy are conducive to developing their intrinsic information-processing capabilities regarding tax-deferred policies, thereby mitigating the adverse impact of rational inattention on their willingness to purchase commercial pension insurance. However, in addition to intrinsic factors such as rational inattention, other factors may also influence household participation in commercial pension insurance. Therefore, this paper further conducts exclusion tests to rule out alternative explanations.

(1) Household Wealth and Participation in Commercial Pension Insurance

The heterogeneity in the effects of the tax-deferred pilot policy may well align with another

Table 6: PSM, Subsample and Excluding the interference of other policies: Additional robustness tests.

Dependent variable: Y_{ipt}	(1) PSM-DDD	(2) Subsample	(3) Excluding the interference of other policy
$Treat_p \times Post_t \times Finit_i^h$	0.011^{***} (0.003)	0.017^{**} (0.008)	0.016^{**} (0.007)
$Income_i^h \times Tax_{2018}$			-0.001 (0.004)
Controls	Yes	Yes	Yes
Family fixed effects	Yes	Yes	Yes
Province-year fixed effects	Yes	Yes	Yes
Observations	2409	34279	35168
Adjusted R^2	0.382	0.266	0.266

explanation: households with outstanding bank loans or holdings of financial products tend to be wealthier. A substantial body of empirical research indicates that household asset levels are a critical determinant of their willingness and capacity to acquire pension insurance. The most direct evidence comes from Brown and Poterba (2005), who found that total household wealth and financial assets are effective predictors of variable annuity holdings, suggesting that substantial asset accumulation is a prerequisite for participating in such complex retirement products. Although informational interventions can influence retirement planning decisions, Duflo and Saez (2003) discovered that employees in high-paying sectorswho already possess higher incomes and greater accumulated assetsare more inclined to proactively adjust their pension arrangements. Collectively, these studies reveal a clear logical chain: greater assets directly enhance purchasing power, leading to better pension decisions, and thereby confirm that asset-rich households face fewer constraints in acquiring pension insurance.

To examine whether household wealth serves as a crucial channel through which tax incentives influence participation in commercial pension insurance, this paper measures household wealth using assets, given that the impact of income has already been accounted for in the control variables. Let $Asset$ denote the natural logarithm of a household's total assets in the current year. Households are grouped based on the median value of $Asset$: if the natural logarithm of household i's total assets exceeds the median of all sample households, $Asset_i^h$ is set to 1, indicating that household i has relatively higher assets and is wealthier; otherwise, it is set to 0. The results, as shown in Column (1) of Table 7, reveal that the estimated coefficient for $Treat_p \times Post_t \times Asset_i^h$ is statistically insignificant. This suggests that the tax-deferred pilot does not promote the purchase of commercial pension insurance among households with higher assets, thereby providing no support

for the notion that household wealth is a significant channel through which tax incentives affect participation in commercial pension insurance.

(2) Social Networks and Participation in Commercial Pension Insurance

Under the assumption of bounded rationality, households' processing of financial information is also influenced by external information acquisition. That is, through more extensive information searching, investors may obtain valuable financial information from external sources, thereby optimizing their intrinsic information-processing capabilities during financial activities and enhancing their financial awareness of tax-deferred commercial pension insurance. Households primarily acquire information through various social networks, including television, the internet, radio, newspapers, periodicals, mobile phones, and word-of-mouth from others. The more extensive a household's social networks, the more likely it is to gather useful information, which in turn affects its financial decisions (Bogan, 2008; Brown et al., 2008; Jiang et al., 2022; Wang et al., 2023; Hvide et al., 2024). However, as a traditional relationship-based society, China places significant emphasis on interpersonal connections characterized by reciprocal exchanges of favors. Households with broader social networks tend to have higher expenditures on gifts for weddings, funerals, and other ceremonial occasions (Yang et al., 2011; Ma et al., 2011; Yin et al., 2021). Generally, households with more developed social networks interact more frequently with the external environment, often incurring higher gift expenditures, which facilitates the exchange of financial information and communication about investment opportunities.

To examine whether social networks serve as a crucial channel through which tax incentives influence household participation in commercial pension insurance, this paper uses gift expenditures *Favpay* as a proxy for social networks. Let *Favpay* denote the natural logarithm of a household's gift expenditures in the current year. Households are grouped based on the median value of *Favpay*: if the natural logarithm of household i's gift expenditures exceeds the median of all sample households, $Favpay_i^h$ is set to 1, indicating that household i has relatively higher gift expenditures and a more developed social network; otherwise, it is set to 0. The results, as shown in Column (2) of Table 7, reveal that the estimated coefficient for $Treat_p \times Post_t \times Favpay_i^h$ is statistically insignificant. This suggests that the tax-deferred pilot does not promote the purchase of commercial pension insurance among households with higher gift expenditures, providing no support for the notion that social networks are a significant channel through which tax incentives affect household participation in commercial pension insurance.

Table 7: Exclude alternative explanations: Household wealth and social networks.

Dependent variable :Y_{ipt}	(1)	(2)
$Treat_p \times Post_t \times Asset_i^h$	0.007 (0.005)	
$Treat_p \times Post_t \times Favpay_i^h$		−0.001 (0.813)
Controls	Yes	Yes
Family fixed effects	Yes	Yes
Province-year fixed effects	Yes	Yes
Observations	35168	35168
Centered R^2	0.266	0.266

2.4.5. Further discussion: Possible factors driving rational inattention

Given that enhancing individual financial literacy helps reduce the degree of rational inattention and thereby promotes the policy incentive effect of the tax-deferred pilot program on households' purchase of commercial pension insurance, a subsequent question arises: What factors drive rational inattention? Existing research has provided compelling evidence that both wage growth uncertainty and tax complexity are significant determinants of investors' rational inattention. Luo (2017) introduced information frictions and information-processing constraints to explore optimal investment, consumption, and precautionary saving behaviors under the joint influence of preference robustness and rational inattention. The study found that both types of uncertainty reduce the proportion of investments in risky assets, offering an explanation for the lower stock market participation rates among low-income households. Luo et al. (2022) further delved into the specific channel of income uncertainty, revealing that in environments characterized by pervasive information deficiencies and uncertain future income streams, households optimally adopt simplified, robust consumption and saving rules. This behavior more clearly reflects rational inattention toward complex dynamic programming problems, as the costs of processing complex signals about volatile wage growth outweigh the benefits. Complementing these findings, Farhi and Gabaix (2020) conducted pioneering research by formally incorporating behavioral agents (including those subject to attention constraints) into optimal taxation theory. They discovered that when taxpayers face complex tax codes, their limited attention capacity endogenously influences their economic decisions, demonstrating that tax complexity is not merely an inconvenience but a primary factor distorting cognitive effort. Therefore, this paper further examines whether wage uncertainty and perceptions of tax complexity affect household financial literacy, delving deeper into the potential driving factors behind rational inattention. This exploration also provides richer empirical evidence

for subsequent theoretical modeling.

Before conducting regression analysis, it is first necessary to select measurement indicators for wage growth uncertainty and tax complexity. Measuring wage uncertainty is relatively straightforward. Based on the logarithmic growth rate of total after-tax wages, we calculate the standard deviation (i.e., volatility) of the growth rate for each household, resulting in *WageSD*. A higher *WageSD* indicates greater volatility and uncertainty in wage growth.

Measuring tax complexity, however, poses challenges due to difficulties in obtaining direct indicators or standardizing the metric. In this paper, we define variables based on respondents' answers in the CFPS questionnaire regarding their awareness of wages, monthly average bonuses, and monthly average pensions. The specific questions include: (1) "Including wages, bonuses, subsidies, in-kind benefits, etc., what was the total after-tax wage income of all wage earners in your household over the past 12 months?"; (2) "Excluding year-end bonuses, what was the average monthly after-tax amount of your monthly or quarterly bonuses over the past 12 months?"; and (3) "How much after-tax income do you currently receive each month in total?" If the respondent or household head can provide specific monetary amounts for these questions, the variable is assigned a value of 1; otherwise, it is assigned a value of 0. This generates a continuous variable, *Taxcomplex*, ranging from 0 to 3, which comprehensively reflects whether households understand the calculation of various complex tax rates, including wages, monthly average bonuses, and monthly average pensions. A higher *Taxcomplex* indicates relatively lower tax rate complexity.

Next, we use whether household i belongs to a high-financial-literacy group $Finit_i^h$ as the independent variable and gradually incorporate wage growth uncertainty (*WageSD*) and tax complexity (*Taxcomplex*) as dependent variables. Simultaneously, we include income (*Income*), other housing rental expenses (*Housrent*), the dependency ratio (*Depration*), and gift expenditures (*Favpay*) as control variables in a conditional Logit model regression. The results, as shown in Table 8, reveal the following: The estimated coefficient for *WageSD* is significantly negative, indicating that as *WageSD* increases and wage uncertainty rises, the likelihood of a household belonging to the high-financial-literacy group decreases, and its degree of rational inattention increases. Conversely, the estimated coefficient for *Taxcomplex* is significantly positive, suggesting that as *Taxcomplex* decreases and tax rate complexity rises, the likelihood of a household belonging to the high-financial-literacy group decreases, and its degree of rational inattention increases. Furthermore, the chi-square statistic from the joint significance test of *WageSD* and *Taxcomplex* is $\chi(2) = 10.57$, significant at the $p = 0.05$ level.

Combining these empirical results with earlier findings, we outline a clear empirical fact: Rational inattention serves as a critical pathway through which tax incentives influence household

participation in commercial pension insurance, while wage growth uncertainty and tax complexity jointly affect the degree of rational inattention.[8]

Table 8: Possible factors driving rational inattention: Wage growth uncertainty and tax complexity.

Dependent variable : $Finit_i^h$	(1)	(2)
WageSD	-0.079^{**}	-0.077^{**}
	(0.039)	(0.039)
Taxcomplex		0.098^{*}
		(0.051)
Controls	Yes	Yes
Family fixed effects	Yes	Yes
Province-year fixed effects	Yes	Yes
Observations	7534	7534
Centered R^2	0.070	0.071

Note: Given that conditional likelihood estimation requires complete nesting of groups within clusters, the conditional Logit model cannot accommodate non-nested fixed effects structures. As a compromise, this paper controls for province-level fixed effects while reporting robust clustered standard errors at the provincial level.

3. Theoretical mechanism

In the empirical section of this paper, we demonstrate that rational inattention influences the purchase behavior of commercial pension insurance. In the modeling section, we introduce wage growth uncertainty and tax computation complexity to further investigate the mechanism underlying rational inattention.

Specifically, we construct a pension investment portfolio model incorporating tax deferral policies, and introduce tax rates and wage growth uncertainty to characterize consumers' rational inattention traits. We then derive the optimal investment-consumption (or contribution) strategy. By analyzing the optimal strategy, we explore the working mechanism of rational inattention and elucidate its economic implications. Mathematically, we consider a complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$, where $\mathcal{F}_t$ represents the information set available to consumers up to time t. For simplicity, we assume no transaction costs.

3.1. Rational inattention

Within the framework of rational inattention, this paper establishes a third-pillar individual commercial pension insurance account and incorporates a tax deferral incentive policy to discuss

[8]The factual evidence presented in this subsection provides a foundational reference for our subsequent theoretical analysis in both Section 3 and Section 5.

consumers' optimal decision-making problem.[9] For third-pillar individual commercial pension insurance in China, a tax incentive policy is adopted. A typical example is the Exempting-Exempting-Taxing (EET) tax deferral policy, where contributions and investment returns are not taxed, but withdrawals are subject to taxation. This policy provides a tax incentive for policyholders, thereby enhancing individuals' enthusiasm for purchasing commercial pension insurance. On the other hand, China's personal income tax structure is progressive, with seven brackets, and the tax rate varies depending on the income level. Furthermore, the calculation of personal income tax involves various deductions, such as the basic deduction, special deductions, and additional special deductions (specific amounts and eligibility criteria can be referred to the latest regulations published by the State Taxation Administration[10]).

Based on this, for the representative consumer participating in a third-pillar commercial pension plan, this paper assumes uncertainty in future wage growth rates. Consumers exhibit characteristics of rational inattention when confronted with a progressive income tax structure and uncertainty regarding wage growth. Typically, the financial market comprises a risk-free bond and a stock. The price process of the risk-free bond is described by the following ordinary differential equation:

$$dS_0(t) \quad = \quad rS_0(t)dt, \quad S_0(0) = s_0, \tag{5}$$

where r represents a constant denoting the risk-free interest rate. The stock price process is characterized by the following geometric Brownian motion:

$$dS(t) \quad = \quad S(t)(\mu_S dt + \sigma_S dW_S(t)), \quad S(0) = s, \tag{6}$$

where $\mu_S > r$ denotes the drift rate of the stock, and σ_S represents the volatility of the stock. Both μ_S and σ_S are constants. $W_S(t)$ is a one-dimensional standard Brownian motion, embodying the systematic risk in the financial market. Let $\theta = \frac{\mu_S - r}{\sigma_S}$ signify the market price of risk.

Assuming that the consumer begins working at time 0 and retires at time T, during which period she receives a continuous stream of wage income. We posit that the wage process is characterized

[9]The model framework presented in this paper can be further extended to a dual-account model for individual pension investments. Specifically, the dual accounts refer to the first-pillar basic pension account and the second/third-pillar (enterprise annuity and individual pension) accounts. Under China's current pension system, the basic pension has always been tax-exempt, with contribution rates mandated by the state and primarily invested in safe assets such as government bonds (risk-free assets). Therefore, consumers have no control over the basic pension account and only need to consider the statutory proportion b of their wages deducted before taxes for basic pension contributions; hence, consumers primarily focus on optimizing their third-pillar individual pension accounts. In particular, the general setup and computational results for the dual-account model are provided in Appendix D.

[10]https://www.chinatax.gov.cn/.

by the following geometric Brownian motion:

$$dY(t) \quad = \quad (\mu_Y(t) - \kappa Y(t))dt + \sigma_Y dW_Y(t), \quad Y(0) = y_0 > 0, \tag{7}$$

where $\mu_Y(t)$ represents the expected growth rate of wages, indicating the time-varying increase in income. σ_Y denotes the volatility of wages, while κ and σ_Y are constants that characterize the speed of mean reversion and the instantaneous volatility of wage income stochastic components, respectively. Here, the standard Brownian motion $W_Y(t)$ is used to capture wage risk. For simplicity, it is assumed to be independent of $W_S(t)$.

Herein, the specification and estimation of $\mu_Y(t)$ are crucial. Following Wang (2009), we assume that the consumer is uncertain about the future expected growth rate of their wages. Therefore, before selecting the optimal investment portfolio and consumption strategy, the consumer must first estimate their income growth rate. However, departing from Wang (2009), this paper assumes that consumers have the opportunity to actively acquire information from the financial market and, by gathering such information, gain a better understanding of the true wage growth rate, i.e., obtaining accurate information regarding wage growth. Drawing on Andrei and Hasler (2019), it is assumed that the information (signal) obtained by the consumer is

$$dI(t) \quad = \quad \mu_Y(t)dt + \frac{1}{\phi(t)}dW_S(t), \tag{8}$$

where $W_S(t)$ indicates that the consumer pays attention to information or signals pertaining to the financial market, through which they learn and estimate the growth rate of their wage income. The condition $(\phi(t))^2 > 0$ signifies that the consumer's attention to information or signals is limited, with $\phi(t)$ representing the consumer's response to financial market uncertainty. Following Andrei and Hasler (2019), when the consumer gathers this information or signals, they incur a certain cost. For simplicity, the information cost function is modeled in a simple quadratic form, denoted as $\chi(\phi) \equiv \chi(\phi(t)) = \frac{\omega}{2}(\phi(t))^2$.

Remark 3.1 (Real costs vs. entropy constraint)**.** While our model explicitly incorporates real information gathering and processing costs into the household's budget constrainta different technical device from the shadow-price-based entropy constraint in Sims (2003)-the two formulations capture the same core economic intuition of rational inattention. As discussed in Veldkamp (2009), Myatt and Wallace (2012) and Caplin and Dean (2015), models with explicit real costs and those with an information-capacity constraint can be viewed as complementary: both imply that agents optimally limit their attention to noisy signals, leading to gradual and incomplete absorption of new information. In our setting, the real cost parameter implicitly disciplines the shadow cost of reducing uncertainty, thereby generating dynamics in beliefs and actions that are qualitatively consistent with Sims' foundational framework. Hence, although the micro-foundation differs,

the resulting predictions for slow information diffusion and state-dependent attention are closely aligned, allowing our specification to serve as a tractable, wealth-based representation of the same class of informational frictions.

Similar to Emms (2012), under the structure of China's individual income tax system, let the initial income tax rate be denoted as τ and the contribution rate as $\xi(t)$. When tax incentives such as tax deferral policies are implemented, contributions to the third-pillar commercial pension insurance are made before tax deductions. Thus, her contribution to the commercial pension insurance at time t is $Cont(t) \equiv \xi(t)Y(t)$, and the income subject to personal income tax is the income level after deducting both the basic pension and the commercial pension insurance, i.e., $(1-\xi(t))Y(t)$. When there are information-processing costs, it follows that her consumption level at time t is $C(t) = (1-\tau)(1-\xi(t))Y(t) - \chi(\phi)$. Based on the quantitative relationship, it can be further derived that the contribution is $Cont(t) = Y(t) - \frac{C(t)}{1-\tau} - \frac{\chi(\phi)}{1-\tau}$.[11]

In addition to the tax deferral policy incentive, another fundamental element of the third-pillar commercial pension insurance is the individual's autonomy in choosing pension allocation strategies. Specifically, individuals can independently select various pension investment products based on their risk preferences, which facilitates the transition from saving for retirement to investing for retirement in China. Therefore, for individual commercial pensions, enhancing the returns of the investment portfolio is of great significance for pension wealth accumulation. At time t, the consumer's pension wealth is denoted as $X(t)$, and the wealth invested in stocks is denoted as $\pi(t)$. Consequently, the wealth invested in risk-free bonds is $X(t) - \pi(t)$. Thus, $X(t)$ follows

$$\begin{aligned} dX(t) &= \pi(t) \cdot \frac{dS(t)}{S(t)} + (X(t) - \pi(t)) \cdot \frac{dS_0(t)}{S_0(t)} + Cont(t)dt \\ &= [rX(t)dt + \pi(t)\sigma_S(\theta dt + dW_S(t))] + \left(Y(t) - \frac{C(t)}{1-\tau} - \frac{\chi(\phi)}{1-\tau}\right) dt. \end{aligned} \tag{9}$$

Based on the aforementioned analysis, we next incorporate the characteristic of consumer rational inattention when calculating the impact of the tax deferral policy on disposable income and tax rate structures. Specifically, when confronted with the actual tax system, consumers typically

[11] This analysis can also be extended to a dual-account model of individual pension investments that includes a basic pension account. Assume the contribution rate for the basic pension is b (deducted before tax). In this case, the contribution rate for the third-pillar commercial pension insurance applies to the wage level after deducting the basic pension, i.e., $(1-b)Y(t)$. Thus, her contribution to the commercial pension insurance at time t is $Cont(t) \equiv \xi(t)(1-b)Y(t)$, and the income subject to personal income tax is the income level after deducting both the basic pension and the commercial pension insurance, i.e., $(1-\xi(t))(1-b)Y(t)$. When there are information search costs, it follows that her consumption level at time t is $C(t) = (1-\tau)(1-\xi(t))(1-b)Y(t) - \chi(\phi)$. Based on the quantitative relationship, it can be further derived that the contribution is $Cont(t) = (1-b)Y(t) - \frac{C(t)}{1-\tau} - \frac{\chi(\phi)}{1-\tau}$.

cannot accurately calculate the value of their disposable income. The main reasons are as follows: First, the complexity of the progressive tax structure;[12] Second, the uncertainty of future wages and the cost of information search.[13] In addition, these factors, combined with various tax incentives such as special deductions and additional special deductions in China, collectively lead consumers to be unable to accurately calculate the specific value of their disposable income after tax deferral. This reflects consumers' "subjective pursuit of rationality, but objectively, they can only limitedly and imprecisely calculate the aforementioned commercial pension insurance contribution $Cont(t) \equiv Y(t) - \frac{C(t)}{1-\tau} - \frac{\chi(\phi)}{1-\tau}$".

Furthermore, within the model framework, whether consumers estimate their disposable income by searching for information or calculate the specific value of how tax deferral policies and tax rate structures affect their disposable income, consumers with higher financial literacy are more likely to achieve these tasks compared to those with lower financial literacy.

Remark 3.2 (Progressive tax rates)**.** When there is only one tax rate, the calculation of the commercial pension insurance contribution term $Cont(t) \equiv Y(t) - \frac{C(t)}{1-\tau} - \frac{\chi(\phi)}{1-\tau}$ is not complex. Here, we provide an example with two tax rates and one wage threshold to illustrate that calculating the specific value of how tax deferral policies and progressive tax rate structures affect disposable income is relatively complex. Therefore, we consider the characteristic of investor rational inattention. Suppose the tax rate is $\tau_t = \{\tau_{L,t}, \tau_{H,t}\}$, where $\tau_{L,t}$ and $\tau_{H,t}$ represent different tax rates faced when wages are low and high, respectively, and the threshold for applying different tax rates to wages is Γ. First, at time t, the consumer's wage is $Y(t)$. With the tax deferral policy, the contribution to the third-pillar commercial pension insurance account is $Cont(t) \equiv \xi(t)Y(t)$, and the remaining portion of the wage is consumed as $C(t)$, but it needs to be taxed first. Thus, the specific level of consumption is:

$$C(t) = \begin{cases} (1-\xi)(1-\tau_{L,t})Y(t) - \chi(\phi), & \text{if} \quad (1-\xi)Y(t) < \Gamma, \\ (1-\tau_{L,t})\Gamma + [(1-\xi)Y(t) - \Gamma](1-\tau_{H,t}) - \chi(\phi), & \text{if} \quad (1-\xi)Y(t) \geq \Gamma. \end{cases} \tag{10}$$

Drawing from Herzog (2019), consider an example concerning taxation. Suppose the price of a commodity is denoted by p, and the tax rate on this commodity is t. Consequently, the full price of the commodity is $q = p + t$. However, consumers only partially attend to the tax, thus the perceived

[12] See the footnote below.

