How Good is Generative AI Personal Financial Advice?*

Taha Choukhmane† Tim de Silva‡ Weidong Lin§ Matthew Akuzawa¶

December 1, 2025

Latest Version

Abstract

How does AI-generated personal financial advice from Large Language Models (LLMs) compare to economists' normative models? We develop and implement a method to evaluate generative AI financial advice by simulating thousands of life cycle paths for consumption, saving, and portfolio choices under realistic income, employment, and asset return scenarios. Our approach compares these LLM-generated paths against optimal choices from a standard life cycle model and can parsimoniously summarize LLM advice across prompts and models by estimating structural time and risk preferences. Applying our method to OpenAI's `GPT-5 mini`, we find that the advice qualitatively aligns with standard life cycle theory but deviates systematically in four key ways: (i) recommended consumption and saving paths imply unrealistically high patience, with estimated intertemporal discount factors well above one; (ii) recommended choices often reflect simple heuristics, such as round savings rates, fixed-percentage withdrawal rules in retirement, and common asset allocation rules-of-thumb; (iii) LLM recommendations exhibit substantially more inertia in portfolio rebalancing and moderately less consumption-smoothing in unemployment than our normative benchmark; and (iv) holding all else constant, recommendations vary systematically with demographics (e.g., recommending lower equity shares for women) and between repeated identical queries.

*First draft: November 2025. We thank James Choi, Eric So, and Jonathan Parker for helpful feedback and Stanford's AI Playground for financial support and assistance.

†MIT Sloan School of Management and NBER, tahac@mit.edu.

‡Stanford University, Graduate School of Business and HAI, tdesilva@stanford.edu.

§MIT Sloan School of Management, wel@mit.edu.

¶MIT Sloan School of Management, maku49@mit.edu.

Individuals are increasingly seeking financial advice from large language models (LLMs). Recent industry surveys suggest that over half of adults in the U.S. and the U.K. have used LLMs for financial guidance (Lloyds Banking Group 2025; J.D. Power 2025), a rate that may already exceed the share of people who consult a human financial advisor (Gallup 2025). Given the rapid growth of Artificial Intelligence (AI), the promise of affordable, widely accessible, and quality personal financial advice may be closer than ever. Yet, there is little systematic evidence on the quality of this advice, largely due to the difficulty in setting a clear benchmark. Financial decisions are dynamic, involve significant uncertainty, and depend on individuals' evolving circumstances and preferences, all of which make it difficult to define what constitutes "good" advice.

In this paper, we ask: how does LLM personal financial advice compare to the prescriptions of economists' normative models? We develop and implement a method to answer this question, using a standard life cycle portfolio choice model as our normative benchmark. Applying this method to a widely used LLM, OpenAI's `GPT-5 mini`, we simulate thousands of life cycle paths for consumption, saving, and portfolio choices under realistic income, employment, and asset return scenarios, and compare these LLM-generated paths against optimal choices derived from a life cycle model. In effect, we trace the key financial decisions of individuals who follow LLM advice during every year of their lives. This approach allows us to characterize LLMs' financial advice across varying labor and asset market conditions, estimate the structural preference parameters that best fit the LLM recommendations, and identify systematic deviations of this advice from our normative benchmark.

Our results suggest that `GPT-5 mini` can already provide financial advice that qualitatively aligns with life cycle theory, but this advice deviates systematically in four important ways: (i) the LLM's recommendations reflect unrealistically high patience, with savings behavior that our model can only rationalize with discount factors above one; (ii) recommended choices often rely on simple heuristics; (iii) recommendations exhibit substantial inertia in portfolio rebalancing and less consumption-smoothing in unemployment than our life cycle model; (iv) the advice exhibits variation across demographics (e.g., recommending lower equity exposure for women or higher savings rate for Asian individuals) and between identical repeated queries that is not well understood. Our interpretation of these findings, particularly the reliance on heuristics and demographic variation, is that they are consistent with LLM recommendations largely reflecting popular financial advice from books, websites, and news articles, which often varies across demographics and uses rules of thumb (Choi 2022). Additionally, while our results apply to a specific LLM and set of prompts at one

point in time, the core contribution of this paper is to provide a general framework for evaluating the quality of LLM financial advice that can be applied to compare other prompts and models, especially to track the evolution of LLM advice as these technologies improve.