[13] Consumers, unable to accurately discern their wage growth rate, can only estimate it based on market signals. However, in the estimation process, on the one hand, their information processing capacity or attention is limited, manifested as $(\phi(t))^2 > 0$; on the other hand, searching for information incurs a cost $\chi(\phi)$, further increasing the difficulty of estimating their disposable income.

tax is $t_s = mt$, and the perceived price becomes $q_s = p + mt$, where $m \in [0, 1]$ measures the level of attention. Intuitively, under bounded rationality, the perceived full price of the commodity diminishes. Such psychological misperceptions and misjudgments exert a significant influence on optimal taxation and tax evasion theories. Building upon this, we posit that consumers apply a distorted tax rate, τ', when calculating their disposable income in the context of tax deferral. Therefore, it is equally justifiable to incorporate consumer bounded rationality characteristics when assessing the impact of tax deferral policies on disposable income and tax rate structures.

In summary, in the subsequent sections, we transform a progressive tax rate structure into a flat tax rate by calculating the stable effective tax burden, assuming a uniform personal income tax rate applies to all wage levels. Consequently, an intuitive approach to setting this tax rate, τ', is to base it on the actual wage tax rate faced by consumers in their daily lives. This aligns with the concepts of bounded rationality and rational inattention in behavioral finance. It is worth noting that when adopting a progressive tax structure, as indicated in Footnote 1, the calculation of multiple tax rate parameters becomes more complex. Ordinary consumers either neglect these calculations and continue to use a single wage tax rate or invest time and effort, or utilize specialized tax rate calculators available online. Nevertheless, consumers generally understand that tax deferral appropriately reduces the original wage tax rate, τ. Therefore, the relationship between τ' and the original tax rate τ satisfies $\tau \geq \tau'$. That is, for tax-advantaged third-pillar commercial pension insurance, consumers face a lower tax rate, reflecting the characteristic of tax incentives. Hence, in the subsequent modeling of this paper, the tax rate adopted by consumers is set as τ'.

Building upon the optimal portfolio models of Merton (1971) and Luo (2017), we construct a continuous-time investment-consumption model where consumers aim to maximize the expected utility of lifetime consumption.

$$U(C) = \mathbb{E}\left[\int_0^\infty e^{-\beta s} u(C(s))ds\right], \tag{11}$$

Where $\beta > 0$ represents the consumer's subjective discount rate, and the instantaneous utility function $u(c)$ satisfies $u(c) = -\frac{e^{-\gamma c}}{\gamma}$, with $\gamma > 0$ denoting the coefficient of absolute risk aversion.

3.2. Filtering

To solve the optimal problem involving the unobservable $\mu_Y(t)$, we first employ the Kalman-Bucy filtering method to infer the estimated parameter $\hat{\mu_Y}(t)$. Subsequently, we use $\hat{\mu_Y}(t)$ instead of $\mu_Y(t)$ to solve the optimal investment-consumption problem.

Based on the aforementioned assumptions, for the consumer, the information set $\mathcal{F}$ available at time t only includes the historical records of realized wage income $Y(s) : \ 0 \leq s \leq t$ and the

acquired information or signals $I(s):\ 0 \leq s \leq t$. The consumer's estimated expected wage growth rate is denoted by $\hat{\mu_Y}(t) \equiv \mathbb{E}[\mu_Y(t)|\mathcal{F}_t]$.

Before estimating $\hat{\mu_Y}(t)$, we need to describe some statistical distributions or characteristics of $\mu_Y(t)$. Here, we consider a relatively simple scenario where the wage growth rate $\mu_Y(t)$ follows a two-state Markov process. Specifically, $\mu_Y(t) \in \{\mu_{LY}, \mu_{HY}\}$, where $\mu_{LY} < \mu_{HY}$ represent the consumer having a lower or higher wage growth rate, respectively. For simplicity, let $\delta = \mu_{HY} - \mu_{LY}$. For this Markov process, let λ_i denote the transition probability from state i to the other state, where $i = L, H$.

Let $p(t)$ represent the probability that the consumer believes the income growth rate is high given the information set $\mathcal{F}$, which is expressed as follows:

$$p(t) \quad = \quad \mathbb{P}(\mu_Y(t) = \mu_{HY}|\mathcal{F}), \tag{12}$$

a straightforward calculation yields $\hat{\mu_Y}(t) = \mu_{LY} + (\mu_{HY} - \mu_{LY})p(t) = \mu_{LY} + \delta p(t)$. Furthermore, based on Liptser and Shiryaev (2001), the stochastic processes governing the consumer's beliefs and related state variables can be derived as follows:

$$dp(t) \quad = \quad [\lambda_L - (\lambda_L + \lambda_H)p(t)]dt + \delta p(t)(1 - p(t)) \left(\frac{1}{\sigma_Y} d\hat{W}_Y(t) + \phi(t) d\hat{W}_S(t) \right). \tag{13}$$

Accordingly,

$$\frac{dS(t)}{S(t)} \quad = \quad \mu_S dt + \sigma_S d\hat{W}_S(t), \tag{14}$$

$$dY(t) \quad = \quad (\mu_{LY} + \delta p(t) - \kappa Y(t))dt + \sigma_Y d\hat{W}_Y(t), \tag{15}$$

where the definitions and derivation processes of $\hat{W}_S(t)$ and $\hat{W}_Y(t)$are provided in Appendix B.

3.3. Optimal strategy

In the solution process of this paper, we employ the method of stochastic dynamic programming. Specifically, the optimization problem faced by the consumer is formulated as:

$$\max_{C(t),\pi(t),\phi(t)} \mathbb{E}\left[\int_t^\infty e^{-\beta s} u(C(s)) ds | \mathcal{F}_t \right], \tag{16}$$

where the state variables $X(t)$, $Y(t)$ and $p(t)$ (for convenience, hereinafter referred to as x, y and p respectively) satisfy equations

$$dX(t) = \left[rX(t)dt + \pi(t)\sigma_S(\theta dt + d\hat{W}_S(t)) \right] + \left(Y(t) - \frac{C(t)}{1-\tau'} - \frac{\chi(\phi)}{1-\tau'} \right) dt,$$

(15) and (13) respectively. It is worth noting that the consumer's decisions encompass consumption $C(t)$, investment $\pi(t)$, and the response to financial market uncertainty $\phi(t)$ (or, following Andrei and Hasler (2019), the allocation of attention to market information $(\phi(t))^2$.

To analytically and focusedly examine the impact of rational inattention and information acquisition on the investment-consumption decisions of consumers facing uncertainty in wage growth, we consider a simplified scenario. The consumer faces an unknown but constant wage growth rate, i.e., the wage growth μ_Y is constant but unknown, hence $\lambda_L = \lambda_H = 0$.[14] Consequently, the consumer's beliefs evolve according to

$$dp(t) \quad = \quad \delta p(t)(1-p(t)) \left(\frac{1}{\sigma_Y} d\hat{W}_Y(t) + \phi(t) d\hat{W}_S(t) \right). \tag{17}$$

The above equation indicates that the consumer's estimate of the wage growth rate only exhibits a stochastic term, which is reasonable given that the wage growth rate is actually constant in this context. Consequently, the evolution of the agent's beliefs must be unpredictable, rendering the process a martingale. Based on this formula, we observe that the instantaneous volatility of belief updating is symmetric with respect to $p(t)$ and $1-p(t)$. Furthermore, the updating of consumer beliefs hinges on the difference $\delta = \mu_{HY} - \mu_{LY}$ between the two possible values of the unobserved growth rate. As δ increases, the volatility of consumer belief updating becomes greater. On the other hand, a higher income volatility σ_Y implies a lower volatility in belief updating. Intuitively, if the income process exhibits less volatility, the realized values of income provide more information about the unknown wage growth (see Wang (2009)).

Following Wang (2009), we employ the dynamic programming approach to derive the Hamilton-Jacobi-Bellman (HJB) equation and the optimal investment-consumption strategy. The detailed mathematical derivation steps are provided in Appendix C, and here we present the final results.

Proposition 3.3 (Analytical solutions)**.** *If the wage growth rate is constant but unknown to the consumer, the consumer estimates and learns about it by searching for information and incurring costs. In this case, the consumer's optimal investment, consumption, and response to financial market uncertainty are given by:*

$$C^*(t) = r(1-\tau') \left(x + h(y; \mu_{LY}) + f(p) \right), \tag{18}$$

$$\pi^*(t) = \frac{\theta}{(1-\tau')\gamma r \sigma_S} - \frac{\theta \delta^2 p^2 (1-p)^2}{\sigma_S [\delta^2 p^2 (1-p)^2 f''(p) - \frac{\omega}{1-\tau'}]} (f'(p))^2, \tag{19}$$

$$\phi^*(t) = \frac{\theta \delta p (1-p) f'(p)}{\delta^2 p^2 (1-p)^2 f''(p) - \frac{\omega}{1-\tau'}}, \tag{20}$$

[14] For related assumptions on wage growth rates and further discussion, see Wang (2009).

where the function $h(y;\mu_{LY})$ represents the certainty equivalent human capital level induced by wages (based on y), determined by the following equation:

$$h(y;\mu_{LY}) = \frac{1}{r+\kappa}\left(y + \frac{\mu_{LY}}{r} - \frac{\gamma(1-\tau')\sigma_Y^2}{2(r+\kappa)}\right) + \frac{\beta - r}{(1-\tau')\gamma r^2} + \frac{\theta^2}{2(1-\tau')\gamma r^2}, \tag{21}$$

the function $f(p):\ 0 \le p \le 1$ represents the certainty equivalent wealth level induced by learning (based on $p(t)$), which satisfies the following nonlinear ordinary differential equation:

$$\begin{aligned} rf(p) = \frac{\delta p}{r+\kappa} - (1-\tau')\frac{\gamma r\delta p(1-p)}{r+\kappa}f'(p) - (1-\tau')\frac{\gamma r\delta^2 p^2(1-p)^2}{2\sigma_Y^2}(f'(p))^2 \\ -\frac{\theta^2\delta^2 p^2(1-p)^2}{2[\delta^2 p^2(1-p)^2 f''(p) - \frac{\omega}{1-\tau'}]}(f'(p))^2 + \frac{\delta^2 p^2(1-p)^2}{2\sigma_Y^2}f''(p), \end{aligned} \tag{22}$$

with boundary conditions satisfying

$$f(0) = 0,\ \ f(1) = \frac{\delta}{r(r+\kappa)}. \tag{23}$$

Under these conditions, the consumer's optimal value function is given by

$$J(x,y,p) = -\frac{e^{-\gamma r(1-\tau')(x+h(y;\mu_{LY})+f(p))}}{\gamma r}. \tag{24}$$

The consumer's optimal allocation of attention to financial market information, $(\phi^*(t))^2$, is given by

$$(\phi^*(t))^2 = \left(\frac{\theta\delta p(1-p)f'(p)}{\delta^2 p^2(1-p)^2 f''(p) - \frac{\omega}{1-\tau'}}\right)^2. \tag{25}$$

Based on equation (18), it can be observed that when consumers face wage uncertainty and rational inattention, their optimal consumption strategy is influenced by three factors: (1) the wealth accumulation x in the third-pillar pension account itself; (2) the certainty equivalent human capital level $h(y;\mu_{LY})$, which reflects the future contributions to the third-pillar pension account discounted under the influence of wages and contributions; and (3) the certainty equivalent wealth level $f(p)$, which is affected by incomplete information, information search costs, and the updating of beliefs about the wage growth rate μ_Y.

Through equations (18), (21), (22), and (25), we find that both the optimal consumption and the optimal level of attention allocation to market information by consumers decrease due to the presence of taxes. Specifically, the optimal consumption level is directly affected by the tax adjustment factor $1-\tau'$ acting on the term $x + h(y;\mu_{LY}) + f(p)$. Furthermore, based on the definitions of $h(y;\mu_{LY})$ and $f(p)$, it can be readily calculated (through first-order conditions or numerical verification) that both $h(y;\mu_{LY})$ and $f(p)$ increase when τ' increases. However, the decreasing effect of $1-\tau'$ is more pronounced, leading to a reduction in the optimal consumption

level $C^*(t)$ (as also verified in the numerical section). The optimal level of attention allocation to market information by consumers is not directly affected by the tax adjustment factor $1-\tau'$ in the same way as the consumption adjustment factor. Therefore, when τ' increases, although the term $-\frac{\omega}{1-\tau'}$ decreases, causing the fraction in (25) to increase, the combined effect of τ' on $f'(p)$ and $f''(p)$ results in a similar reduction in the optimal attention allocation $(\phi^*(t))^2$.

Luo (2017) demonstrated that the disposable income of consumers decreases when personal income tax is imposed. With the continuous rise of tax rates, consumers tend to increase their investment proportion in risky assets and reduce their level of precautionary savings for the future. After the introduction of personal income tax, a significant increase occurs in consumers' proportion of risky asset investments. However, this investment proportion decreases when a tax deferral policy is implemented. The reason is that personal income tax reduces consumers' disposable income while increasing their liquid wealth. As previously analyzed, consumers will confront greater investment risks while increasing their investment proportion. When a tax deferral policy for individual pensions is introduced, the attractiveness of commercial pension insurance plans increases, and the discounted value of future contributions rises because of the tax deferral policy, which leads to a decrease in the investment proportion and an improvement in the risk situation faced by consumers. In this paper, we further validated these conclusions.

Specifically, the adverse effects of personal income tax are mitigated when a tax deferral policy for individual pensions is established. The tax deferral policy helps increase consumers' disposable income (the portion of disposable income used to contribute to third-pillar individual pension accounts is not subject to tax, as detailed in Section 3.1). With an increase in disposable income, consumers can choose to increase their consumption or purchase more commercial pension insurance. Nevertheless, the specific decision is up to the level of wage growth and tax rates. When a tax deferral policy is implemented (or the policy is more generous or wages exhibit high growth rates), the certainty equivalent human capital level $h(y;\mu_{LY})$ increases, which increases optimal consumption and contribution levels. On the other hand, the tax adjustment factor $1-\tau'$ approaching 1 indicates that consumers enjoy more favorable tax policies, which also results in higher optimal consumption and contribution levels, and thereby enhances their welfare. Based on the above analysis, we summarized the main conclusions in the following corollary.

Corollary 3.4 (Tax effects)**.** *(1) When consumers need to pay income tax, both their optimal consumption and optimal attention allocation levels decline.*

(2) In the transition from a scenario without tax incentives to one with tax preferential policies, the disposable income of consumers increases. Nonetheless, whether consumers increase their

purchase of commercial pension insurance depends on wage growth levels and tax rates.

This paper provides a deeper mechanistic analysis of the implementation effects of tax deferral policies by incorporating the concept of rational inattention. Specifically, based on equations (19) and (20), we obtain $\pi^*(t) = \frac{\theta}{(1-\tau')\gamma r\sigma_S} - \phi^*(t) * \frac{\delta p(1-p)f'(p)}{\sigma_S}$. The first term in this equation represents the traditional optimal investment strategy in the absence of information gathering and rational inattention; the second term reflects the additional investment demand arising from rational inattentive consumers acquiring information from the financial market, which depends on their attention to financial market signals. If consumers' attention to information is positive (i.e., $\phi^*(t) > 0$), the investment demand induced by searching for and acquiring this signal or information, $-\phi^*(t) * \frac{\delta p(1-p)f'(p)}{\sigma_S}$, is negative. Consequently, consumers allocate less wealth to the market portfolio. Furthermore, consumers' rational inattention influences the conversion of a progressive tax rate structure into an effective proportional tax rate τ'. As consumers' level of bounded rationality decreases, their perception of the tax rate becomes more accurate, enabling them to make more rational financial decisions. For instance, as τ' gradually approaches the true τ, consumers' consumption decreases, leading to higher contributions and promoting the purchase of commercial pension insurance. Thus, rational inattention helps to enhance the effectiveness of tax deferral policies on consumers.

According to Wang et al. (2022), consumers' active information-seeking and learning behavior enhance the learning-induced certainty equivalent wealth level $f(p)$, thereby increasing their consumption level accordingly. In situations of high uncertainty regarding income growth rates, agents increase their search for and attention to financial market information, leading to additional investment demand for risky assets. Based on the definition of the belief p, when p is close to 0.5, consumers face the greatest uncertainty about future wage growth rates. At this point, the optimal allocation of attention to financial market information $(\phi(t))^2$ is maximized. As p gradually decreases or increases from 0.5, the uncertainty about wage growth rates diminishes, and the optimal allocation of attention to financial market information gradually declines (as further verified by subsequent sensitivity analysis). This results in a reduction in third-pillar commercial pension contributions $Cont(t)$, indicating that when bounded rationality and wage uncertainty are present, consumers allocate more effort to learning about wage growth rate parameters. Conversely, if consumers' level of bounded rationality decreases, their perception of wage growth rates becomes more accurate. Further, based on the quantitative relationship between pension contributions, consumption, and information search, namely $Cont(t) = Y(t) - \frac{C(t)}{1-\tau'} - \frac{\chi(\phi)}{1-\tau'}$, it follows that consumers' information search costs will decrease, leading them to contribute more to third-pillar commercial

pension plans and thereby promoting the accumulation of pension wealth. Based on the above analysis, we summarize the main conclusions in the following proposition.

Corollary 3.5 (Wage uncertainty and rationality effects)**.** *(1) With the increase in consumers' beliefs about the uncertainty of wage growth rates, their optimal allocation of attention to financial market information grows accordingly.*

(2) When consumers' level of bounded rationality decreases and their perception of wage growth rates becomes more accurate, the cost of information search decreases, which facilitates consumers' purchase behavior of commercial pension insurance.

4. Numerical example and economic analysis

In this section, we present specific numerical examples to illustrate the mechanisms through which rational inattention influences the policy effectiveness of tax deferral on consumer decision-making behavior.

4.1. Calibration

In this paper, most parameters in the financial models are calibrated to real data, and remaining parameters are adopted from widely recognized values in numerous prior works.

We utilize real-world data from China to calibrate or estimate the model parameters: the risk-free interest rate r, the average growth rate μ_S and volatility σ_S of stock prices, as well as the average growth rate μ_Y and volatility σ_Y of wages.

(1) For the risk-free interest rate r, we calibrate it using China's interbank 7-day repo, with a sample period spanning from January 1, 2005, to December 31, 2024. The data frequency is daily (trading days), sourced from the Wind Financial Terminal database. Based on the historical average of the China's interbank 7-day repo, the calibrated value is $\hat{r} = 0.028$. As a result, we have the subjective discount factor $\beta = 1 - 1/(1 + r) = 0.027$.

(2) For the parameters of the stock price model, we utilize the representative CSI 300 Index of the Chinese stock market for estimation. The sample period spans from January 1, 2005, to December 31, 2024, with a daily (trading day) data frequency sourced from the Wind Financial Terminal database. Specifically, the estimation steps are as follows: First, we calculate the logarithmic returns of the CSI 300 Index, which conform to the assumption of the Geometric Brownian Motion (GBM) that logarithmic returns follow a normal distribution, as illustrated in Figure 9 (in Appendix A3). Subsequently, based on the stock price model equation and assuming 252 trading days in a year (i.e., a time interval of $dt = 1/252$), we estimate the parameters as $\hat{\mu}_S = 0.105$ and $\hat{\sigma}_S = 0.257$. As a result, we have the market price of risk $\theta = \frac{\mu_S - r}{\sigma_S} = 0.300$.

(3) For the parameters of the wage process, we conduct estimations using micro-survey data on Chinese households. Consistent with the preceding empirical analysis section, the data are sourced from the China Family Panel Studies (CFPS) conducted by the Institute of Social Science Survey (ISSS) at Peking University. We utilize households surveyed in the CFPS databases from the years 2014, 2016, 2018, 2020, and 2022 as our research sample. Similarly, we perform analogous data processing steps, including merging individual and household questionnaires, removing duplicate observations, deleting samples with severe missing values, excluding samples that only exist during the period of, prior to, or following the implementation of the tax-deferred policy, and applying a 1% winsorization to all continuous variables at both tails. Specifically, the parameter estimation process for the wage process is as follows.

First, we construct the main variables. We take the natural logarithm of the household's total after-tax wage income (including wages, bonuses, subsidies, in-kind benefits, etc.) to obtain the wage variable ($Wage$). The control variables include household size ($Size$), assets ($Asset$), income ($Income$), urban-rural status ($Urban$), other housing rentals ($Housrent$), dependency ratio ($Depration$), as well as the age of the household head (Age), the square of the household head's age ($Age2$) (denoted as $Age^2/100$), gender ($Male$), marital status ($Marriage$), and educational attainment (Edu). The remaining variables are consistent with those used in the preceding empirical analysis and will not be elaborated on further here.

Next, we employ a fixed-effects model to estimate the average growth rate of wages, with the model specified as follows:

$$Wage_{it} = \alpha_0^w + \beta_1^w Year_{2016} + \beta_2^w Year_{2018} + \beta_3^w Year_{2020} + \beta_4^w Year_{2022} + \mu_i^w + \alpha_i^w X_{it} + \varepsilon_{it}.$$

Panel A in Table 9 reports the fixed-effects regression results for the average growth rate of wages, where columns (1) and (2) present the regression outcomes without and with control variables, respectively. Given that the inclusion of Age, $Age2$, and other control variables allows us to account for life-cycle effects (i.e., the inverted U-shaped curve of wages rising and then falling with age) as well as wage variations attributable to other confounding factors at both the household and household head levels, we utilize the estimated coefficient of $Year_{2022}$ from the model with control variables to calculate the average growth rate of wages. The estimated coefficient of $Year_{2022}$ is 0.329, statistically significant at the 1% level, representing the cumulative wage growth from Period 1 (2014) to Period 5 (2022). Based on this, we calculate the compound annual growth rate of wages to be 0.042, i.e., $\hat{\mu}_Y = 0.042$.

Thirdly, we utilize the residuals from the aforementioned fixed-effects regression to capture unobserved wage fluctuations. By analyzing the variance-covariance structure of these residuals

through the Generalized Method of Moments (GMM), we decompose the fluctuations into persistent and transitory components, thereby estimating the total volatility of wages. Assuming that the residuals, denoted as ε_{it}, are composed of transitory and persistent shocks, i.e., $\varepsilon_{it} = v_{it} + w_{it}$, where v_{it} represents the persistent shock, following an AR(1) process, i.e., $v_{it} = \rho_v v_{it-1} + \varepsilon_{it}^v$, with $\varepsilon_{it}^v \sim i.i.d.(0, \sigma_v^2)$; and w_{it} represents the transitory shock, following a white noise process, i.e., $w_{it} \sim i.i.d.(0, \sigma_w^2)$. Based on this, we derive theoretical moments, namely the variance and k-th order covariance of ε_{it}, and select sample moments to estimate the parameters. Before conducting GMM estimation, it is necessary to first determine ρ_v and σ_w^2. This is because the multi-period survey data from CFPS contain numerous missing values, rendering the transitory variance sensitive to measurement errors. In this paper, we employ direct moment estimation to calculate the variance and covariance, taking the median of all possible autocorrelation coefficients as the estimate for ρ_v, denoted as $\hat{\rho}_v = -0.516$, satisfying $|\rho_v| < 1$. We then use eigenvalue decomposition of the variance-covariance matrix to estimate the variance components, with the largest eigenvalue corresponding to the persistent variance and the transitory variance equal to the total variance minus the persistent variance, i.e., $\hat{\sigma}_w^2 = 0.126$. Based on the determined ρ_v and σ_w^2, we select five sample moments (variances) for estimating σ_v^2. The estimation results $\hat{\sigma}_v^2 = 0.062$ are shown in Panel B of Table 9. Ultimately, the overall wage volatility can be expressed as $\sigma_Y^2 = \sigma_v^2/(1-\rho_v)^2 + \sigma_w^2$, yielding $\hat{\sigma}_Y = 0.458$.

(4) For the remaining parameters of the model, without loss of generality, $x_0 = 10$ and $y_0 = 1$ are chosen to be the initial wealth in her pension account and her initial salary level, respectively. The risk aversion coefficient of the investor is $\gamma = 2$. Following the benchmark parameters (cf. Wang (2009)), we assume the wage parameter $\kappa = 0.05$ and wage growth rate variation $\delta = 0.02$. The information-processing costs parameter is $\omega = 0.02$[15], and the personal income tax rate under tax preferences is $\tau' = 0.2$.

[15]In calibrating the information-processing costs parameter ω in the function $\frac{\omega}{2}\phi^2$, we follow the suggestion to link it directly to the entropy-based information-processing constraint of Sims (2003) and the corresponding Kalman filter. Specifically, the Kalman gain G-which measures how much uncertainty is removed upon receiving new signals-provides an empirically grounded metric for setting ω. Drawing on the estimates in Coibion and Gorodnichenko (2015), the implied Kalman gain for professional forecasters ranges between about 0.3 and 0.7, corresponding to an average frequency of information updating of once every 4-7 months in a sticky-information interpretation, or equivalently to agents placing roughly 30%-70% weight on new information in an imperfect-information setting. We calibrate ω so that the model's endogenous information-acquisition rate matches this empirically observed pace of updating, thereby ensuring that the degree of information rigidity in the model aligns with the micro-foundations of rational inattention and the documented behavior of forecasters. This calibration is critical not only for generating realistic consumption-investment rules but also for obtaining reliable welfare implications of rational inattention.

Table 9: Parameter Estimation Results of Wage Process

Panel A:		
Dependent variable: $Wage_{it}$	(1)	(2)
$Year_{2016}$	0.129***	0.014
	(0.021)	(0.020)
$Year_{2018}$	0.292***	0.055***
	(0.022)	(0.024)
$Year_{2020}$	0.483***	0.218***
	(0.023)	(0.026)
$Year_{2022}$	0.645***	0.329***
	(0.023)	(0.028)
Constant	10.280***	3.301***
	(0.016)	(0.295)
Controls	No	Yes
Observations	15201	15201
Within R2	0.113	0.230
Panel B:		
σ_v^2	0.062***	
	(0.005)	

Note: Robust standard errors are reported in parentheses; *, **, *** indicate significance at 10%, 5%, and 1% levels, respectively.

To clearly illustrate the mechanisms through which rational inattention affects consumer decision-making behavior within the model, particularly the influence of core model parameters (information acquisition and tax preferences), we consider three sets of values for the information search cost ω and the personal income tax rate τ', as detailed below.

A. Baseline scenario: $\omega = 0.02$, $\tau' = 0.2$;

B. Reduced information search cost: $\omega = 0.01$, $\tau' = 0.2$;

C. Increased income tax rate: $\omega = 0.02$, $\tau' = 0.4$.

In the figures that follow, these three scenarios are represented by a solid black line, a dashed blue line, and a dotted red line, respectively. By comparing scenarios A and B, we can assess the impact of changes in financial market information search costs on consumer decisions under wage growth uncertainty. By contrasting scenarios A and C, we can discern the influence of variations in personal income tax rates on consumer decisions amidst the complexity of tax calculation.

4.2. Attention allocation mechanism underlying rational inattention

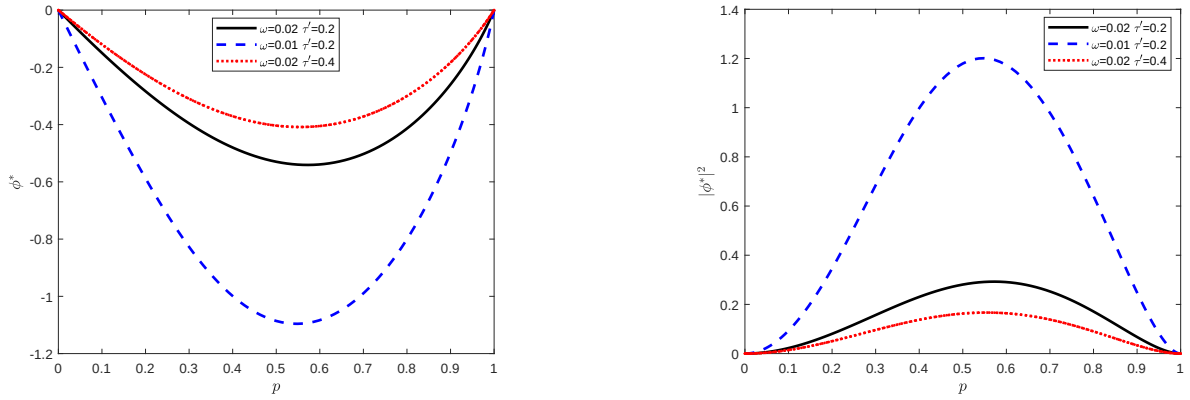


Figure 3: Reaction to financial market uncertainty and attention allocation to information.

Figure 3 unveils the attention allocation mechanism underlying rational inattention. As illustrated, panel (a) demonstrates that, under baseline parameter values, consumers exhibit a negative response to financial market uncertainty. This negative response implies that consumers perceive a negative correlation between higher wage growth and the returns offered by risky assets, thereby inducing a positive investment demand for risky assets (as depicted in Figure 3 below). Furthermore, according to the definition of the belief p, when p approaches 0.5, consumers' uncertainty about future wage growth rates is maximized. As consumers become more uncertain about whether wage growth is in a high or low state (p gradually approaches 0.5), agents exhibit a greater unexpected response to the financial market, increasing their attention allocation to counteract the

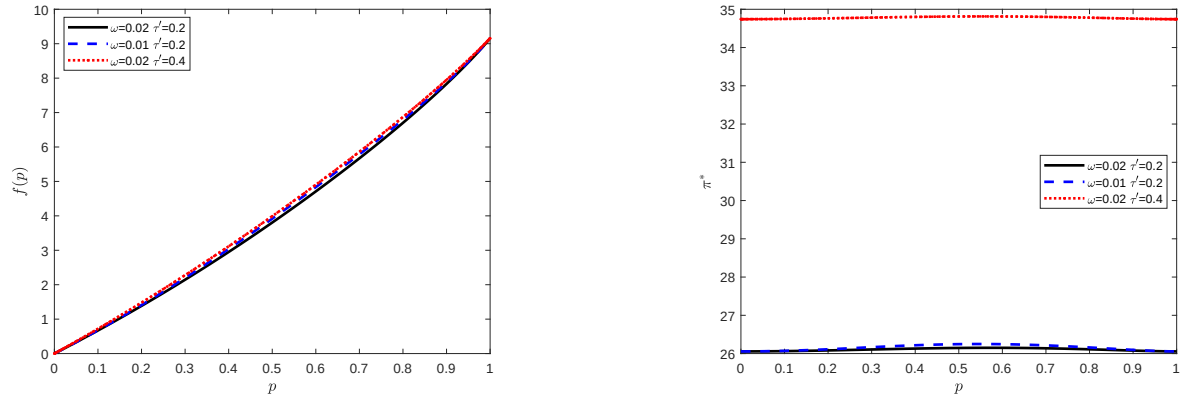
adverse effects of income growth uncertainty. Panel (b) presents the relationship between the optimal attention allocation $(\phi^*(t))^2$ chosen by agents and consumers' beliefs about wage growth. For rationally inattentive consumers, when their degree of rational inattention is higher, i.e., they are more uncertain about the state of income growth rates, they are willing to incur greater costs to acquire more information to update their beliefs, thereby selecting information with higher precision.

Comparing Scenarios A and B, we find that when the cost of searching for financial market information decreases, consumers' attention allocation to financial market information increases (from the solid black line to the dashed blue line). This is because a reduction in information search costs implies lower costs faced by consumers, partially mitigating the adverse effects of rational inattention or indicating higher financial literacy among consumers, making it easier for them to acquire information. Consequently, they choose to pay attention to more information. Further analysis reveals that the gap between the solid black line and the dashed blue line is largest near $p = 0.5$, where consumers face the greatest uncertainty about wage growth rates. As p approaches 0 or 1, consumers' uncertainty gradually decreases, reducing the need for information search and causing the gap to narrow and approach zero.

Comparing Scenarios A and C, we observe that when the individual income tax rate increases, consumers' attention allocation to financial market information decreases (from the solid black line to the dotted red line). An increase in the individual income tax rate raises consumers' welfare costs, exacerbating the adverse effects of rational inattention. Although participating in a tax-deferred individual third-pillar pension plan can provide tax benefits, the disposable income portion will bear additional taxes. As a result, the overall tax cost borne by consumers increases, leading to a reduction in their attention allocation to financial market information. This also provides an explanation, based on the attention allocation perspective of rational inattention, for the empirical finding that agents with higher degrees of rational inattention purchase less commercial pension insurance.

4.3. Impact of rational inattention on investment

Figure 4 elucidates the impact of consumers' rational inattention characteristics on the learning-induced certainty equivalent wealth level and optimal investment strategy. According to the definition of the belief p, when p approaches 0.5, consumers' uncertainty about future wage growth rates is maximized. Panel (a) demonstrates that as consumers' belief p increases, the learning-induced certainty equivalent wealth level $f(p)$ rises. Since a higher p implies a higher perceived wage growth rate, according to the definition of $f(p)$, an increasing p indicates that consumers



(a) (b)

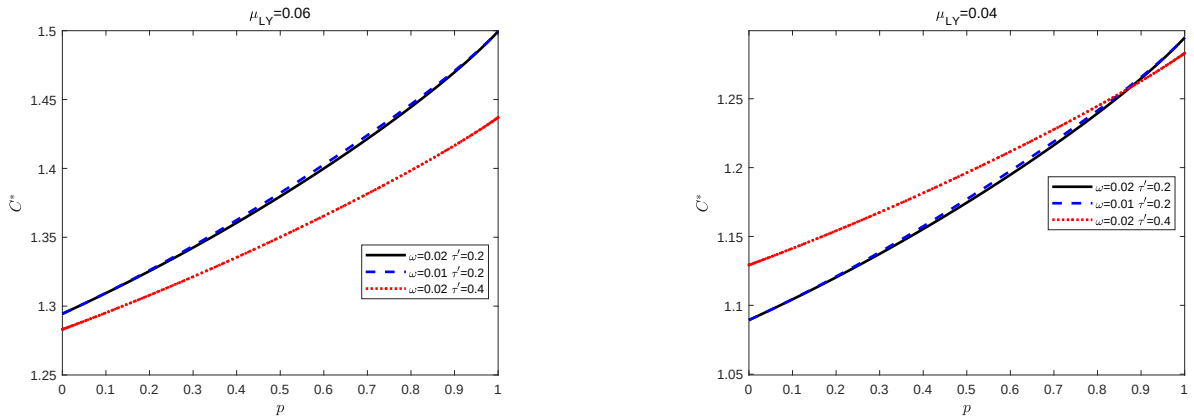
Figure 4: Learning-induced deterministic equivalent wealth level and optimal investment strategy.

anticipate a significant increase in their future wage income, thus leading to a higher certainty equivalent wealth level $f(p)$. Panel (b) illustrates that when consumers' belief p is around 0.5, their investment in stocks reaches its maximum. This is because when p is close to 0.5, consumers' uncertainty about future wage growth rates is at its peak, corresponding to the highest degree of rational inattention among investors. Based on previous analysis, the optimal investment strategy $\pi^*(t) = \frac{\theta}{(1-\tau')\gamma r\sigma_S} - \phi^*(t) * \frac{\delta p(1-p)f'(p)}{\sigma_S}$. The first term represents the traditional optimal investment strategy in the absence of information gathering and rational inattention, while the second term captures the additional investment demand induced by limitedly rational consumers acquiring information from the financial market, which depends on their attention to financial market signals. Therefore, when investors exhibit a high degree of rational inattention, they need to incur greater costs to acquire more information to update their beliefs, thereby choosing higher-precision information. This increased attention allocation leads to a greater influence of financial market signals on investment strategies. Consequently, in conjunction with the analysis in Figure 2, the optimal investment strategy behaves as depicted in the figure.

Comparing Scenarios A and B, we find that when the cost of searching for financial market information decreases, both the learning-induced certainty equivalent wealth level $f(p)$ and the optimal investment wealth in the stock market $\pi^*(t)$ increase (from the solid black line to the dashed blue line). This result naturally stems from the reduction in information search costs, which mitigates the adverse effects of consumers' rational inattention. Further analysis reveals that the gap between the solid black line and the dashed blue line is largest near $p = 0.5$, where consumers face the greatest uncertainty about wage growth rates. As p approaches 0 or 1, consumers' uncertainty gradually decreases, and the degree of rational inattention diminishes, reducing the need for information search. Consequently, the impact of rational inattention gradually lessens and converges.

Comparing Scenarios A and C, we observe that when the individual income tax rate increases, both the learning-induced certainty equivalent wealth level $f(p)$ and the optimal investment wealth in the stock market $\pi^*(t)$ rise (from the solid black line to the dotted red line). In the expression for the optimal investment strategy $\pi^*(t)$, when τ' increases, the denominator of the first term decreases, amplifying the overall impact of the first term. Although the second term decreases overall, the influence of τ' is diluted by various factors, making its overall impact less significant than that of the first term. Therefore, this indicates that when the tax rate increases, rationally inattentive consumers increase both their learning-induced certainty equivalent wealth level and their investment in the stock market.

4.4. Impact of rational inattention on consumption



(a) (b)

Figure 5: Optimal consumption under different wage growth rates for consumers.

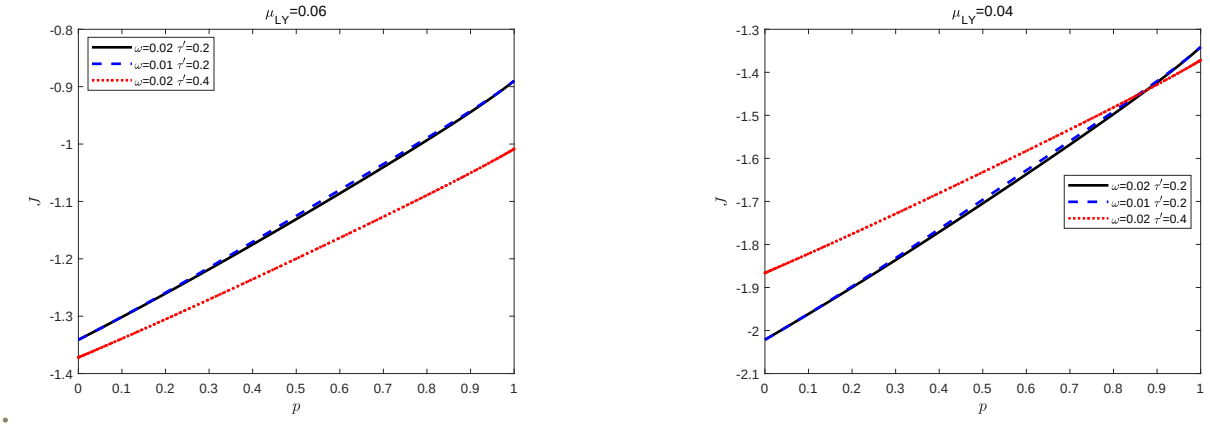
Figure 5 elucidates the impact of consumers' rational inattention characteristics on their consumption behavior. Panels (a) and (b) of Figure 5 present consumers' consumption levels under different wage growth rates (μ_{LY}). Both panels demonstrate that as consumers' belief p increases, the certainty-equivalent wealth level induced by learning, denoted as $f(p)$, rises. According to the consumption strategy expressed in Equation (18), a higher $f(p)$ leads to elevated consumption levels.

On one hand, a comparison of panels (a) and (b) reveals two key findings. First, at higher wage growth rates μ_{LY}, consumers exhibit higher consumption levels, indicating that an increase in wage income boosts disposable income, thereby stimulating consumption. Second, across different wage growth rates (μ_{LY}), optimal consumption increases as the cost of information search decreases. This occurs because a reduction in search costs enhances the efficiency of attention allocation, diminishes the degree of rational inattention, and naturally elevates consumption levels. Consequently, the

adverse effects of rational inattention diminish, reducing expenditures on information search and increasing both consumption and contributions. However, the influence of individual income tax rates on optimal consumption varies, as evidenced by the inconsistent relationship between optimal consumption and tax rate changes across different belief probabilities and wage growth rates. This suggests that under tax deferral policies, despite the presence of tax incentives, significant differences in wage growth rate prediction accuracy and tax rate perception accuracy among consumers with varying degrees of rational inattention lead to divergent decision-making behaviors. This aligns with empirical observations, and the theoretical model further provides a clearer theoretical mechanism.

On the other hand, a comparison of the three curves in either panel (a) or (b) (taking (a) as an example) illustrates the impact of changes in financial market information search costs on consumption. As ω decreases, signifying lower information search costs, consumers' utility improves, leading to increased consumption. From the black solid line to the blue dashed line, comparing Scenario A and Scenario B, reveals the effect of changes in information search costs. From the black solid line to the red dotted-dashed line, comparing Scenario A and Scenario C, reveals the impact of changes in individual income tax rates on consumption: (1) When consumers anticipate a higher wage growth rate (μ_{LY}), an increase in the tax rate enhances their willingness to participate in the third-pillar pension plan but also increases the tax cost of disposable income, resulting in a decrease in optimal consumption; (2) When consumers anticipate a lower wage growth rate (μ_{LY}), particularly when p is close to 0, they perceive the wage growth rate as too low ($\mu_{LY} = 0.04$). In this case, consumers recognize a decline in the intertemporal substitution effect of disposable income, leading to an increase in optimal consumption as the tax rate rises; (3) As p gradually increases, the probability of consumers expecting a high income growth rate (μ_{HY}) rises (as defined in Equation (12)). Consequently, the situation aligns with (1), where an increase in the tax rate leads to a decrease in optimal consumption.

Since consumers' wage income comprises three parts: consumption, pension contributions, and information search expenditures, it is evident that consumers with different degrees of rational inattention exhibit significant differences in their prediction accuracy of wage growth rates and the accuracy of their tax rate perceptions. These differences represent various scenarios mentioned above, leading to the conditional nature of whether "tax deferral policies promote the purchase of commercial pension insurance." This behavior is not always unconditionally valid and is considerably influenced by rational inattention.



(a) (b)

Figure 6: Value function under different wage growth rates for consumers.

4.5. Impact of rational inattention on value function

Figure 6 illustrates the impact of consumer rational inattention on the value function. Based on the expression of the value function (24), it is evident that the value function is primarily influenced by the consumer's optimal consumption. Consequently, the overall characteristics of the value function exhibit similarities to those depicted in Figure 4. Specifically: (1) As the consumer's belief p increases, indicating a greater anticipated growth in future wage income, the certainty-equivalent wealth level $f(p)$ induced by learning rises, leading to an elevation in the consumer's value function. (2) Under higher wage growth rates (μ_{LY}), the consumer's value function is also elevated. Naturally, higher income levels result in higher levels of the value function (utility). (3) As ω decreases, signifying a reduction in information search costs, the detrimental effects of consumer rational inattention diminish, subsequently enhancing the consumer's utility, i.e., increasing the value function. (4) When consumers experience a higher wage growth rate (μ_{LY}), an increase in the personal income tax rate leads to an augmentation of the value function. (5) Conversely, when consumers face a lower wage growth rate (μ_{LY}), the impact of the personal income tax rate on the value function varies across different p values. This economic analysis aligns with the findings presented in Figure 5. Notably, when further examining the impact of information search costs in conjunction with Figures 6 and 5, it is observed that the disparity between the black solid line and the blue dashed line is maximal near $p = 0.5$, where consumers encounter the greatest uncertainty regarding wage growth rates. As p approaches 0 or 1, consumer uncertainty gradually decreases, thereby reducing the demand for information search. This results in a narrowing utility gap, which tends towards zero.