Methodology. Our method involves three steps. First, we build a quantitative life cycle model similar to Gourinchas and Parker (2002) and Cocco et al. (2005) with a stochastic income process, labor market transitions, and asset returns calibrated to match nationally representative U.S. data. The model most closely resembles Choukhmane and de Silva (Forthcoming) and features liquid safe and risky assets, income and mortality risk, job-to-job to unemployment transitions, as well as a realistic tax and social insurance system. Individuals choose how much to consume, how much to save, and how to allocate their savings between stocks and bonds. After solving the model, we simulate thousands of individuals' decisions over their full life cycles.

In the second step, we design a detailed prompt in which we ask an LLM to provide financial advice to a hypothetical individual. The prompt elicits the LLM's recommended choices, providing as input the key state variables in our life cycle model and information about the economic and demographic environment, such as assumptions about asset returns and mortality risk. When designing this prompt, we balance two objectives: mimicking how a typical individual, perhaps with guidance from a human financial advisor, would interact with an LLM, while ensuring the LLM has sufficient information to make decisions comparable to the life cycle model. We highlight constraints in the prompt to ensure sensible recommendations, such as requiring choices that satisfy an individual's budget constraint. We then vary the prompt to test how information about the economic environment, individuals' preferences, and demographics affect the LLM's financial advice.

Finally, we simulate life cycle paths of LLM advice by querying the LLM at every age from 22 to 89 for each individual in our model. Each query is independent, with no memory of past responses; the only link across years is the evolution of state variables that depend on prior choices. We then compare these LLM-generated paths against the optimal choices implied by our model for the same sequences of exogenous shocks. This approach builds on emerging methods for assessing LLM financial advice (Fedyk et al. 2025; Ouyang et al. 2025; Ross et al. 2024) by extending them to a dynamic stochastic environment, which allows us to establish a realistic normative benchmark. To our knowledge, we are the first to combine LLM responses with a quantitative life cycle model.

Main findings. Applying this method to OpenAI's `GPT-5 mini` delivers four main results.[1] Our first result is that the LLM's financial advice aligns qualitatively with standard life cycle theory. Focusing on consumption and saving decisions, the LLM recommends that individuals save during working years and consume less than their income, while recommending that individuals dissave to fund consumption in retirement. In terms of portfolio choices, the LLM recommends that over 98% of individuals participate in the stock market and choose high equity shares that decline with age from around 80% early in life to 60% later in life. Additionally, the LLM is able to incorporate time and risk preferences according to standard theory: recommended wealth accumulation and equity shares increase for individuals described as more patient or less risk-averse.

Second, we find that the LLM's advice is consistent with moderate risk aversion but unrealistically high patience. To show this, we jointly estimate the preference parameters that minimize the distance between simulated life cycle profiles recommended by the LLM and those from our life cycle model. We find that the equity shares recommended by the LLM imply a relative risk aversion coefficient of 4.7, which is similar to standard estimates. In contrast, the LLM consumption-savings choices exhibit unrealistically high patience: our model only replicates the very high savings rates recommended during working years with an intertemporal discount factor of 1.06, far exceeding typical values estimated in observational and experimental data. Even when our prompt explicitly states that individuals have no bequest motive and do not care about wealth after death, the LLM recommends minimal wealth decumulation in retirement that our model cannot match for any reasonable preference parameters. We then repeat our estimation for each set of time and risk preferences stated in the prompt. Despite the extreme patience on average, we find the LLM is able to incorporate the stated preferences quantitatively: the estimated discount factors and coefficients of risk aversion increase monotonically with the provided preferences. More broadly, we view our structural estimation of preference parameters as a parsimonious way to summarize an LLM's advice that can be estimated across different models and prompt variations.