5. Theoretical mechanism: Supplement

In this section, we further extend the model to supplement and refine the theoretical mechanisms. We examine the impacts of multiple saving channels and reframing rational inattention (with policy complexity as the core friction) on pension contribution and investment behaviors. First, we consider how the fundamental results of the model change when investors are confronted with multiple investment and saving channels. Second, we delve deeper by assuming that investors do not exhibit rational inattention solely with respect to individual parameters such as tax rate parameter calculations and uncertainty regarding wage growth. Instead, they display characteristics of rational inattention towards the entire framework of tax-deferred pension plans. Specifically, consumers face cognitive costs in acquiring and processing information about third-pillar reform policies, including eligibility criteria, tax benefits, and contribution rules. We discuss how the rational inattention traits of investors are influenced by the complexity arising from these common factors.

5.1. A multi-channel savings framework

In our baseline model, the sole saving channel is through commercial pension contributions. This design mechanically links any factor discouraging savings (like wage uncertainty or high information costs) to reduced pension contributions, making it impossible to isolate the specific drivers of pension take-up versus general savings. To provide a credible framework for analyzing pension participation, we propose a revised model incorporating multiple saving channels. Specifically, we consider two types of wealth processes for investors: the pension wealth (wealth within the tax-deferred commercial pension account) and the liquid wealth (liquid account). Investors earn a stochastic wage, and a certain proportion of this wage is contributed to the pension account as pension contributions, which are tax-exempt. The after-tax wage is then used for consumption, investment, and the allocation of attention (to cover information search costs) in the liquid account. Investors optimally allocate their assets between risky and risk-free assets within both accounts. The utility of investors derives from their consumption during the working period and the terminal wealth levels in both accounts upon retirement. The specific modeling is as follows:

We extend the standard life-cycle model with three distinct wealth components, and denote these wealth components as the investor's liquid wealth $W(t)$ (taxable assets invested across risk-free bonds and stocks), pension wealth $X(t)$ (tax-deferred commercial pension assets), wage process $Y(t)$ (stochastic labor income stream). For the control variables: $C(t)$ is the consumption rate, $\pi_W(t)$ is the risky asset allocation percentage for liquid wealth, $\pi_X(t)$ is the risky asset allocation

percentage for pension wealth, $Cont(t)$ is the pension contribution rate, and $\phi(t)$ is investor's attention devoted to policy understanding.

As a result, the wealth dynamics are governed by:

$$dW(t) = [(r + \pi_W(t)(\mu_S - r))W(t) + (1 - \tau)(Y(t) - Cont(t)) - C(t) - \chi(\phi(t))]\, dt + \pi_W(t)\sigma_S W(t) dW_S(t), \tag{26}$$

$$dX(t) = [(r + \pi_X(t)(\mu_S - r))X(t) + Cont(t)]\, dt + \pi_X(t)\sigma_S X(t) dW_S(t), \tag{27}$$

$$dY(t) = \mu_Y Y(t)dt + \sigma_Y Y(t) dW_Y(t), \tag{28}$$

where τ is the marginal tax rate, $Cont(t)$ denotes pension contributions, and $\chi(\phi(t))$ represents information acquisition costs.

In this extended model, the main innovations include these three characteristics. Firstly, we construct explicit separation of taxable vs. tax-deferred savings channels in a life-cycle model. Secondly, we model endogenous pension contributions ($Cont(t) \geq 0$) as distinct choice variable. Thirdly, we characterize the accurate EET tax treatment: Contributions reduce current taxable income.

5.2. Reframing rational inattention

The investor displays characteristics of rational inattention towards the entire framework of tax-deferred pension plans.[16] Especially, the rational inattention of processing policy-specific information contains two aspects of issues. First, the investor has imperfect knowledge of the true tax benefit parameter τ_{eff}, which is the effective net present value (NPV) of the tax benefit per unit contribution, and represents the effective tax saving rate. It depends on the statutory marginal tax rate τ, the deferral period, the expected return within the pension, the tax rate at withdrawal, and potential future policy changes. Therefore, calculating τ_{eff} precisely is complex. Second, the investor has imperfect knowledge of complexity cost χ_{eff}, which is the complexity cost due to the diversity of third-pillar reform policies (such as eligibility criteria, tax benefits or not, and contribution rules). It captures the cognitive effort, time, and potential hassle costs associated with understanding eligibility rules, choosing a provider, selecting funds, managing the account, and anticipating future withdrawal rules. χ_{eff} is inherently subjective but likely correlated with observable policy design complexity.

[16] Inattention to the wage growth process, while relevant for general savings and portfolio choice, is less central to explaining pension take-up (contribution) than inattention to the policy itself. The revised model in this section shifts the focus of rational inattention to the cognitive costs of understanding the third-pillar policy.

Next, we characterize information acquisition. Notice that τ_{eff} and χ_{eff} are unobserved state variables, showing the investor's facing uncertainty about key policy parameters crucial for evaluating the pension's attractiveness. As a result, the investor can exert attention effort $\phi(t)$ (time spent researching, consulting advisors, filling forms) to learn about τ_{eff} and χ_{eff}. This effort is costly.[17]

Therefore, we model rational inattention as cognitive costs of processing policy-specific information:

Assumption 5.1 (Policy parameters)**.** *The investor faces uncertainty about:*

1. *Effective tax benefit τ_{eff}: NPV of tax savings per unit contribution*
2. *Complexity cost χ_{eff}: Cognitive burden of participation*

Next, regarding how to depict the operational mechanism of rational inattention, the approach is similar to the construction of the benchmark model in Section 3. It is assumed that the learning process and the information (signal) obtained by the investor is

$$dI_\tau(t) = \tau_{\text{eff}}dt + \frac{\sigma_\tau}{\phi(t)}dW_\tau(t), \tag{29}$$

$$dI_\chi(t) = \chi_{\text{eff}}dt + \frac{\sigma_\chi}{\phi(t)}dW_\chi(t), \tag{30}$$

where σ_τ and σ_χ are fundamental uncertainty about policy parameters; $\frac{1}{\phi(t)}$ shows that signal noise decreases with effort (when $\phi(t)$ increases, the noise decreases, then signal precision increases). This captures the idea that more effort yields better information; $dW_\tau(t)$ and $dW_\chi(t)$ are independent Brownian motions representing signal noise.

In this case, the investor's belief formation is also filtering: She update her beliefs $p(t) = \mathbb{E}[\tau_{\text{eff}}|\mathcal{F}_t]$ and $q(t) = \mathbb{E}[\chi_{\text{eff}}|\mathcal{F}_t]$ using Kalman-Bucy filtering or Bayesian updating based on the signals $I_\tau(t)$, $I_\chi(t)$ and their prior beliefs. The dynamics of $p(t)$ and $q(t)$ will depend on $\phi(t)$, σ_τ, σ_χ, and the signal history. Higher effort $\phi(t)$ leads to faster belief updating and lower posterior variance (reduced uncertainty).

[17] For example, we assume the cost function is $\chi(\phi) \equiv \frac{\omega}{2}(\phi(t))^2$ in section 3. In fact, Dewan and Neligh (2020) investigate and estimate information cost functions in models of rational inattention. They find that most subjects exhibit monotonicity in performance with respect to potential rewards, and there is mixed evidence on continuity and convexity of costs. Moreover, a significant portion of subjects are likelier to make small mistakes than large ones, contrary to the predictions of mutual information. This suggests that while people are generally rationally inattentive, their cost functions may display non-convexities or discontinuities, or they may incorporate some notion of perceptual distance.

5.3. Filtering and HJB equation

Define value function:

$$J(t,\mathbf{s}) = \max_{\mathbf{u}} \mathbb{E}_t \left[\int_t^T e^{-\beta(s-t)} U(C_s) ds + e^{-\beta(T-t)} B(W_T, X_T) \right],$$

where $\mathbf{s} = (W, X, Y, p, q)$, $\mathbf{u} = (C, \pi_W, \pi_X, Cont, \phi)$.

The Hamilton-Jacobi-Bellman equation characterizes optimal policies:

$$\begin{aligned} 0 = \max_{C,\pi_W,\pi_X,Cont,\phi} \Big\{ & U(C) - \beta J + J_W \left[(r + \pi_W(\mu_S - r))W + (1-\tau)(Y - Cont) - C - \chi(\phi) \right] \\ & + J_X \left[(r + \pi_X(\mu_S - r))X + Cont \right] + J_Y \mu_Y Y \\ & + \frac{1}{2} J_{WW} \pi_W^2 \sigma_S^2 W^2 + \frac{1}{2} J_{XX} \pi_X^2 \sigma_S^2 X^2 + \frac{1}{2} J_{YY} \sigma_Y^2 Y^2 \\ & + J_{XW} \pi_W \pi_X \sigma_S^2 W X + \mathcal{D}(p(t), q(t)) \Big\}, \end{aligned} \tag{31}$$

where $\mathcal{D}(p(t), q(t))$ captures belief evolution dynamics.

In the rational inattention model, the expression for the belief updating process $D(p_t, q_t)$ needs to be derived using the Kalman-Bucy filtering theory. Below is the main derivation process and the resulting expression. We assume that the true parameters τ_{eff} and χ_{eff} are constants, but unknown to investors. The signal processes are given by Eqs. (29)-(30), and $\phi(t)$ represents attention effort (a control variable). The evolution of the investor's beliefs $p(t) = \mathbb{E}[\tau_{\text{eff}}|\mathcal{F}_t]$ and $q(t) = \mathbb{E}[\chi_{\text{eff}}|\mathcal{F}_t]$ follows the Kalman-Bucy filtering equations:

$$\begin{aligned} dp(t) &= \underbrace{\frac{V_p(t)\phi^2(t)}{\sigma_\tau^2}}_{\text{Kalman gain}} \left(dI_\tau(t) - p(t)dt \right), \\ dq(t) &= \frac{V_q(t)\phi^2(t)}{\sigma_\chi^2} \left(dI_\chi(t) - q(t)dt \right), \end{aligned}$$

where $V_p(t) = \mathbb{E}[(\tau_{\text{eff}} - p(t))^2|\mathcal{F}_t]$ and $V_q(t) = \mathbb{E}[(\chi_{\text{eff}} - q(t))^2|\mathcal{F}_t]$ are the posterior variances.[18]

As a result, in the Hamilton-Jacobi-Bellman (HJB) equation for the value function $J(t, W, X, Y, p, q)$, $D(p(t), q(t))$ captures the contribution of belief updates. Its expression is:[19]

$$D(p(t), q(t)) = \frac{1}{2} \frac{V_p^2(t)\phi^2(t)}{\sigma_\tau^2} \frac{\partial^2 J}{\partial p^2} + \frac{1}{2} \frac{V_q^2(t)\phi^2(t)}{\sigma_\chi^2} \frac{\partial^2 J}{\partial q^2}.$$

[18]The posterior variances satisfy the following dynamic evolution processes.

$$\frac{dV_p(t)}{dt} = -\frac{V_p^2(t)\phi^2(t)}{\sigma_\tau^2}, \frac{dV_q(t)}{dt} = -\frac{V_q^2(t)\phi^2(t)}{\sigma_\chi^2}.$$

[19]Please refer to Appendix B for complete derivation process.

We have some economic implications for previous analysis as follows. (1) In the belief evolution dynamics $\mathcal{D}(p(t), q(t))$, the terms $\frac{V_p^2\phi^2}{\sigma_\tau^2}$ and $\frac{V_q^2\phi^2}{\sigma_\chi^2}$ represent the learning efficiency. Specifically, the following three factors will affect this learning efficiency. The first one is attention effort $\phi(t)$, higher effort accelerates belief updates; the second one is signal noise σ_τ, σ_χ, greater noise reduces learning efficiency; the third one is uncertainty decay V_p, V_q, smaller posterior variance slows learning. (2) The second derivatives $\frac{\partial^2 J}{\partial p^2}$ and $\frac{\partial^2 J}{\partial q^2}$ capture the curvature effects (convexity/concavity), which show the impact of belief uncertainty on marginal value. For example, if J is convex in p, higher uncertainty increases the value function. (3) There are endogenous learning incentives. The optimal attention $\phi^*(t)$ is determined in the HJB maximization:

$$\phi^* = \arg\max_{\phi} \left\{ -J_W \chi(\phi) + \frac{1}{2}\left(\frac{V_p^2\phi^2}{\sigma_\tau^2}\frac{\partial^2 J}{\partial p^2} + \frac{V_q^2\phi^2}{\sigma_\chi^2}\frac{\partial^2 J}{\partial q^2} \right) \right\},$$

where $\chi(\phi)$ is the attention cost function.

In conclusion, $D(p(t), q(t))$ quantifies the impact of belief updates on investment decisions under rational inattention. Its core lies in the trade-off between learning efficiency and uncertainty. The second-derivative terms reveal how belief uncertainty affects optimal strategies through the curvature of the value function, providing a theoretical basis for policy design (e.g., reducing complexity $\sigma_\chi \downarrow$ or enhancing transparency $\sigma_\tau \downarrow$).

5.4. Theoretical insights

Given that our primary focus lies on the variables of pension contribution *Cont* and attention allocation strategy ϕ, we derive the following two core propositions regarding these two decision variables based on the HJB equation and its corresponding first-order conditions. The detailed mathematical derivation steps are provided in Appendix C, and here we present the final results.

Proposition 5.2 (Optimal pension contributions)**.** *The optimal contribution rule satisfies:*[20]

$$Cont^*(t) = \Psi\left(\underbrace{p(t)}_{\text{Tax benefit}}, \underbrace{q(t)}_{\text{Policy cost}}, \underbrace{J_W(1-\tau)}_{\text{Marginal cost of reducing liquid wealth}}, \underbrace{J_X}_{\text{Benefit of contribution}} \right) \tag{32}$$

with $\partial\Psi/\partial p(t) > 0$, $\partial\Psi/\partial q(t) < 0$.

[20]Notice that since $J_X = (1-\tau)J_W$ by F.O.C., the term "$J_W(1-\tau)$" is actually redundant. The proposition can also be stated as follows: $Cont^*(t) = \Psi(p(t), q(t), \text{state variables})$. It shows that the marginal benefit of contributions (marginal value of pension wealth) equals the after-tax marginal cost (marginal value of liquid wealth). What is more, it should be noted that the theoretical results of this proposition and the following one can both be verified by the numerical solution results. For specific details, please refer to the analysis in the subsequent section.

In Proposition 5.2, we find that the interplay between tax benefits and complexity costs is pivotal in determining household contribution behavior within pension systems. Contributions effectively balance the advantages derived from tax incentives against the burdens imposed by complexity costs. Specifically, households exhibit a heightened propensity to contribute when the effective tax benefits (τ_{eff}) are well-understood, as evidenced by an upward trend in $p(t)$, reflecting increased awareness or clarity regarding these incentives. Conversely, a reduction in complexity costs (χ_{eff}), denoted by a downward shift in $q(t)$, further encourages contributions by minimizing the cognitive and operational burdens associated with decision-making. Thus, optimizing the design of pension systems to enhance the transparency of tax benefits while simultaneously reducing complexity costs is crucial for fostering greater household participation and ensuring the sustainability of retirement savings frameworks.

Proposition 5.3 (Optimal attention allocation)**.** *Optimal attention $\phi(t)^*$ solves:*

$$\chi'(\phi^*) = \underbrace{\frac{1}{J_W}\left(\frac{\partial \mathcal{D}}{\partial \phi}\right)}_{\textit{Value of Information}} \tag{33}$$

and increases with: (1) Financial literacy ($\downarrow \chi'$); (2) Income $Y(t)$; (3) Tax rate τ.

In Proposition 5.3, we find that the extent to which households allocate attention to pension-related decisions is contingent upon several key factors that influence their engagement and investment behavior. Households are more inclined to invest attention when the cost of acquiring relevant knowledge is relatively low, which often correlates with higher levels of financial literacy. This affordability of learning empowers households to make more informed decisions, thereby increasing their willingness to engage with complex financial products. Additionally, the magnitude of potential rewards plays a critical role; households with higher incomes or those subject to favorable tax rates are more likely to invest attention, as the anticipated benefits outweigh the perceived costs. Finally, a reduction in policy complexity, as measured by a decrease in χ_{eff}, further incentivizes households to invest attention by simplifying the decision-making process and reducing the cognitive load associated with understanding intricate policy details. Collectively, these factors underscore the importance of designing pension systems that are both financially rewarding and cognitively accessible, thereby fostering greater household participation and enhancing the overall efficiency of retirement savings mechanisms.

In order to show the structure of the optimal strategies more clearly, we conduct a detailed analysis (decomposition and cross-terms) as follows. In the first part, to clearly identify the impact mechanisms of key model parameters, we omit cross-terms in the baseline analysis scenario. By taking the first-order condition of the HJB equation with respect to the control variable, we obtain:

1. Consumption C^*:

$$U'(C^*) = J_W;$$

2. Risky asset allocation percentage for liquid wealth π_W^*:

$$\pi_W^* = -\frac{J_W}{J_{WW}W} \cdot \frac{\mu_S - r}{\sigma_S^2};$$

3. Risky asset allocation percentage for pension wealth π_X^*:

$$\pi_X^* = -\frac{J_X}{J_{XX}X} \cdot \frac{\mu_S - r}{\sigma_S^2};$$

4. Investor's attention devoted to policy understanding ϕ^*:

$$\chi'(\phi^*) = \frac{1}{J_W}\left(\frac{V_p^2\phi^*}{\sigma_\tau^2}J_{pp} + \frac{V_q^2\phi^*}{\sigma_\chi^2}J_{qq}\right).$$

The analysis of key economic implications reveals multifaceted interactions between rational inattention parameters, multi-channel savings dynamics, and lifecycle decision-making patterns. First, learning efficiency parameters exhibit asymmetric effects: when tax incentive opacity σ_τ increases, agents reduce attention allocation ϕ^*, leading to inaccurate beliefs about policy parameters $(p(t))$ and lower optimal contributions $Cont^*$. Conversely, enhanced transparency in policy complexity σ_χ reduces learning costs, prompting lower attention levels and more accurate perceptions of pension account features $(q(t))$. This relationship is further moderated by financial literacy, as households with lower cognitive capacity face steeper attention costs $\chi(\phi)$, resulting in greater belief errors and lower contributions compared to their literate counterparts.

In the realm of savings allocation, cross-account risk management emerges as a critical mechanism. The optimal risk allocation between liquid and pension accounts satisfies

$$\pi_W^* > \pi_X^* \quad \text{when} \quad \left|\frac{J_W}{J_{WW}W}\right| > \left|\frac{J_X}{J_{XX}X}\right|,$$

reflecting the risk-mitigating effect of pension accounts' tax advantages and illiquidity. The marginal contribution condition $J_X = (1-\tau)J_W$ formalizes the trade-off between pre-tax returns and after-tax costs, where higher tax incentives $(\tau \uparrow)$ directly stimulate contributions by aligning private and social returns.

Thus far, we have reached the primary conclusions of the first part of this section.

Corollary 5.4 (Pension participation barriers and policy tool efficacy)**.** *(1) Barriers to pension participation: Multiple savings channels dilute pension tax incentives, and underestimation of effective tax benefits (τ_{eff}) creates invisible barriers, especially for low-literacy households, leading them to prioritize other forms of wealth over pension contributions. High complexity costs (χ_{eff}) also suppress attention investment, causing knowledge gaps and reducing contribution rates.*

(2) Policy instrument effectiveness: Policy tools are substitutable. Enhancing τ_{eff} is more effective for high-income groups, while reducing χ_{eff} is universally effective but costly. Smart default options offer a balanced approach, combining efficiency and equity.[21]

Remark 5.5 (Smart default options)**.** Traditional pension systems require individuals to make active decisions (e.g., contribution rates, investment portfolios), but most people face issues of "decision fatigue" or limited information-processing capacity. Smart default options alleviate individual decision-making burdens by pre-setting reasonable parameters (e.g., automatically adjusting contribution rates based on age and income), thereby avoiding low participation rates caused by hesitation or lack of knowledge, and enhancing the overall operational efficiency of the system. Leveraging big data analytics, smart defaults can match optimal solutions for different groups. This differentiated design gradually improves participation quality through "auto-enrollment plus auto-escalation" mechanisms (Chetty et al., 2014; Choukhmane, 2025).[22]

In the second part, we retain all potential cross-terms in the model (generated by non-independent Brownian motions) to further explore the principal economic implications embedded in the results, with a particular focus on analyzing cross-term effects. The corrected first-order conditions and the expressions for the optimal strategies are as follows:

1. Consumption C^*: $U'(C^*) = J_W$, which shows that at the optimal consumption point, the marginal utility of consumption equals the marginal value of wealth.

2. Risky asset allocation percentage for liquid wealth π_W^*:

$$\pi_W^* = -\frac{J_W}{J_{WW}W} \cdot \frac{\mu_S - r}{\sigma_S^2} - \frac{J_{WX}}{J_{WW}W}\pi_X X - \frac{J_{WY}}{J_{WW}W}\frac{\sigma_Y Y \rho_{WY}}{\sigma_S};$$

3. Risky asset allocation percentage for pension wealth π_X^*:

$$\pi_X^* = -\frac{J_X}{J_{XX}X} \cdot \frac{\mu_S - r}{\sigma_S^2} - \frac{J_{WX}}{J_{XX}X}\pi_W W - \frac{J_{XY}}{J_{XX}X}\frac{\sigma_Y Y \rho_{XY}}{\sigma_S};$$

4. Investor's attention devoted to policy understanding ϕ^*:

$$\chi'(\phi^*) = \frac{1}{J_W}\left[J_{pp}\frac{V_p^2\phi^*}{\sigma_\tau^2} + J_{qq}\frac{V_q^2\phi^*}{\sigma_\chi^2} + J_{pW}\frac{V_p}{\sigma_\tau}\pi_W \sigma_S W \rho_{\tau S}\right].$$

[21] It should be noted that the theoretical results of this corollary and the following one can both be verified by the numerical solution results. For specific details, please refer to the analysis in the subsequent section.

[22] Numerous real-world examples support this approach. For instance, in U.S. "Auto-Enrollment 401(k) Plans", after raising the default contribution rate from 0% to 3%, participation rates soared from 67% to 92%, with a greater increase among low-income earners compared to high-income earners. This demonstrates that smart defaults can simultaneously expand coverage and promote equity. In UK "NEST Pension" plans, by defaulting to "target-date funds" and dynamically adjusting risk levels, it ensures all income groups receive investment returns aligned with their retirement age, preventing wealth disparities caused by differences in investment capabilities.

The analysis reveals intricate cross-account risk spillover mechanisms, where the marginal adjustment in liquid account risk allocation (π_W^*) responds to pension account risk changes (π_X) through the cross-derivative term $-\frac{J_{WX}}{J_{WW}W}X$. When wealth substitution effects dominate ($J_{WX} < 0$), increased pension risk exposure paradoxically induces higher liquid account risk-taking to maintain optimal portfolio risk levels, reflecting a"risk compensation" behavior quantified as $\Delta\text{Total Risk} = \omega\Delta\pi_X + (1-\omega)\Delta\pi_W$. This dynamic interacts with income-market risk hedging through a newly identified cross-term $-\frac{J_{WY}}{J_{WW}W}\frac{\sigma_Y Y \rho_{WY}}{\sigma_S}$, where positive income-market correlation ($\rho_{WY} > 0$) triggers portfolio de-risking ($\pi_W^* \downarrow$) as income streams already carry market exposure.

Attention allocation decisions exhibit cyclical patterns driven by the learning-risk interaction term $J_{pW}\frac{V_p}{\sigma_\tau}\pi_W\sigma_S W\rho_{\tau S}$. Negative correlation between tax incentives and market returns ($\rho_{\tau S} < 0$) creates a counter-cyclical learning effect: market volatility spikes ($\sigma_S \uparrow$) prompt increased attention efforts ($\phi^* \uparrow$) as households seek to accurately assess tax-advantaged accounts' hedging value during turbulent periods. This aligns with lifecycle observations where younger agents demonstrate higher learning efforts during volatile markets, while older agents reduce attention in stable conditions, with temporary learning surges during economic downturns.

Contribution decisions reveal hidden adjustment channels through the indirect effect $\frac{\partial Cont^*}{\partial \pi_X} = -\frac{J_{WX}}{J_{XX}X}\pi_W W$. When pension account risk increases ($J_{WX} < 0$) and liquid accounts hold risky positions ($\pi_W > 0$), households reduce contributions ($Cont^* \downarrow$) to preserve liquidity against rising aggregate risk exposure. Complexity effects further discriminate against low-income households through the second-order term $\frac{\partial^2 \phi^*}{\partial \chi \partial W}$, where limited wealth ($W$) amplifies sensitivity to policy complexity (χ). This creates a vicious cycle: lower financial literacy reduces attention ($\phi^* \downarrow$), leading to suboptimal contributions, while high-income households' larger wealth buffers insulate them from complexity-induced decision errors, maintaining stable contribution patterns.

Targeted learning subsidies should prioritize households with high exposure to market risk ($\sigma_S\pi_W W$, substantial wealth holdings (W), and strong risk correlations (J_{pW}), as quantified by the subsidy priority index $\frac{|J_{pW}|\sigma_S\pi_W W}{\chi''}$. This approach addresses the complexity discrimination phenomenon, where low-income groups face amplified decision-making costs from policy opacity, trapping them in vicious cycles of inadequate learning and suboptimal contributions.

The integration of cross-terms uncovers systemic vulnerabilities in pension decision-making: 1) inter-account risk contagion channels that transmit shocks between pension and liquid accounts; 2) pro-cyclical attention allocation patterns where market crises paradoxically heighten cognitive efforts; and 3) the low-learning trap exacerbating wealth inequality through differential policy responsiveness. These findings underscore the necessity for policy synergy, as isolated interventions risk being neutralized by cross-effect counteractions.

Reform priorities should focus on developing inter-account risk hedging frameworks, implementing state-contingent automatic stabilizers, and creating simplified pension products for low-income earners. As the analysis reveals,"cross-terms expose the hidden transmission channels of economic fragilityignoring them is tantamount to overlooking the subterranean currents of crisis propagation." This necessitates a paradigm shift from siloed policy design to systems-thinking approaches that account for dynamic interdependencies across decision variables.

Next, to sum up, we obtain the main conclusions of the second part of this section.

Corollary 5.6 (Pension participation: Policy-market dynamics and household risk-incentive adaptation)**.** *(1) Policy and market factors shape pension participation: Policy complexity creates barriers, especially for disadvantaged and low-wealth households, reducing their participation and widening wealth gaps by discouraging learning efforts due to cognitive costs. Conversely, market turbulence during economic downturns prompts households to intensify efforts to understand pension benefits as a risk-mitigation strategy.*

(2) Household behavior adapts to risk and incentives: Risk-taking decisions across savings channels show compensatory patterns, with households adjusting exposure in liquid accounts when pension investments become riskier. Tax incentives primarily benefit financially literate households, whereas reducing complexity and implementing automatic enrollment are more effective in boosting participation among vulnerable groups.

5.5. Numerical solutions (with additional verification and explanation of theoretical propositions and corollaries)

Corresponding to the aforementioned theoretical analysis, we have also provided the numerical solution results for this optimization problem. Furthermore, the corresponding numerical findings (presented in the main text and appendix) and the analysis of their economic implications indicate that the aforementioned theoretical analysis is consistent with the numerical results. Specifically, the presented figures (presented in the main text and appendix) offer a comprehensive visualization of the numerical solutions to the optimization problem, providing critical insights into the dynamics of wealth accumulation, belief evolution, and optimal control variables within the context of the model. The results depicted in Figure 7 illustrate the optimal control variables as functions of key state variables, offering crucial insights into the economic mechanisms governing sophisticated financial decision-making under uncertainty. The investment allocation π_W displays a hump-shaped relationship with wealth, initially increasing due to greater risk-bearing capacity before declining

at higher wealth levelsa pattern consistent with decreasing relative risk aversion.[23] Simultaneously, the contribution $Cont$ exhibits strategic complementarity with investment decisions, reflecting the intricate balancing of consumption smoothing, investment opportunities, and precautionary savings motives.

We also find that consumption from the liquid account increases as the wealth level rises, which aligns perfectly with intuition. Additionally, contributions to the pension account increase with income, which, as analyzed earlier, holds true when income has not reached a certain threshold. As discussed previously, the optimal attention allocation generally exhibits a U-shaped time evolution trend, which is also verified in the numerical solution result plots in the appendix.

It is noteworthy that Proposition 5.2 states that the optimal pension contribution $Cont$ increases with belief p and decreases with belief q. In other words, the more effective the tax incentives are, the higher the contribution; conversely, the more complex the tax system is, the lower the contribution. Proposition 5.3 points out that the optimal investor attention ϕ rises with income growth. Moreover, the optimal investor attention and the optimal pension contribution change in the same direction. These theoretical findings have all been effectively validated in the numerical analysis in Figure 7.

The analysis of the theoretical results for Corollary 5.4 follows a similar pattern. Corollary 5.4 further establishes that enhancing τ_{eff} (the effective tax incentive parameter) generates progressively larger effects for higher-income groups, while reducing χ_{eff} (the effective complexity parameter of the tax system) benefits all households but exhibits decreasing marginal effectiveness.. This result requires analysis with the aid of Figure 12 in Appendix A, which reveals that as wealth increases, the investment proportion tends to be relatively lower. Meanwhile, Figure 7 shows that the overall investment proportion is higher under belief q, corresponding to lower wealth levels and thus having a greater impact on low-income individuals; in contrast, the investment proportion is generally lower under belief p, which more significantly affects high-income individuals. Therefore, these findings also provide robust validation for the aforementioned conclusion.

Furthermore, the analysis of conclusion (1) in Corollary 5.6 follows a similar pattern: as q increases, the overall investment proportion rises, corresponding to low-income individuals. Moreover, pension contributions decline, and attention allocation ϕ also decreases. Consequently, policy

[23]This non-monotonic pattern emerges from the interaction between wealth effects and the convex costs of information acquisition. As wealth increases, investors initially expand their risky asset exposure due to greater absolute risk tolerance, but eventually reduce allocations as the marginal benefit of active portfolio management diminishes relative to information costs.

complexity poses a barrier, particularly for vulnerable groups and low-income households, as cognitive costs inhibit learning efforts, leading to a decline in participation rates and exacerbating wealth inequality.

The analysis of conclusion (2) in Corollary 5.6 is even more profound and intriguing. First, integrating the preceding empirical evidence, it can be observed that households with higher financial literacy possess a more accurate perception of tax rates, meaning they are more sensitive to tax incentive parameters and have a clearer understanding of tax system costs. This facilitates better cost control, both of which collectively contribute to increased pension contributions and higher participation rates in commercial pension plans[24]. Second, regarding the statement that "reducing complexity and implementing automatic enrollment are more effective in boosting participation among vulnerable groups," the model does not provide direct evidence. However, it constitutes an intuitive and valid inference, as streamlining processes and implementing automatic enrollment more clearly highlight the advantages of tax incentives while mitigating the negative impacts of tax system complexity, all of which contribute to enhancing participation rates in commercial pension insurance plans.

The belief-contingent control variations reveal particularly profound economic implications. Optimal policies demonstrate significant sensitivity to beliefs (p, q), with information acquisition ϕ intensifying during periods of heightened epistemic uncertainty. This endogenous learning behavior represents a fundamental departure from traditional portfolio models, as agents actively manage their information sets rather than passively accepting exogenous information structures. The resulting dynamics suggest that belief-dependent hedging motives substantially alter conventional asset allocation rules, with agents reducing exposure to ambiguous assets while simultaneously investing in uncertainty resolution.

For enhanced readability and logical coherence, the three-dimensional representations of the value function and control surfaces, control variables time evolution under different wealth levels and different beliefs, Belief Evolution and Control Relationships have been placed in the Appendix A.

Finally, we provide a summary of economic mechanisms in this section, which is shown by the following table 10.

[24]Since the model does not directly incorporate financial literacy as a variable, this conclusion can only be indirectly validated through Figure 7.

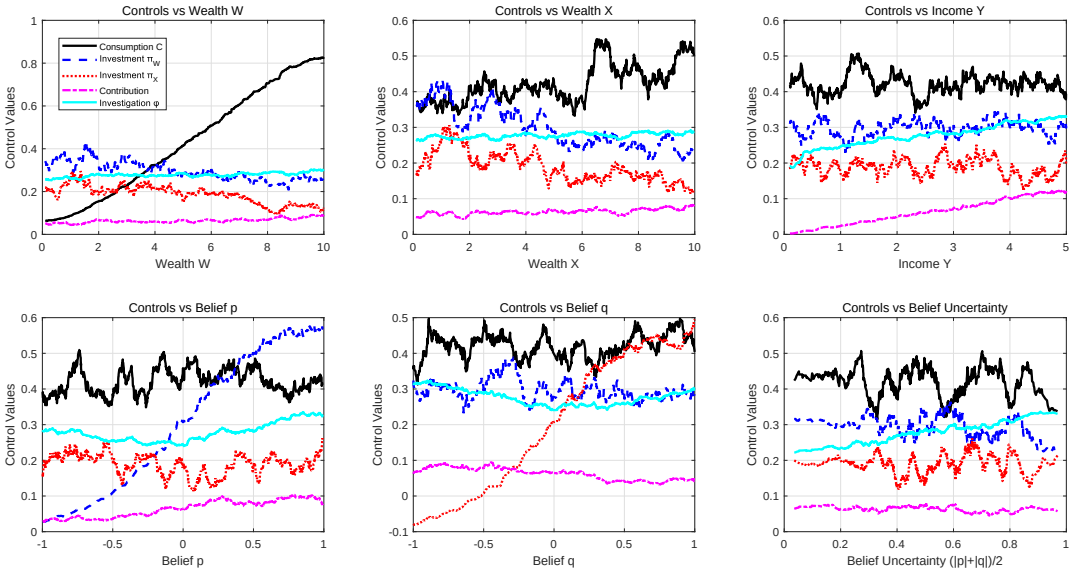


Figure 7: Control variables relationships analysis.

6. Welfare analysis with CARA utility

6.1. Equivalent consumption variation derivation

While Sims (2003) and subsequent work (Luo, 2017; Luo & Young, 2010; Miao et al., 2022; etc) focus on general information constraints, our policy-specific inattention framework reveals how complexity costs (χ_{eff}) create distinctive barriers in retirement planning. Unlike macroeconomic inattention, which generates aggregate inertia, policy complexity induces distributional effects that disproportionately affect cognitively constrained householdsa novel mechanism with important welfare implications.

The measurement of welfare effects under behavioral frictions requires careful consideration of utility comparability. We employ equivalent consumption variation (ECV for short) as our principal welfare metric, which quantifies the permanent consumption increment rendering households indifferent between baseline and reform policy regimes. For households with constant absolute risk aversion (CARA) preferences (consistent with Eq. (11) in Section 3) represented by $u(c) = -\frac{e^{-\gamma c}}{\gamma}$, the ECV admits a closed-form solution that circumvents numerical approximation errors while preserving cardinal utility properties.

Formally, let $U^b(\mathbf{s}_0)$ and $U^k(\mathbf{s}_0)$ denote the lifetime utilities under baseline and reform policies respectively, for a household with initial state vector $\mathbf{s}_0 = (W_0, X_0, Y_0, p_0, q_0)$. The ECV solves:

$$U^k(\mathbf{s}_0) = \mathbb{E}_0 \left[\int_0^T e^{-\beta t} u\left((1 + \text{ECV}) c_t^b \right) dt \right] \tag{34}$$

Table 10: Economic mechanisms and policy implications.

Decision	Key Drivers	Effect
Pension Contributions	↑ Perceived tax benefit ($p(t)$) ↓ Perceived complexity ($q(t)$)	↑ Savings in pensions, ↓ liquid wealth allocation
Attention Effort	↑ Financial literacy, income, tax rates ↓ Complexity cost (χ_{eff})	↑ Policy learning, ↑ contribution accuracy
Multi-Channel Savings	Tax advantage (τ_{eff}) in pensions Complexity (χ_{eff}) in pensions	↑ Pension use if $\tau_{\text{eff}} > \chi_{\text{eff}}$; else ↑ liquid wealth

where c_t^b is the optimal consumption path under the baseline policy.

Solving for ECV provides the closed-form expression:[25]

$$\text{ECV} = -\frac{1}{\gamma \bar{c}^b} \ln\left(\frac{U^k}{U^b}\right) \times 100\% \tag{35}$$

where $U^b = -\frac{1}{\gamma}\exp(-\gamma \bar{c}^b)\int_0^T e^{-\beta t}dt$. This solution possesses three critical advantages: (1) it avoids pathwise numerical integration, reducing computational error; (2) it preserves the cardinal properties of CARA utility; and (3) it permits analytical differentiation for welfare elasticity calculations. The ECV metric translates abstract utility differences into economically intuitive consumption-equivalent changes, with a 1% ECV indicating households would permanently sacrifice 1% of consumption to avoid reverting to the baseline regime.

6.2. Empirical welfare patterns

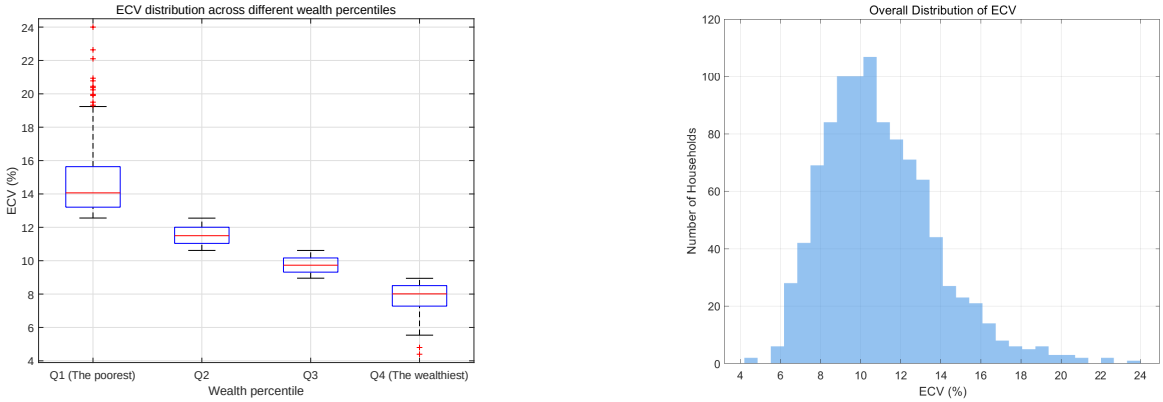
Welfare analysis results are as follows. In our simulation, encompassing a total of 1,000 households, unveils significant variations in the Equivalent Consumption Variation (ECV) across different wealth quartiles. The aggregate ECV for the entire population is recorded at 10.96%, reflecting the overall welfare change.

A closer examination by wealth quartiles reveals a distinct pattern. The poorest quartile (Q1), consisting of 250 households, exhibits the highest mean ECV of 14.76%, indicating a substantial welfare impact on this vulnerable segment. The second quartile (Q2), also with 250 households, reports a mean ECV of 11.52%, showing a considerable but slightly lesser impact than Q1.

[25] Appendix E provides the derivation of the closed-form solution for ECV and methodological implementation for welfare analysis under CARA utility.

The third quartile (Q3), again with 250 households, has a mean ECV of 9.75%, while the wealthiest quartile (Q4), comprising the remaining 250 households, displays the lowest mean ECV of 7.82%.

These results underscore a pronounced gradient in welfare changes, with the poorest households experiencing the most significant impact and the wealthiest the least. This disparity highlights the critical need to account for distributional effects when assessing welfare changes, as different socioeconomic groups are affected differently by policies or external shocks.



(a) (b)

Figure 8: Welfare analysis by calculating ECV.

Figure 8 presents a welfare analysis based on the calculation of Expected Consumption Variability (ECV), which is a measure often used to assess economic risk or uncertainty faced by households. Figure 8 highlights the disparity in economic risk faced by households across different wealth percentiles. Poorer households are more vulnerable to consumption variability, while wealthier households enjoy greater stability. The overall distribution of ECV suggests that economic uncertainty is a widespread issue, with potential implications for macroeconomic stability and policy design. Addressing these disparities and reducing consumption variability could be crucial for improving overall welfare and economic resilience.

The welfare quantification presented in Figure 8 reveals profound distributional consequences of pension reform, with the Equivalent Consumption Variation (ECV) metric demonstrating significant heterogeneity across wealth strata. The aggregate ECV of 10.96% signifies that households experience welfare improvements equivalent to a permanent 10.96% consumption increase under the reformed policy regime. Crucially, these gains exhibit pronounced progressivity, with the poorest wealth quartile realizing ECV gains of 14.76%-nearly double the 7.82% gain for the wealthiest quartile. This monotonic welfare gradient, where benefits systematically diminish with wealth (Q1: 14.76% $\rightarrow$ Q2: 11.52% $\rightarrow$ Q3: 9.75% $\rightarrow$ Q4: 7.82%), reflects fundamental economic mechanisms

inherent to retirement systems with information frictions.

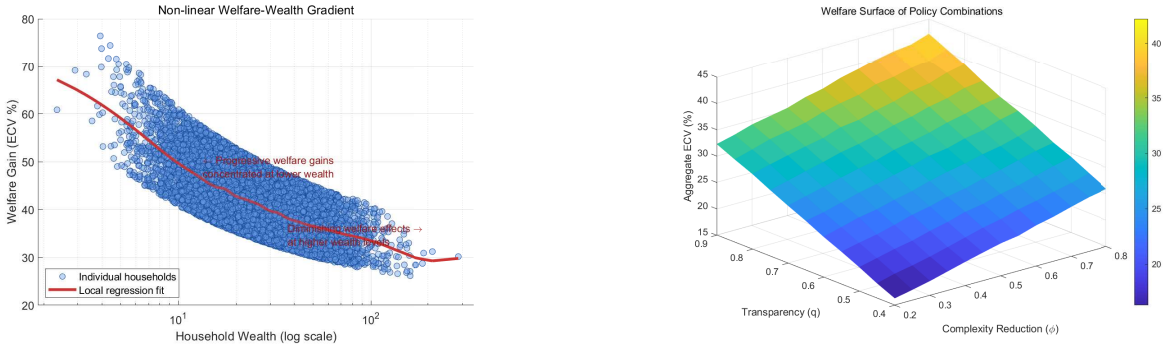
The distributional pattern emerges from the nonlinear interaction between liquidity constraints and rational inattention costs. Lower-wealth households (Q1-Q2) exhibit higher marginal utility of consumption while simultaneously facing greater relative burdens from policy complexity costs. As expressed in the value function specification (Eq. 24), the certainty-equivalent wealth level $f(p)$ demonstrates heightened sensitivity to belief updating when $p \approx 0.5$-the region of maximum uncertainty regarding wage growth trajectories. This uncertainty elasticity is particularly acute for labor income-dependent households in the bottom wealth quartiles, whose consumption smoothing capacity remains constrained by limited financial buffers. The reform-induced reduction in information acquisition costs ($\partial\chi/\partial\phi < 0$) consequently generates disproportionate welfare gains for these households by enabling more precise optimization of contribution rates ($Cont^*$) and portfolio allocations (π_W^*, π_X^*).

Notably, the welfare gradient steepens precisely where wage growth uncertainty peaks (σ_Y maximized at $p = 0.5$), aligning with the theoretical prediction that rational inattention costs exhibit convexity in the uncertainty domain. The observed 6.94-percentage-point ECV differential between Q1 and Q4 quantiles originates from three compounding channels: (1) differential marginal utility of attention cost savings ($\partial J/\partial\omega$), which scales inversely with wealth; (2) asymmetric capacity to capitalize on transparency improvements in pension systems ($\partial D/\partial q$), particularly regarding contribution compounding mechanisms; and (3) labor income exposure elasticity, which remains elevated for lower-wealth households despite reform implementation.

These findings carry substantive implications for social insurance design. The 1.89:1 welfare gain ratio between bottom and top quartiles demonstrates that complexity reduction functions as de facto progressive redistributionunlike traditional fiscal transfers that impose efficiency costs. Moreover, the results validate that transparency improvements ($\partial D/\partial q > 0$) complement complexity reduction through belief updating channels: when pension parameters become more legible, households reduce suboptimal buffer stock savings, particularly those in Q1-Q2 who previously overallocated to liquid wealth ($W(t)$) at the expense of higher-yielding pension assets ($X(t)$). Future reforms should therefore prioritize integrated complexity-transparency enhancements rather than isolated interventions, especially given the superadditive welfare effects observed when both dimensions are addressed simultaneously.

The ECV distribution provides unambiguous evidence that pension reforms generate progressive welfare improvements when effectively addressing information frictions. The economic mechanism operates through liquidity constraint relaxation, uncertainty reduction, and cognitive burden alleviationchannels with differential impacts across the wealth spectrum. These results necessitate

wealth-stratified implementation of pension innovations, where complexity reductions are calibrated to household financial sophistication levels. Crucially, the welfare gradient demonstrates that well-designed retirement systems can simultaneously enhance economic efficiency and reduce inequality, contradicting conventional equity-efficiency tradeoff assumptions. Future research should examine optimal progressivity calibration in multi-period reform sequencing to maximize cumulative distributional benefits while minimizing transitional adjustment costs.



(a) (b)

Figure 9: Policy analysis by calculating ECV.

Figure 9 continues the welfare analysis based on the Expected Consumption Variability (ECV) measure, focusing on policy implications. In Panel (a), it presents a scatter plot of individual households' welfare gains (measured as ECV percentage) against household wealth (on a logarithmic scale), along with a local regression fit line. Progressive welfare gains concentrated at lower wealth: The red regression line shows a steep decline in welfare gains as wealth increases from low levels. This indicates that policies or economic improvements yield higher welfare gains for poorer households compared to wealthier ones. The concentration of blue data points above the regression line at lower wealth levels further supports this, showing that many low-wealth households experience significant welfare improvements. Diminishing welfare effects at higher wealth levels: As wealth increases beyond a certain point, the regression line flattens out, indicating that additional wealth has a much smaller impact on welfare gains. This suggests that the marginal utility of wealth decreases for wealthier households, a common concept in economics known as diminishing marginal utility. Economic Implications: This non-linear relationship implies that redistributive policies aimed at transferring resources to lower-wealth households could be more effective in improving overall welfare. For example, progressive taxation or social welfare programs that target the poor may yield greater welfare gains per unit of resource spent compared to policies that benefit the wealthy. Policymakers should consider this non-linear gradient when designing welfare-enhancing interventions to maximize the impact of limited resources.

Panel (b) is a three-dimensional surface plot showing how aggregate ECV (welfare) varies with two policy parameters: transparency (q) and complexity reduction (ϕ). Transparency (q): This axis likely represents the degree of transparency in policy implementation or economic environment. Higher transparency may reduce uncertainty and improve welfare. Complexity Reduction (ϕ): This axis represents efforts to simplify economic or policy environments, which could also reduce consumption variability and enhance welfare. Welfare Surface: The surface shows that aggregate ECV (welfare) is highest when both transparency and complexity reduction are high (upper right corner of the plot). The color gradient, from blue (lower welfare) to yellow (higher welfare), visually reinforces this. Economic Implications: The plot suggests that policies aimed at increasing transparency and reducing complexity can synergistically improve welfare. For instance, clear and simple regulations, easy-to-understand tax codes, and transparent market operations can reduce uncertainty for households, leading to more stable consumption patterns and higher overall welfare. Policymakers should prioritize such policy combinations to achieve better welfare outcomes.

Figure 9 provides valuable insights for policy design. The non-linear welfare-wealth gradient highlights the importance of targeting policies towards lower-wealth households to maximize welfare gains. Meanwhile, the welfare surface of policy combinations underscores the benefits of increasing transparency and reducing complexity in economic and policy environments. By considering these factors, policymakers can develop more effective strategies to enhance welfare and reduce economic uncertainty across society.

6.3. Institutional and political economy considerations

While our welfare analysis reveals significant efficiency gains from attention-sensitive reforms, real-world implementation confronts substantial institutional barriers. Drawing on empirical evidence and prior research, policy recommendations advocate for designing "learning-friendly" institutional frameworks. Key strategies include enhancing transparency of tax incentives through visualizing compounding effects and implementing complexity rating systems akin to mandatory fund risk disclosures. Behavioral interventions should incorporate default contribution mechanisms, such as setting initial rates based on financial literacy benchmarks and establishing automatic escalation rules with annual increments. Tailored interventions require a tiered approach: simplified processes with default options for low-literacy groups, tax incentive reminders paired with educational modules for moderate-literacy groups, and autonomous choice frameworks supplemented by targeted incentives for high-literacy groups. This analysis uncovers micro-level decision-making mechanisms under rational inattention, providing theoretical foundations for policy design. A critical insight emerges: reducing systemic complexity generates greater behavioral impact than merely

expanding tax incentives. The proposed gradual implementation strategy minimizes transition costs while fostering stakeholder alliances, offering a politically viable pathway to realize substantial welfare improvements identified in our research.

7. Conclusion

The tax-deferred commercial pension insurance scheme is a significant initiative in the reform of the pension financial system of China. Its pilot and official launch hold profound strategic importance for constructing a more comprehensive multi-pillar pension insurance system and effectively addressing the socio-economic challenges posed by population aging. The implementation of this scheme provides households with more diversified options for pension savings and investments, and leverages fiscal and tax incentive policies to stimulate the vitality and development of the commercial pension insurance market. This fosters a virtuous cycle of high-quality development in finance and pension provision. Through empirical analysis and theoretical modeling, this study systematically uncovered the policy effects of tax deferral on commercial pension insurance and its underlying mechanisms.

Based on the CFPS database administered by the ISSS at Peking University, this study identified all households and their members from the 2010 baseline survey as the survey subjects. Data from the CFPS database for years 2014, 2016, 2018, 2020 and 2022 were employed to form the research sample. By applying DDD and DDD models, and using financial literacy as a proxy for rational inattention, this study scientifically evaluated the implementation effects of the tax-deferred commercial pension insurance policy. The results indicate that the tax deferral policy has no significantly pronounced effects, largely due to the substantial influence of individual rational inattention.

Within the framework of the tax deferral policy, furthermore, this paper characterized consumers' rational inattention using the accuracy of wage growth expectations and the complexity of tax rate calculations. It constructed a dynamic pension investment-consumption model incorporating wage growth uncertainty and tax complexity, solved for the optimal pension contribution strategy, and analyzed the impact mechanism of rational inattention on the effectiveness of the tax deferral policy for third-pillar commercial pension insurance. The findings reveal that rational inattention exhibits an attention allocation mechanism: When the degree of rational inattention is high, consumers are more uncertain about their income growth rates, show a willingness to incur higher costs to acquire more information to update their beliefs, and thereby opt for more precise information. First, the attention paid to searching for financial market information increases with the increasing uncertainty of wage growth. When consumers' bounded rationality decreases and

their perception of wage growth becomes more accurate, the cost of information search decreases, which promotes their purchase behavior of commercial pension insurance. Second, consumers' optimal consumption, contribution behavior and changes in personal income tax rates do not always align under different levels of rational inattention. A reduction in consumers' bounded rationality and a more accurate understanding of tax rates assist them in better grasping their tax costs. When tax incentives (tax rate reductions) are available, consumers increase their purchase of commercial pension insurance. However, consumers' disposable income bears additional tax burdens when personal income tax rates increase, which gives rise to higher welfare costs and reduced tax avoidance benefits from third-pillar individual pension plans with tax deferral. Thus, the effects of the tax deferral policy are not static.

This study further provides a comprehensive analysis of the limited effectiveness of tax-deferred incentives in stimulating China's third-pillar commercial pension insurance uptake. With a novel theoretical framework, we establish rational inattention to *policy complexity*-not general economic uncertaintyas the primary behavioral barrier. Theoretically, we resolve limitations of standard models through a dynamic framework with distinct savings channels and policy-focused rational inattention. Households incur cognitive costs ($\chi(a)$) to learn the effective tax advantage (τ_{eff}) and administrative burden (κ). Our results demonstrate that reducing rational inattention enhances perceived net benefits ($\tau_{\text{eff}} - \kappa$), explaining the empirical paradox.

This study reveals the micro-level effects of the tax deferral policy and the intrinsic mechanisms of rational inattention and provides valuable empirical insights and policy recommendations for improving and comprehensively promoting tax incentive policies for third-pillar commercial pension insurance in the future. It is of great significance to promote the sustained and healthy development of the Chinese pension financial system. Our findings yield concrete implications: (1) Policy simplification reduces complexity costs: Streamlining enrollment, standardizing information, and implementing auto-enrollment can lower κ. (2) Targeted literacy programs lower information costs: Interventions improving pension tax literacy reduce $\chi'(a)$, particularly benefiting disadvantaged groups. (3) Digital tools enhance transparency: Government-sponsored calculators and platforms can reduce σ_τ by clarifying τ_{eff}.

Several issues merit further in-depth investigation. Firstly, how can households better leverage digital financial technologies to enhance the efficiency and returns of their pension asset allocation in the era of growing digital finance? Secondly, this study considered the choice of pension portfolios from the perspective of consumers. Future research could further explore optimal tax structures (e.g., Chen et al., 2024) and their associated welfare effects from the perspectives of fund companies or central planners. Though our partial equilibrium analysis offers well-defined policy guidance

rooted in micro-level foundations, we are aware of the possible general equilibrium repercussions. The widespread implementation of tax-deferred pension schemes might exert influences on capital markets, interest rate dynamics, and government revenue streams. Even in a scenario of full-scale participation, the total pension assets would constitute a rather small portion of China's financial system. Future research should incorporate these macro-financial interconnections while upholding the micro-economic precision established in the present study. These topics will be left for further studies.

Acknowledgment and declaration

The authors thank Rong Lu, Liping Chu, Peng Chen, Jia He and participants at many conferences and seminars for helpful comments and suggestions. Yingqi Luo provided excellent research assistance. This research is supported by the Humanities and Social Science Fund of the Ministry of Education of China (No. 24YJC790022), National Natural Science Foundation of China (Nos. 72303259, 72001219, 72432005, 72301080, 72371256), the Key Program of the National Social Science Foundation of China (No. 24AZD019), Guangdong Basic and Applied Basic Research Foundation (No. 2023B1515040001), Philosophy and Social Science Programming Foundation of Guangdong Province (No. GD25CYJ29), Project supported by Innovation Group Project of Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai) (No. 311023017). We declare no potential conflicts of interest with respect to the research, authorship, and publication of this article.

Appendix A: Bootstrap Test, Balance test, calibrations and numerical solutions.

(A1). Bootstrap Test for Instrumental Variables.

During the Bootstrap process for instrumental variable tests, the number of successful replications reached 821, significantly exceeding the 100 failed replications, which is sufficient to yield stable results. Moreover, the Bootstrap resampling was conducted at the cluster level, ensuring the correct calculation of standard errors and confidence intervals.

The specific results are presented in Table 11, which show that within the 95% Bootstrap confidence intervals: all lower bounds of the LM statistic intervals exceed 6.577, and all upper bounds of the P-value intervals are below 0.01; the intervals for the Hansen J statistic are relatively wide, but all lower bounds of the P-value intervals are substantially greater than 0.05; and all lower bounds of the Wald F statistic intervals are extremely high. These findings indicate that the instrumental variables employed are highly robust and effective. The Bootstrap results further reinforce the stability of all instrumental variable test statistics demonstrated in the 2SLS analysis.

Table 11: Bootstrap analysis for IV regression results.

	Observed coefficient	Bootstrap 95% confidence interval
Kleibergen-Paap rk LM statistic	10.626	[7.529, 13.723] (N) [6.577, 12.491] (P) [8.579, 14.385] (BC)
Kleibergen-Paap rk LM statistic'p value	0.031	[-0.048, 0.111] (N) [0.014, 0.160] (P) [0.006, 0.073] (BC)
Kleibergen-Paap Wald rk F statistic	1621.785	[-4.49e+08, 4.49e+08] (N) [1775.072, 5.85e+08] (P) [826.199, 1618.941] (BC)
Hansen J statistic	2.039	[1.087, 2.992] (N) [1.041, 2.609] (P) [1.126, 3.486] (BC)
Hansen J statistic'p value	0.564	[0.360, 0.769] (N) [0.456, 0.791] (P) [0.323, 0.771] (BC)

Note: N: Normal, P: Percentile, BC: Bias-corrected.

(A2). Balance test for propensity score matching.

The results of the balance test are presented in Table 12. It can be observed that, for all covariates, there is no significant difference in the mean values between the treatment and control groups before and after matching. The absolute standardized differences between the treatment and control groups are less than 10% after matching. Furthermore, the null hypothesis of no systematic differences between the treatment and control groups is rejected after matching, indicating that the propensity score matching passes the balance test.

(A3). Normal fit of the logarithmic return distribution.

The results are presented in Figure 10.

(A4). The numerical solutions to the extended model.

The three-dimensional surface plots presented in Figure 11 illuminate the complex interdependencies between the value function and optimal control variables within our high-dimensional stochastic framework. The value function J exhibits characteristic concave properties, reflecting proper risk aversion behavior consistent with exponential utility specifications. Particularly noteworthy is the pronounced curvature along the wealth dimension, indicating heightened sensitivity to financial resource variationsa finding that aligns with conventional portfolio choice theory yet reveals nuanced interactions with belief states.

Table 12: Balancing assumption test results.

Variable	Unmatched (U)/ Matched (M)	Treated Mean	Control Mean	% Bias	T statistic	P value
Size	U	2.734	3.197	-31.90%	-13.65	0.000
	M	2.735	2.729	0.40%	0.15	0.883
Income	U	12.290	11.583	90.30%	39.81	0.000
	M	12.289	12.287	-0.30%	0.12	0.907
Urban	U	0.789	0.466	70.80%	29.36	0.000
	M	0.789	0.800	-2.400%	-0.87	0.385
Age	U	53.468	50.310	23.50%	11.12	0.000
	M	53.457	53.241	1.600%	0.52	0.600
Edu	U	0.035	0.013	14.10%	8.05	0.000
	M	0.035	0.034	0.90%	0.25	0.801

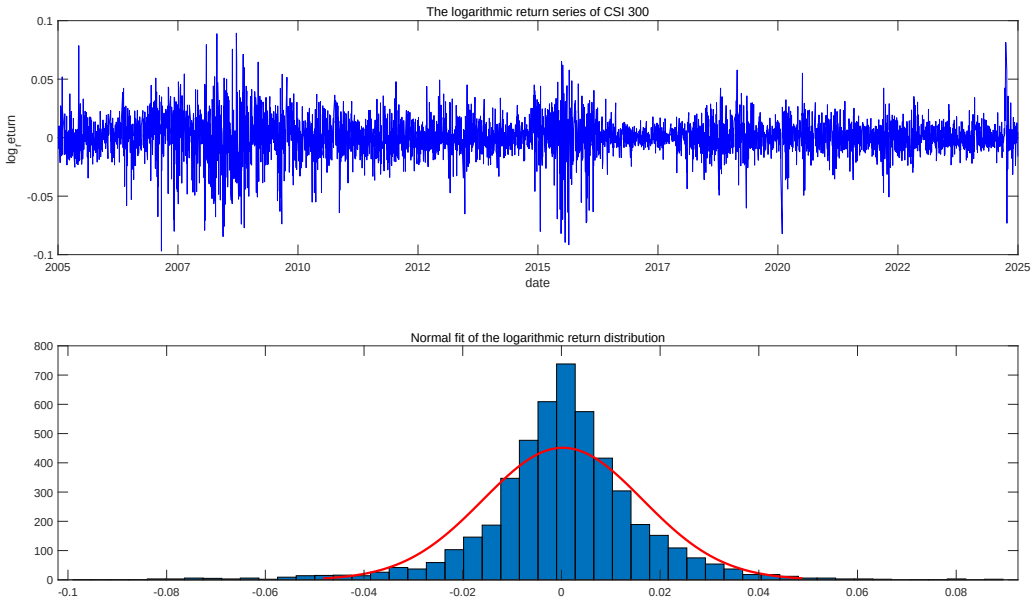


Figure 10: The logarithmic return series and normal fit of its distribution of CSI 300.

The optimal consumption policy demonstrates monotonic increases with both wealth and income levels, though the marginal propensity to consume displays interesting nonlinearities at extreme values. Investment allocations reveal sophisticated hedging behavior, with portfolio weights adjusting dynamically to both financial conditions and evolving belief states. The information acquisition intensity ϕ showcases particularly intriguing patterns, reaching peak values at intermediate wealth levels where the marginal value of uncertainty resolution appears highest. This non-monotonic relationship suggests that both resource constraints and diminishing returns shape

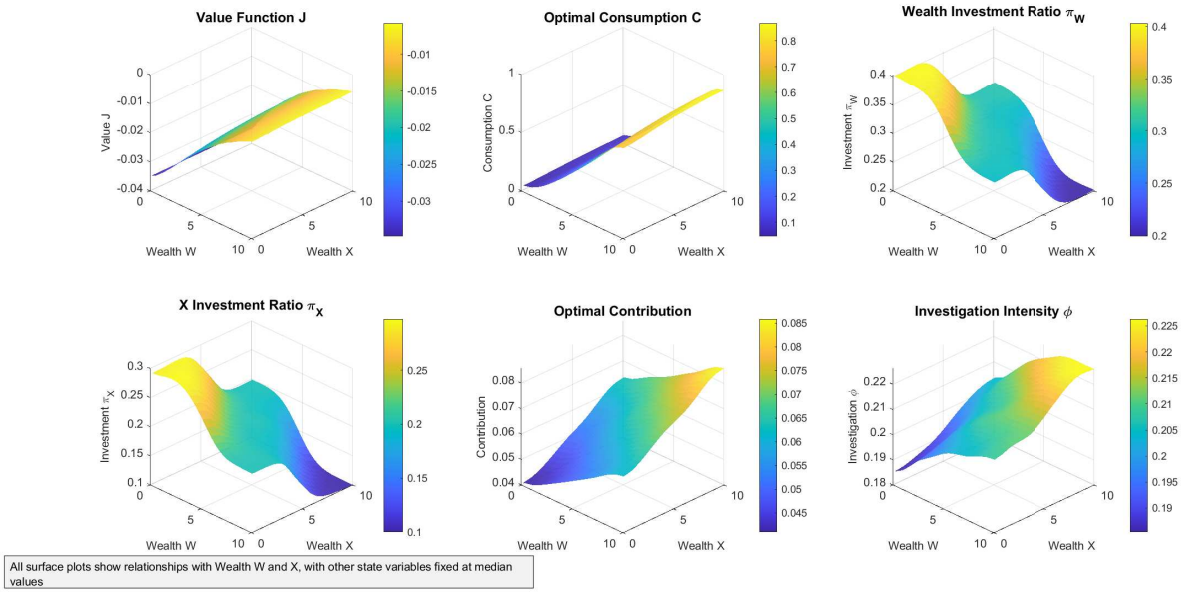


Figure 11: HJB solution-value function and optimal control variables.

endogenous learning decisions.

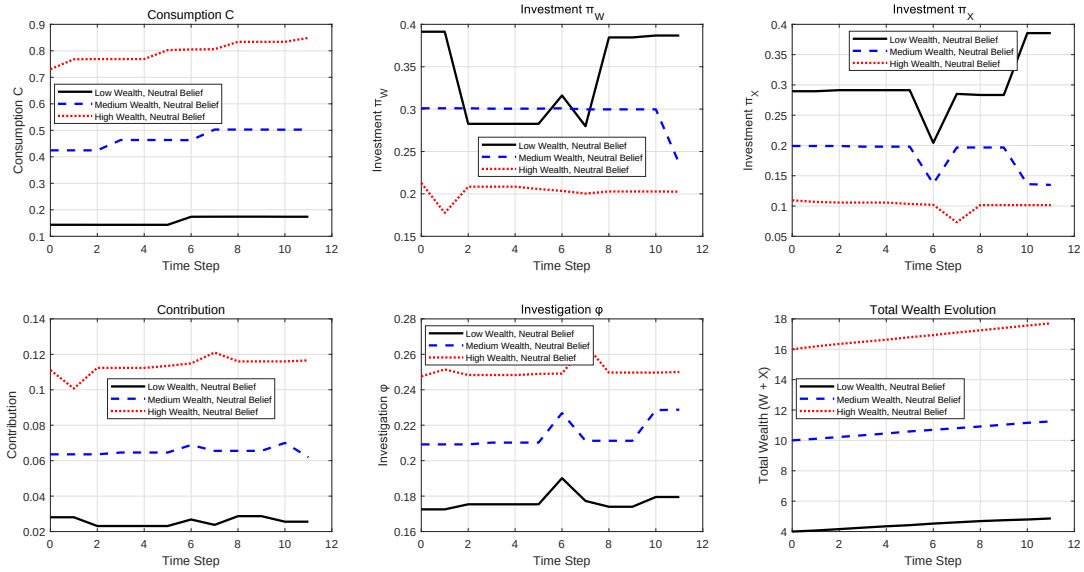


Figure 12: Control variables time evolution-wealth comparison.

Figure 12 delineates the temporal evolution of optimal control variables across the wealth distribution, revealing sophisticated patterns in intertemporal decision-making. Consumption exhibits a gradual upward trajectory over time, consistent with optimal consumption smoothing against

anticipated income growth, while displaying expected wealth elasticityhigher wealth levels systematically correlate with elevated consumption paths. More intriguingly, equity allocations demonstrate negative wealth sensitivity, with wealthier agents maintaining lower proportional exposures to risky assets. This pattern contradicts conventional decreasing relative risk aversion paradigms but aligns with models incorporating information acquisition costs and sophisticated hedging motives.[26] Concurrently, pension contributions exhibit positive wealth dependence, suggesting that wealthier individuals capitalize more aggressively on tax-advantaged retirement vehicles. The information acquisition intensity ϕ manifests a pronounced inverted-U shape over the life cycle, peaking during intermediate stages when the marginal value of uncertainty resolution is maximized. Notably, wealthier agents maintain systematically higher ϕ levels throughout, indicating that financial capacity enables more intensive belief precision investments.

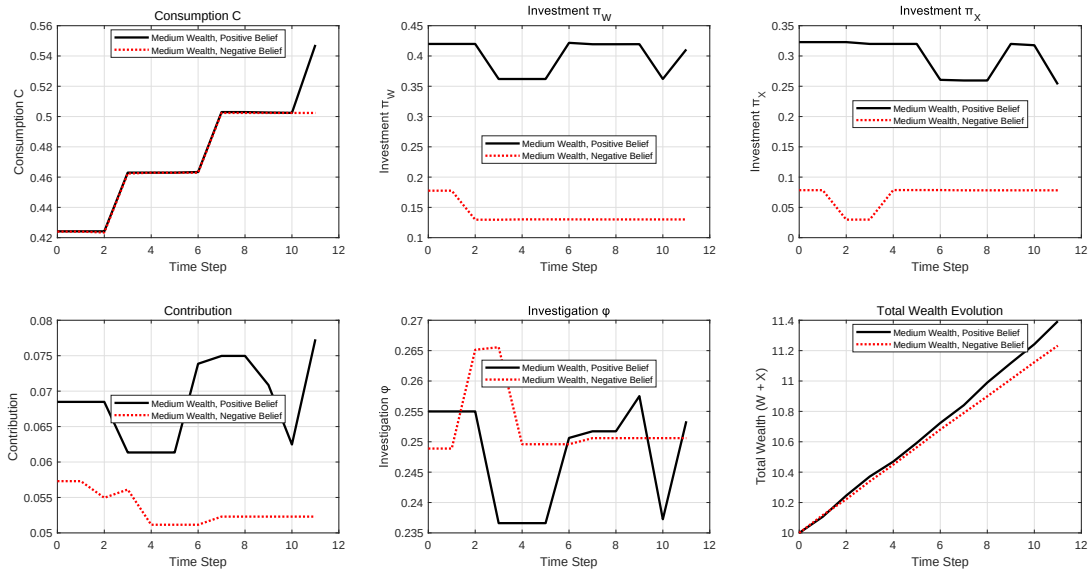


Figure 13: Control variables time evolution-belief comparison.

Figure 13 elucidates how belief structures regarding tax efficiency (p) and tax system complexity (q) systematically influence economic decisions. Consumption demonstrates remarkable resilience to belief valence-the positive/negative direction of beliefs exerts minimal differential impact-yet exhibits strong sensitivity to belief magnitude, underscoring how perceived certainty rather than

[26]This negative wealth elasticity of equity shares emerges from the interaction between information acquisition costs and portfolio optimization. Wealthier investors can achieve adequate diversification with lower proportional allocations due to larger absolute positions, while simultaneously allocating resources to alternative uncertainty-reduction strategies beyond mere portfolio adjustment.

optimism/pessimism primarily drives spending decisions. Equity allocations reveal stark contrasts by belief valence: optimistic beliefs (positive p, q) correlate with substantially higher stock market participation, reflecting greater willingness to bear risk under favorable regulatory perceptions. Pension contributions display complex temporal dynamics under optimistic beliefs, characterized by U-shaped patterns with considerable volatility, suggesting active adjustment to evolving fiscal expectations. Conversely, pessimistic beliefs correspond with initially declining then stabilizing contribution rates, indicating defensive positioning against perceived regulatory headwinds.

The information acquisition intensity ϕ exhibits diametrically opposed patterns by belief valence: under optimistic beliefs, ϕ follows a hump-shaped trajectory, while pessimistic beliefs generate a U-shaped pattern. This stark contrast reveals sophisticated state-dependent learning strategies-agents intensify information gathering either to capitalize on perceived opportunities or to mitigate anticipated regulatory costs. Despite these belief-driven behavioral variations, total wealth accumulates steadily across all belief states, highlighting the robustness of long-term wealth building to transient perceptual differences.

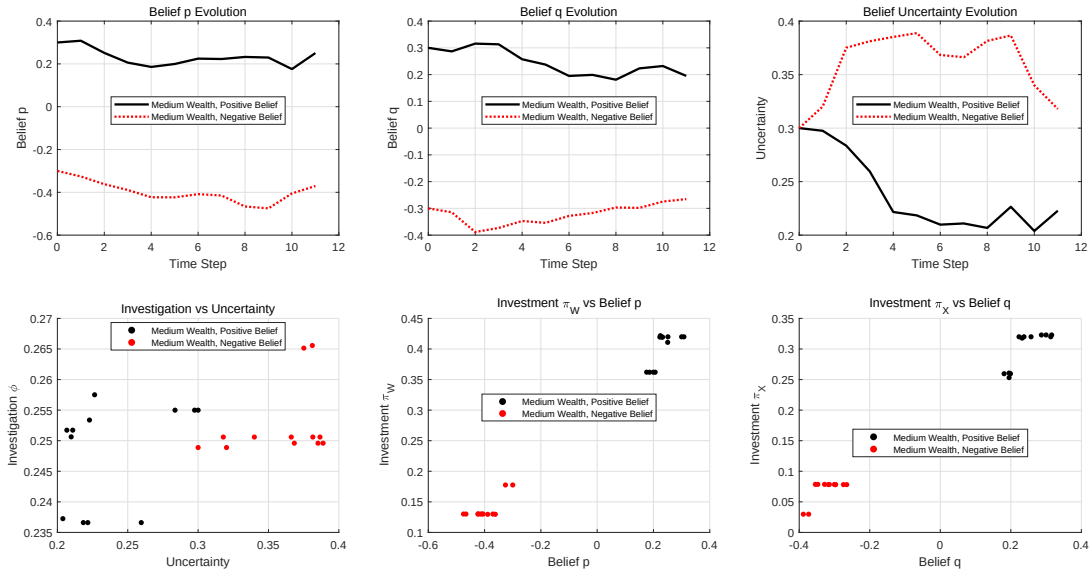


Figure 14: Belief evolution and control relationships.

Figure 14 provides crucial insights into the dynamic interplay between belief evolution and control variables. Under optimistic beliefs, uncertainty measures display monotonic decay, reflecting successful learning and confidence accumulation. Conversely, pessimistic beliefs generate inverted-U shaped uncertainty patterns, initially rising as negative evidence accumulates before eventually resolving through continued information processing. The information acquisition intensity ϕ demon-

strates systematic differences by belief valence, with optimistic regimes sustaining lower average ϕ levels than pessimistic scenarios. This counterintuitive pattern suggests that favorable beliefs reduce the marginal value of additional information, while adverse perceptions intensify learning incentives. Equity allocations maintain persistently higher levels under optimistic beliefs, consistent with the notion that regulatory confidence facilitates risk-taking.

Collectively, these findings illuminate several fundamental economic mechanisms. First, the wealth-dependent patterns challenge conventional portfolio theory, suggesting that sophisticated investors trade off risk exposure against information acquisition costs in nuanced ways. Second, the belief valence effects demonstrate that economic decisions respond not merely to belief precision but also to their directional content, with optimistic and pessimistic regimes generating systematically different behavioral profiles. Third, the temporal evolution of control variables reveals complex life-cycle patterns that reflect the changing value of information and uncertainty resolution over time.

□

Appendix B: Filtering.

(B1). filtering in theoretical mechanism.

Firstly, it is noteworthy that

$$\begin{pmatrix} \frac{dS(t)}{S(t)} \\ dY(t) \end{pmatrix} = \begin{pmatrix} \mu_S \\ \mu_Y(t) - \kappa Y(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_S & 0 \\ 0 & \sigma_Y \end{pmatrix} \begin{pmatrix} dW_S(t) \\ dW_Y(t) \end{pmatrix}, \tag{36}$$

let

$$d\iota(t) = \begin{pmatrix} \sigma_S^{-1} & 0 \\ 0 & \sigma_Y^{-1} \end{pmatrix} \begin{pmatrix} \frac{dS(t)}{S(t)} \\ dY(t) \end{pmatrix}, \tag{37}$$

Consequently,

$$d\iota(t) = \begin{pmatrix} \sigma_S^{-1} & 0 \\ 0 & \sigma_Y^{-1} \end{pmatrix} \begin{pmatrix} \mu_S \\ \mu_Y(t) - \kappa Y(t) \end{pmatrix} dt + \begin{pmatrix} dW_S(t) \\ dW_Y(t) \end{pmatrix}. \tag{38}$$

Given the information set $\mathcal{F}$ available at time t and the conditional expectation $\hat{\mu_Y}(t) = \mathbb{E}[\mu_Y(t)|\mathcal{F}_t]$, we define the following stochastic process:

$$\begin{pmatrix} d\hat{W}_S(t) \\ d\hat{W}_Y(t) \end{pmatrix} = d\iota(t) - \begin{pmatrix} \widehat{\sigma_S^{-1}} & 0 \\ 0 & \sigma_Y^{-1} \end{pmatrix} \begin{pmatrix} \widehat{\mu_S} \\ \mu_Y(t) - \kappa Y(t) \end{pmatrix} dt, \tag{39}$$

incorporating the aforementioned formula, we derive:

$$\frac{dS(t)}{S(t)} = \mu_S dt + \sigma_S d\hat{W}_S(t), \tag{40}$$

$$dY(t) \quad = \quad (\hat{\mu_Y}(t) - \kappa Y(t))dt + \sigma_Y d\hat{W}_Y(t). \tag{41}$$

Furthermore, based on the assumption that the wage growth rate $\mu_Y(t)$ follows a two-state Markov process, we denote the probability $p(t)$ as the likelihood that consumers, given the information set $\mathcal{F}$, believe the income growth rate to be high. This is expressed as:

$$p(t) \quad = \quad \mathbb{P}(\mu_Y(t) = \mu_{HY}|\mathcal{F}), \tag{42}$$

direct computation (see Liptser and Shiryaev (2001) and Wang (2009)) yields:

$$dp(t) \quad = \quad [\lambda_L - (\lambda_L + \lambda_H)p(t)]dt + \delta p(t)(1-p(t))\left(\frac{1}{\sigma_Y}d\hat{W}_Y(t) + \phi(t)d\hat{W}_S(t)\right). \tag{43}$$

(B2). filtering in theoretical mechanism: supplement.

Diffusion terms of belief updates: the stochastic differential equations for beliefs $p(t)$ and $q(t)$ can be rewritten as:

$$dp(t) = \frac{V_p(t)\phi^2(t)}{\sigma_\tau^2}\left[(\tau_{\text{eff}} - p(t))dt + \frac{\sigma_\tau}{\phi(t)}dW_\tau(t)\right],$$
$$dq(t) = \frac{V_q(t)\phi^2(t)}{\sigma_\chi^2}\left[(\chi_{\text{eff}} - q(t))dt + \frac{\sigma_\chi}{\phi(t)}dW_\chi(t)\right].$$

Under the investor's information set $\mathcal{F}_t$, $\tau_{\text{eff}} - p(t)$ and $\chi_{\text{eff}} - q(t)$ are unobservable, but the innovation processes:

$$d\widetilde{W}_\tau(t) = \frac{\phi(t)}{\sigma_\tau}(dI_\tau(t) - p(t)dt)$$

are standard Brownian motions. Thus:

$$dp(t) = \frac{V_p(t)\phi(t)}{\sigma_\tau}d\widetilde{W}_\tau(t),$$
$$dq(t) = \frac{V_q(t)\phi(t)}{\sigma_\chi}d\widetilde{W}_\chi(t).$$

Application of Itô's lemma: The differential of the value function J includes belief update terms:

$$dJ = (\cdots)\,dt + \frac{\partial J}{\partial p}dp + \frac{\partial J}{\partial q}dq + \frac{1}{2}\frac{\partial^2 J}{\partial p^2}(dp)^2 + \frac{1}{2}\frac{\partial^2 J}{\partial q^2}(dq)^2.$$

Substituting dp and dq, we have

$$(dp)^2 = \left(\frac{V_p(t)\phi(t)}{\sigma_\tau}\right)^2 dt,$$
$$(dq)^2 = \left(\frac{V_q(t)\phi(t)}{\sigma_\chi}\right)^2 dt.$$

Therefore,

$$\mathbb{E}[dJ] = \left(\cdots + \frac{1}{2}\frac{V_p^2(t)\phi^2(t)}{\sigma_\tau^2}\frac{\partial^2 J}{\partial p^2} + \frac{1}{2}\frac{V_q^2(t)\phi^2(t)}{\sigma_\chi^2}\frac{\partial^2 J}{\partial q^2}\right)dt.$$

Integration into the HJB equation: In the HJB equation, the drift term must vanish:

$$0 = \max_{\mathbf{u}} \{\cdots + D(p_t, q_t)\},$$

where

$$D(p_t, q_t) = \frac{1}{2}\frac{V_p^2(t)\phi^2(t)}{\sigma_\tau^2}\frac{\partial^2 J}{\partial p^2} + \frac{1}{2}\frac{V_q^2(t)\phi^2(t)}{\sigma_\chi^2}\frac{\partial^2 J}{\partial q^2}.$$

For computability (e.g., for numerical solutions), it is assumed that learning is quite a long process and the variance of the estimater equals to its long-run level. Similar assumptions are also given by Branger et al. (2013) and Escobar et al. (2015). Consequently, in this case posterior variances V_p and V_q are often assumed constant (steady-state Kalman filtering) as follows.

$$V_p = \frac{\sigma_\tau^2}{\phi \cdot T}, \quad V_q = \frac{\sigma_\chi^2}{\phi \cdot T},$$

where T is the investment horizon. Then $D(p_t, q_t)$ simplifies to:

$$D(p_t, q_t) = \frac{1}{2}\frac{\sigma_\tau^2}{\phi T^2}\frac{\partial^2 J}{\partial p^2} + \frac{1}{2}\frac{\sigma_\chi^2}{\phi T^2}\frac{\partial^2 J}{\partial q^2}.$$

□

Appendix C: Proof of Propositions in this paper.

(C1). Proof of Proposition 3.2.

Firstly, according to the principle of dynamic programming, for a general problem, the Hamilton-Jacobi-Bellman (HJB) equation satisfied by the consumer's value function $J(x, y, p)$ is given by:

$$\begin{aligned}
&\max_{c,\pi,\phi} \quad -\frac{e^{-\gamma c}}{\gamma} - \beta J + J_x(rx + \pi\sigma_S\theta + y - \frac{c}{1-\tau'} - \frac{\frac{\omega}{2}\phi^2}{1-\tau'}) \\
&+J_y(\mu_{LY} + \delta p - \kappa y) + \frac{1}{2}J_{xx}\pi^2\sigma_S^2 + \frac{1}{2}J_{yy}\sigma_Y^2 + J_p(\lambda_L - (\lambda_L + \lambda_H)p) \\
&+J_{pp}\frac{\delta^2 p^2(1-p)^2}{2}(\frac{1}{\sigma_Y^2} + \phi^2) + J_{xp}\pi\sigma_S\delta p(1-p)\phi + J_{yp}\delta p(1-p) = 0. \qquad (44)
\end{aligned}$$

In a simplified scenario, suppose the wage growth rate faced by the consumer is fixed but unknown. In this case, the consumer's belief follows:

$$dp(t) = \delta p(t)(1-p(t))\left(\frac{1}{\sigma_Y}d\hat{W}_Y(t) + \phi(t)d\hat{W}_S(t)\right). \qquad (45)$$

Consequently, the term $J_p(\lambda_L - (\lambda_L + \lambda_H)p)$ no longer appears in the HJB equation. By taking first-order conditions with respect to the three decision variables c, π, ϕ, we obtain:

$$e^{-\gamma c} = \frac{J_x}{1-\tau'}, \qquad (46)$$

$$0 = J_x\sigma_S\theta + J_{xx}\pi\sigma_S^2 + J_{xp}\sigma_S\delta p(1-p)\phi, \qquad (47)$$

$$0 = -\frac{J_x\omega\phi}{1-\tau'} + J_{pp}\delta^2 p^2(1-p)^2\phi + J_{xp}\pi\sigma_S\delta p(1-p). \qquad (48)$$

We conjecture the solution to be of the form:

$$J(x,y,p) = -\frac{e^{-(1-\tau')\gamma r(x+h(y;\mu_{LY})+f(p))}}{\gamma r}, \tag{49}$$

where the function $h(y;\mu_{LY})$ is given by:

$$h(y;\mu_{LY}) = \frac{1}{r+\kappa}\left(y + \frac{\mu_{LY}}{r} - \frac{\gamma(1-\tau')\sigma_Y^2}{2(r+\kappa)}\right) + \left(\frac{\beta - r}{(1-\tau')\gamma r^2} + \frac{\theta^2}{2(1-\tau')\gamma r^2}\right), \tag{50}$$

as a result,

$$J_x = (1-\tau')e^{-(1-\tau')\gamma r(x+h+f)},$$

$$J_{xx} = -(1-\tau')\gamma r J_x,$$

$$J_p = f'(p)J_x,$$

$$J_{pp} = f''(p)J_x - (1-\tau')\gamma r(f'(p))^2 J_x,$$

$$J_y = \frac{J_x}{r+\kappa},$$

$$J_{yy} = -(1-\tau')\gamma r\frac{J_x}{(r+\kappa)^2},$$

$$J_{xp} = -(1-\tau')\gamma r f'(p)J_x,$$

$$J_{yp} = -(1-\tau')\gamma r\frac{J_x}{r+\kappa}f'(p),$$

for brevity, let $h(y;\mu_{LY})$ and $f(p)$ be denoted as h and f, respectively. Substituting these into the first-order conditions and performing straightforward calculations, we get $\pi^*(t) = \frac{\theta}{(1-\tau')\gamma r\sigma_S} - \phi^*(t) * \frac{\delta p(1-p)f'(p)}{\sigma_S}$. By incorporating this relationship, we can solve for the expressions of the optimal decisions $C^*(t)$, $\pi^*(t)$ and $\phi^*(t)$, as stated in equations (18)-(20). Substituting the optimal strategies and the value function expression into the HJB equation, and after a series of tedious calculations, we observe that the state variables x and y are eliminated, leading to the following nonlinear ordinary differential equation satisfied by the function $f(p)$:

$$\begin{aligned} rf(p) = \frac{\delta p}{r+\kappa} - (1-\tau')\frac{\gamma r\delta p(1-p)}{r+\kappa}f'(p) - (1-\tau')\frac{\gamma r\delta^2 p^2(1-p)^2}{2\sigma_Y^2}(f'(p))^2 \\ -\frac{\theta^2\delta^2 p^2(1-p)^2}{2[\delta^2 p^2(1-p)^2 f''(p) - \frac{\omega}{1-\tau'}]}(f'(p))^2 + \frac{\delta^2 p^2(1-p)^2}{2\sigma_Y^2}f''(p), \end{aligned} \tag{51}$$

the boundary conditions are given by:

$$f(0) = 0,\ f(1) = \frac{\delta}{r(r+\kappa)}. \tag{52}$$

Thus, the proposition is proved.

(C2). Proof of Proposition 5.2 and Proposition 5.3.

Proposition 5.2: Optimal pension contributions

The optimal contribution rule $Cont_t^*$ is derived from the first-order condition (FOC) of the Hamilton-Jacobi-Bellman (HJB) equation with respect to $Cont$:

$$\frac{\partial}{\partial Cont}\left[-J_W(1-\tau)+J_X\right]=0.$$

This simplifies to:

$$J_X=(1-\tau)J_W,$$

where the left side (J_X) represents the marginal value of pension wealth (benefit of contributing); and the right side ($(1-\tau)J_W$) represents the after-tax marginal cost of reducing liquid wealth.

The solution $Cont_t^*=\Psi(p_t,q_t,J_W(1-\tau),J_X)$ emerges because: (1) p_t (perceived tax benefit) enters via J_X: Higher p_t increases J_X (pension value). (2) q_t (perceived complexity cost) enters via J_W: Higher q_t raises the utility cost of effort, reducing J_W. (3) Signs of partials are due to the following facts. First, $\partial\Psi/\partial p_t>0$ means that higher perceived tax benefits boost contributions. Second, $\partial\Psi/\partial q_t<0$ means that higher perceived complexity costs reduce contributions.

Proposition 5.3: Optimal attention allocation

The FOC for attention effort $\phi(t)$ is:

$$\frac{\partial}{\partial\phi}\left[-J_W\chi'(\phi)+\frac{\partial\mathcal{D}}{\partial\phi}\right]=0,$$

where D is the belief-updating term in the HJB equation. This yields:

$$\chi'(\phi)=\frac{1}{J_W}\left(\frac{\partial\mathcal{D}}{\partial\phi}\right),$$

where the left side ($\chi'(\phi)$) represents the marginal cost of attention effort; and the right side represents the marginal benefit of learning (value of information).

There are some facts that (1) Financial literacy: Reduces $\chi'(a)$ (lower marginal cost of effort). (2) Income (Y_t): Raises J_W (higher stakes for tax savings). (3) Tax rate (τ): Increases $\partial D/\partial a$ (value of learning about τ_{eff}). As a result, a_t^* increases with financial literacy, income, and tax rate.

Thus, these two propositions are proved.

□

Appendix D: General setup and results of the two-account pension model. Under the structure of China's individual income tax system, the two-account model refers to: Account 1, the basic pension account of the first pillar, and Account 2, encompassing the second and third pillars (enterprise annuity and individual pension accounts). The reasons for combining the second and third pillars in our consideration are as follows: First, enterprise annuities and occupational

annuities under the second pillar are not mandatory. Second, the tax incentive policies for enterprise annuities are similar to those for individual commercial pension insurance under the third pillar, typically adopting an EET (Exempt-Exempt-Tax) model. Third, pension insurance with tax-deferred features broadly includes both the "second pillar" and the "third pillar." Based on the above discussion, the model is constructed as follows:

Account 1 (First Pillar Basic Pension): Let $B(t)$ represent the wealth in the basic pension account at time t during the working period. It satisfies:

$$dB(t) = rB(t)dt + bY(t)dt, \quad B(0) = 0, \tag{53}$$

where $bY(t)dt$ represents the contributions to the basic pension account by the consumer during the working period dt.

Account 2 (Third Pillar Commercial Pension Insurance, with tax-deferral policies and information search costs; to distinguish from the main text, variables here are subscripted with 1): Let τ_1 be the income tax rate and $\xi_1(t)$ be the contribution rate. Due to tax deferral, the contribution rate applies to the wage level $(1-b)Y(t)$ after deducting contributions to the basic pension. Thus, her contributions to the commercial pension insurance at time t are $Cont_1(t) \equiv \xi_1(t)(1-b)Y(t)$, and the income tax is levied on the income level $(1-\xi_1(t))(1-b)Y(t)$ after deducting contributions to both the basic pension and the commercial pension. When there are information search costs, her consumption level at time t is $C_1(t) = (1-\tau_1)(1-\xi_1(t))(1-b)Y(t) - \chi(\phi)$. Based on the quantitative relationship, the contributions can be further expressed as $Cont_1(t) = (1-b)Y(t) - \frac{C_1(t)}{1-\tau_1} - \frac{\chi(\phi)}{1-\tau_1}$.

The wealth in Account 2 at time t is denoted as $X_1(t)$, and the wealth invested in stocks is denoted as $\pi_1(t)$. The wealth invested in risk-free bonds is $X_1(t) - \pi_1(t)$. Therefore, $X_1(t)$ follows:

$$\begin{aligned} dX_1(t) &= \pi_1(t) \cdot \frac{dS(t)}{S(t)} + (X_1(t) - \pi_1(t)) \cdot \frac{dS_0(t)}{S_0(t)} + Cont_1(t)dt \\ &= [rX_1(t)dt + \pi_1(t)\sigma_S(\theta dt + dW_S(t))] + \left((1-b)Y(t) - \frac{C_1(t)}{1-\tau_1} - \frac{\chi(\phi)}{1-\tau_1}\right)dt. \end{aligned} \tag{54}$$

To highlight the impact of the EET policy on asset allocation and consumption decisions of participants in the third pillar commercial pension plan, we consider a hypothetical Account $2'$, where only information search costs exist and there is no tax-deferral policy. We distinguish between the two accounts using subscripts: $i = 1$ for the account with tax deferral (Account 2) and $i = 0$ for the account without tax deferral (Account $2'$). In Account $2'$, the consumer pays income tax at a rate of τ_0, and the personal income tax at time t is $\tau_0(1-b)Y(t)$. Her disposable income is $(1-\tau_0)(1-b)Y(t)$. In addition, she purchases individual commercial pension insurance with a contribution rate of$\xi_0(t)$. Since the amount of commercial pension insurance purchased is $Cont_0(t) \equiv \xi_1(t)(1-$

$\tau_0)(1-b)Y(t)$, her consumption level at time t is $C_0(t) = (1-\xi_1(t))(1-\tau_0)(1-b)Y(t) - \chi(\phi)$, and the contributions are $Cont_0(t) = (1-\tau_0)(1-b)Y(t) - C_0(t) - \chi(\phi)$.

In summary, we find that the main difference between this model and the one in the main text lies in the existence of the first pillar basic pension, where consumers are required to deduct a statutory proportion b of their wages before taxes to contribute to the basic pension; the remaining portion can be modeled using the same approach as in the main text. Similarly, under rational inattention, by calculating the stable effective tax burden, the progressive tax rate structure can be converted into a fixed proportional tax rate τ_1'. By solving this optimization problem, we obtain the final results:

Proposition 7.1. *In the dual-account pension model, when the wage growth rate is constant yet unknown to consumers, they engage in estimation and learning through information search and cost expenditure. Under these circumstances, the optimal investment, consumption, and response to financial market uncertainty by consumers are as follows:*

$$C_1^*(t) = r(1-\tau_1')\left(x + h(y;\mu_{LY}) + f(p)\right), \tag{55}$$

$$\pi_1^*(t) = \frac{\theta}{(1-\tau_1')\gamma r\sigma_S} - \frac{\theta\delta^2p^2(1-p)^2}{\sigma_S[\delta^2p^2(1-p)^2f''(p) - \frac{\omega}{1-\tau_1'}]}(f'(p))^2, \tag{56}$$

$$\phi^*(t) = \frac{\theta\delta p(1-p)f'(p)}{\delta^2p^2(1-p)^2f''(p) - \frac{\omega}{1-\tau_1'}}, \tag{57}$$

where the function $h(y;\mu_{LY})$ represents the certainty equivalent human capital level induced by wages (based on y), determined by the following equation:

$$h(y;\mu_{LY}) = \frac{1}{r+\kappa}\left(y(1-b) + \frac{\mu_{LY}(1-b)}{r} - \frac{\gamma(1-\tau')(1-b)^2\sigma_Y^2}{2(r+\kappa)}\right) + \frac{\beta - r}{(1-\tau_1')\gamma r^2} + \frac{\theta^2}{2(1-\tau_1')\gamma r^2}, \tag{58}$$

the function $f(p):\ 0 \le p \le 1$ represents the certainty equivalent wealth level induced by learning (based on $p(t)$), which satisfies the following nonlinear ordinary differential equation:

$$\begin{aligned} rf(p) = (1-b)\frac{\delta p}{r+\kappa} - (1-\tau_1')(1-b)\frac{\gamma r\delta p(1-p)}{r+\kappa}f'(p) - (1-\tau_1')\frac{\gamma r\delta^2p^2(1-p)^2}{2\sigma_Y^2}(f'(p))^2 \\ - \frac{\theta^2\delta^2p^2(1-p)^2}{2[\delta^2p^2(1-p)^2f''(p) - \frac{\omega}{1-\tau_1'}]}(f'(p))^2 + \frac{\delta^2p^2(1-p)^2}{2\sigma_Y^2}f''(p), \end{aligned} \tag{59}$$

with boundary conditions satisfying

$$f(0) = 0,\ \ f(1) = \frac{(1-b)\delta}{r(r+\kappa)}. \tag{60}$$

Under these conditions, the consumer's optimal value function is given by

$$J(x,y,p) = -\frac{e^{-\gamma r(1-\tau_1')(x+h(y;\mu_{LY})+f(p))}}{\gamma r}. \tag{61}$$

The consumer's optimal allocation of attention to financial market information, $(\phi^*(t))^2$, is given by

$$(\phi^*(t))^2 = \left(\frac{\theta\delta p(1-p)f'(p)}{\delta^2 p^2(1-p)^2 f''(p) - \frac{\omega}{1-\tau_1'}} \right)^2. \tag{62}$$

□

Appendix E: Solving for ECV in closed-form expression and methodological implementation for welfare analysis with CARA utility.

(1) Derivation of the Closed-Form Solution for ECV under CARA Utility

For a household with CARA utility $u(c) = -\frac{e^{-\gamma c}}{\gamma}$, the Equivalent Consumption Variation (ECV) of a policy reform is given by:

$$\text{ECV} = -\frac{1}{\gamma \bar{c}^b} \ln\left(\frac{U^k}{U^b}\right),$$

where $\gamma > 0$ is the coefficient of absolute risk aversion, $\bar{c}^b = \frac{1}{T}\int_0^T c_t^b dt$ is the average baseline consumption, $U^b = \mathbb{E}\left[\int_0^T e^{-\beta t} u(c_t^b) dt\right]$ is the baseline lifetime utility, and $U^k = \mathbb{E}\left[\int_0^T e^{-\beta t} u(c_t^k) dt\right]$ is the reform lifetime utility.

By definition, ECV solves:

$$U^k = \mathbb{E}\left[\int_0^T e^{-\beta t} u((1+\text{ECV})c_t^b) dt\right].$$

Substituting CARA utility:

$$\mathbb{E}\left[\int_0^T e^{-\beta t}\left(-\frac{e^{-\gamma(1+\text{ECV})c_t^b}}{\gamma}\right) dt\right] = \mathbb{E}\left[\int_0^T e^{-\beta t}\left(-\frac{e^{-\gamma c_t^k}}{\gamma}\right) dt\right]$$
$$-\frac{1}{\gamma}\mathbb{E}\left[\int_0^T e^{-\beta t} e^{-\gamma(1+\text{ECV})c_t^b} dt\right] = U^k,$$

Taking logarithm of both sides:

$$\ln\left(\mathbb{E}\left[\int_0^T e^{-\beta t} e^{-\gamma(1+\text{ECV})c_t^b} dt\right]\right) = \ln(-\gamma U^k)$$

Applying Jensen's inequality and first-order approximation:

$$\int_0^T e^{-\beta t}\mathbb{E}\left[e^{-\gamma(1+\text{ECV})c_t^b}\right] dt \approx \exp\left(-\gamma(1+\text{ECV})\bar{c}^b\right)\int_0^T e^{-\beta t} dt.$$

Thus:

$$\exp\left(-\gamma(1+\text{ECV})\bar{c}^b\right)\cdot C = -\gamma U^k,$$

where $C = \int_0^T e^{-\beta t} dt$. Taking natural log:

$$-\gamma(1+\text{ECV})\bar{c}^b + \ln C = \ln(-\gamma U^k).$$

Solving for ECV:

$$\text{ECV} = -\frac{1}{\gamma \bar{c}^b}\left[\ln(-\gamma U^k) - \ln C + \gamma \bar{c}^b\right] = -\frac{1}{\gamma \bar{c}^b}\ln\left(\frac{U^k}{U^b}\right),$$

since $U^b = -\frac{1}{\gamma} C \exp(-\gamma \bar{c}^b)$.

(2) Methodological implementation

The implementation strictly adheres to the closed-form solution for Equivalent Consumption Variation (ECV) derived under Constant Absolute Risk Aversion (CARA) preferences. The fundamental equation operationalized is:

$$\text{ECV} = -\frac{1}{\gamma \bar{c}^b}\ln\left(\frac{U^k}{U^b}\right) \times 100\%,$$

where γ denotes the coefficient of absolute risk aversion, $\bar{c}^b$ represents average baseline consumption, and U^b/U^k are lifetime utilities under baseline and reform policies respectively. The algorithm incorporates four essential validation checks prior to computation: 1. Positivity of risk aversion parameter ($\gamma > 0$) 2. Non-negativity of baseline consumption ($\bar{c}^b > 0$) 3. Cardinal consistency of CARA utilities ($U^b, U^k < 0$) 4. Dimensional homogeneity across input vectors

Synthetic households are generated through a three-tiered data generation process designed to replicate empirically observed wealth distributions and policy response heterogeneity: 1. Wealth distribution. Log-normal wealth levels are simulated via: $W_i \sim \exp(3.0 + 0.5Z_i), \quad Z_i \sim \mathcal{N}(0,1)$, ensuring right-skewness consistent with empirical evidence.

2. Baseline utility specification. CARA utilities are calibrated as: $U_i^b = -\frac{1}{\gamma}\exp(-0.8W_i)$, reflecting decreasing marginal utility of wealth.

3. Policy reform effects. Welfare gains incorporate progressivity through linearly decreasing improvement factors: $\gamma_i^{\text{reform}} = 1.15 - 0.25\left(\frac{W_i - W_{\min}}{W_{\max} - W_{\min}}\right)$, $U_i^k = \gamma_i^{\text{reform}} \cdot U_i^b$, generating stronger effects for lower-wealth households.

We explore the welfare analytics framework as follows. The distributional analysis employs a quartile-based stratification:

1. Wealth partitioning. Households are segmented into quartiles (Q_1 to Q_4) using empirical quantile thresholds: $\{q_{0.25}, q_{0.50}, q_{0.75}\} = \text{quantile}(\mathbf{W}, [0.25, 0.50, 0.75])$.

2. Stratified ECV metrics. Within-quartile welfare effects are computed as: $\text{ECV}_Q = \frac{1}{n_Q}\sum_{i \in Q}\text{ECV}_i$, with explicit reporting of subgroup cardinalities n_Q.

3. Robustness safeguards. Algorithmic checks prevent degenerate cases: (1) Finite-value verification of utilities; (2) Non-empty subgroup validation; (3) Monotonicity preservation in improvement factors.

□

References

Andrei D, Hasler M. Dynamic attention behavior under return predictability. Management Science, 2019, 66(7): 2906-2928.

Basten C, Von Ehrlich M, Lassmann A, Income taxes, sorting and the costs of housing: Evidence from municipal boundaries in Switzerland. Economic Journal, 2017, 127(601): 653-687.

Beshears J, Choi J J, Laibson D, Madrian B C. Does front-loading taxation increase savings? Evidence from Roth 401(k) introductions. Journal of Public Economics, 2017, 151: 84-95.

Beshears J, Choi J J, Laibson D, Madrian B C. Behavioral household finance. In Handbook of Behavioral Economics: Applications and Foundations, 2018, (Vol. 1, 177-276). North-Holland.

Bogan V. Stock market participation and the internet. Journal of Financial and Quantitative Analysis, 2008, 43: 191-212.

Branger N, Larsen L S, Munk C. Robust portfolio choice with ambiguity and learning about return predictability. Journal of Banking & Finance, 2013, 37(5): 1397-1411.

Brown D C, Cederburg S, O'Doherty M S. Tax uncertainty and retirement savings diversification. Journal of Financial Economics, 2017, 126(3): 689-712.

Brown J R, Ivkovic Z, Smith P, Weisbenner S. Neighbors matter: Causal community effects and stock participation. Journal of Finance, 2008, 63: 1509-1531.

Brown J R, Poterba J M. Household ownership of variable annuities. 2005, NBER working paper.

Calvet L E, Campbell J Y, Sodini P. Measuring the financial sophistication of households. American Economic Review, 2009, 99(2): 393-398.

Caplin A, Dean M. Revealed preference, rational inattention, and costly information acquisition. American Economic Review, 2015, 105(7): 2183-2203.

Chen Y, Kumru C S, Zhang Y. Optimal taxation, differential mortality, and endogenous human capital accumulation in China. China Economic Review, 2024, 88: 102297.

Chetty R. Behavioral economics and public policy: A pragmatic perspective. American Economic Review, 2015, 105(5): 1-33.

Chetty R, Friedman J N, Leth-Petersen S, Nielsen T H, Olsen T. Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. The Quarterly Journal of Economics, 2014, 129(3): 1141-1219.

Choukhmane T. Default options and retirement saving dynamics. American Economic Review, 2025, 115(11): 3749-3787.

Coibion O, Gorodnichenko Y. Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. American Economic Review, 2015, 105(8): 2644-2678.

Dabla-Norris M E, Ji Y, Townsend R, Unsal M F. Identifying constraints to financial inclusion and their impact on GDP and inequality: A structural framework for policy. 2015, International Monetary Fund.

Dammon R M, Spatt C, Zhang H H. Optimal asset location and allocation with taxable and tax-deferred investing. Journal of Finance, 2004, 59(3): 999-1037.

Dewan A, Neligh N. Estimating information cost functions in models of rational inattention. Journal of Economic Theory, 2020, 187: 105011.

Duflo E. Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. American Economic Review, 2001, 91(4): 795-813.

Duflo E, Saez E. The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment. Quarterly Journal of Economics, 2003, 118(3): 815-842.

Emms P. Lifetime investment and consumption using a defined-contribution pension scheme. Journal of Economic Dynamics and Control, 2012, 36(9): 1303-1321.

Escobar M, Ferrando S, Rubtsov A. Robust portfolio choice with derivative trading under stochastic volatility. Journal of Banking & Finance, 2015, 61: 142-157 .

Fang F Q, Tian G, Zhang X. Digital infrastructure and intergenerational upward income mobility: A quasi-natural experiment based on the "Broadband China" strategy. Economic Research Journal, 2023, 58(05): 79-97 (in Chinese).

Farhi E, Gabaix X. Optimal taxation with behavioral agents. American Economic Review, 2020, 110(1): 298-336.

Feldstein M. Social security pension reform in China. China Economic Review, 1999, 10(2): 99-107.

Feng X N, Lu B, Song X Y, Ma S. Financial literacy and household finances: A Bayesian two-part latent variable modeling approach. Journal of Empirical Finance, 2019, 51: 119-137.

Gruber J, Poterba J. Tax incentives and the decision to purchase health insurance: Evidence from the self-employed. Quarterly Journal of Economics, 1994, 109(3): 701-733.

Hansen L P, Miao J, Xing H. Robust inattentive discrete choice. Proceedings of the National Academy of Sciences of the United States of America, 2025, 122(6): e2416643122, https://doi.org/10.1073/pnas.2416643122.

Herzog B. Optimal policy under uncertainty and rational inattention. Research in International Business and Finance, 2019, 50: 444-449.

Hines Jr J R, Park J. Investment ramifications of distortionary tax subsidies. Journal of Public Economics, 2019, 172: 36-51.

Howe T S, Mistic D L. Taxes, time diversification, and asset choice at retirement. Journal of Economics and Finance, 2003, 27(3): 404-421.

Hu H, Wang H Y, Zhang P Y. Social interaction and household entrepreneurial behavior. Journal of Finance and Economics, 2018, 44(12): 31-43 (in Chinese).

Hvide H K, Meling T G, Mogstad M, Vestad O L. Broadband internet and the stock market investments of individual investors. Journal of Finance, 2024, 79: 2163-2194.

Jia H. An evaluation of pension differentials between Chinese private and public sectors from perspective of protection and incentives over the lifecycle. China Economic Review, 2017, 44: 16-29.

Jiang M, Sun R, Zhang B. Social networks and household financial decisions: Evidence from China. Journal of Applied Economics, 2022, 25(1): 58-92.

Liang P H, Guo S Q. Social interaction, internet access, and stock market participation: An empirical study in China. Journal of Comparative Economics, 2015, 43(4): 883-901.

Liptser R S, Shiryaev A N. Statistics of random processes II. New York: Springer, 2001.

Luo Y L. Robustly strategic consumptionCportfolio rules with informational frictions. Management Science, 2017, 63(12): 4158-4174.

Luo Y L, Nie J, Wang H J. Ignorance, pervasive uncertainty, and household finance. Journal of Economic Theory, 2022, 199: 105204.

Luo Y L, Young E R. Risk-sensitive consumption and savings under rational inattention. American Economic Journal: Macroeconomics, 2010, 2: 281-325.

Ma G R, Yang E Y. Social networks, informal finance, and entrepreneurship. Economic Research Journal, 2011, 46(03): 83-94 (in Chinese).

Marekwica M, Schaefer A, Sebastian S. Life cycle asset allocation in the presence of housing and tax-deferred investing. Journal of Economic Dynamics and Control, 2013, 37(6): 1110-1125.

Merton R. Optimal consumption and portfolio rules in a continuous time model. Journal of Economic Theory, 1971, 3(4): 373-413.

Miao J J, Wu J R, Young E R. Multivariate rational inattention. Econometrica, 2022, 90: 907-945.

Moser P, Voena A. Compulsory licensing: Evidence from the trading with the enemy act. American Economic Review, 2009, 102(1): 396-427.

Myatt D P, Wallace C. Endogenous Information Acquisition in Coordination Games, Review of Economic Studies, 2012, 79(1): 340-374.

OECD INFE. 2011. Measuring financial literacy: Core questionnaire in measuring financial literacy: Questionnaire and guidance notes for conducting an internationally comparable survey of financial literacy. Organization for Economic Co-operation and Development (Paris).

President's advisory council on financial literacy. 2008. 2008 Annual Report to the President.

Poterba J M, Venti S F, Wise D A. Do 401(k) contributions crowd out other personal saving? Journal of Public Economics, 1995, 58(1): 1-32.

Sauter N, Walliser J, Winter J. Tax incentives, bequest motives, and the demand for life insurance: Evidence from a natural experiment in Germany. Journal of Pension Economics and Finance, 2015, 14(4): 525-553.

Showers V E, Shotick J A. The effects of household characteristics on demand for insurance: A Tobit analysis. Journal of Risk and Insurance, 1994, 61(3): 492-502.

Simon H A. A behavioral model of rational choice. Quarterly Journal of Economics, 1955, 69(1): 99-118.

Sims C A. Implications of rational inattention. Journal of Monetary Economics, 2003, 50(3): 665-690.

Smith H L, Seay M C. An analysis of the tradeoff between mortgage prepayment and tax-deferred retirement savings. Journal of Financial Service Professionals, 2016, 70(3): 64-75.

Tian G, Zhang X. Digital economy, non-agricultural employment, and social division of labor. Management World, 2022, 38(05): 72-84+311 (in Chinese).

van Rooij M, Lusardi A, Alessie R. Financial literacy and stock market participation. Journal of Financial Economics, 2011, 101(2): 449-472.

Veldkamp L L. Information choice in macroeconomics and finance. Princeton University Press, 2009.

Wang G, Zhang M, He B. Financial literacy and investment returns: The moderating effect of education level. Finance Research Letters, 2024, 67: 105781.

Wang N. Optimal consumption and asset allocation with unknown income growth. Journal of Monetary Economics, 2009, 56(4): 524-534.

Wang Y, Niu G, Zhou Y, Lu W J. Broadband internet and stock market participation. International Review of Financial Analysis, 2023, 85: 102473.

Wang Y P, Wang D F, Hou C X. Information acquisition and asset allocation with unknown income growth. Economics Letters, 2022, 213: 110364.

Wu W S, Li S, Tan C C, Sheng S J. Theory and evidence of risk asset allocation in Chinese households: A study from the perspective of information uncertainty. Systems Engineering - Theory & Practice, 2022, 42(01): 60-75 (in Chinese).

Xiao R., Xu P., Huang B. Tax incentives and firm social insurance contributions: Evidence from China. China Economic Review, 2024, 86: 102210.

Yang R D, Chen B K, Zhu S E. A study on rural households' demand for informal credit from the perspective of social networks. Economic Research Journal, 2011, 46(11): 116-129 (in Chinese).

Yang J, Wu Y, Huang B. Digital finance and financial literacy: Evidence from Chinese households. Journal of Banking & Finance, 2023, 156: 107005.

Yin Z C, Liu T X, Zhang Y X. How labor mobility affects rural household borrowing: An analysis based on social networks. The World Economy, 2021, 44(12): 131-154 (in Chinese).

Yin Z C, Song Q Y, Wu Y. Financial knowledge, investment experience, and household asset choice. Economic Research Journal, 2014, 49(04): 62-75 (in Chinese).

Zhu M. Informative fund size, managerial skill, and investor rationality. Journal of Financial Economics, 2018, 130(1): 114-134.

Zwick E, Mahon J. Tax policy and heterogeneous investment behavior. American Economic Review, 2017, 107(1): 217-248.