Our third result is that the LLM's advice systematically deviates from our normative benchmark in ways that suggest the use of simple heuristics and passive responses to shocks. The LLM exhibits round-number bias, frequently recommending savings in multiples of 10% of salary or $5,000, as well as withdrawals of exactly 4% of wealth in retirement

[1] We choose OpenAI's model because ChatGPT is individuals' most common choice when searching for financial advice from LLMs (Lloyds Banking Group 2025; J.D. Power 2025). However, because `GPT-5 mini` is a reasoning model, we cannot modify its temperature parameter to eliminate randomness at the inference stage. We analyze this randomness in the final section of the paper.

and a 60-40 stock-bond allocation. Additionally, relative to our life cycle model, the LLM advice produces less consumption smoothing in response to unemployment shocks and substantially more inertia in the form of passive portfolio drift. While these behaviors deviate from our normative benchmark, they quantitatively align with empirical evidence on actual behavior (Gruber 1997; Calvet et al. 2009).

Finally, we show that the LLM's recommendations exhibit two sources of variability that warrant caution. The first source of variability is differences in advice across demographic groups. Recommended equity shares are systematically lower when an individual is described as a woman rather than a man, despite identical income, employment, and asset returns. Additionally, an individual described as an Asian man accumulates approximately $24,000 more wealth by age 60 than one described as a Black man. These differences emerge despite our prompt explicitly stating that labor market outcomes, marital status, and life expectancy are identical across demographic groups. Nevertheless, they align with empirical evidence that women tend to choose lower equity shares (Agnew et al. 2003) and that Asian workers save the most in retirement savings plans and Black workers save the least, conditional on observables (Choukhmane et al. 2025a). The second source of variability is differences in recommendations between repeated queries with identical prompts and inputs. This variation is important to recognize because it raises questions about the reliability of LLMs for high-stakes financial decisions.[2] However, we find that it is not too large economically in the sense that the estimated preference parameters are almost identical across our different simulations with identical prompts.

Related literature. An extensive literature in behavioral economics and household finance has documented numerous ways in which households make suboptimal financial decisions.[3] While this introduces a role for professional financial advisors, the quality of such advice is often limited due to advisors' high cost, mistaken beliefs (Linnainmaa et al. 2021), and conflicts of interest (Mullainathan et al. 2012; Egan et al. 2019). Similarly, while popular personal finance books provide lower-cost and accessible financial guidance, this advice often contains fallacies and deviates from the prescriptions of normative economic models (Choi 2022). Given the rapid growth of AI, a natural question is whether forms of AI, such as LLMs, can provide an affordable, widely accessible source of high-quality personal financial

[2]This final set of findings is consistent with Cen et al. (2025), who show that LLMs exhibit substantial "irregular variability" that cannot be explained and meaningful sensitivity to demographics.

[3]This includes evidence that households do not take full advantage of available 401(k) employer matches (Choi et al. 2011; Choukhmane et al. 2025b), do not efficiently allocate assets across taxable and tax-advantaged accounts (Bergstresser and Poterba 2004), or fail to refinance fixed-rate mortgages when beneficial (Andersen et al. 2020). For reviews, see Beshears et al. (2018), Gomes et al. (2021), and Campbell and Ramadorai (2025).

advice. This paper's main contribution is to introduce and implement a framework for evaluating the quality of this advice that recognizes the inherent complexity of households' financial decisions. In doing so, this paper is part of a growing literature that evaluates generative AI, and more specifically LLMs, in economics and finance (see Mo and Ouyang 2025 for a review).

Most closely related is the emerging literature that studies LLMs as models of human behavior in economic settings. These studies show that LLMs sometimes exhibit similar behavioral biases to humans in classic economic experiments, while behaving more closely to rational utility-maximizing agents in other settings (Horton 2023; Ross et al. 2024; Bini et al. 2025). Part of this literature studies LLMs' behavior specifically in financial settings.[4] For example, Ouyang et al. (2025) study the stability of the risk-taking behavior of LLMs across a variety of behavioral tasks, including a stylized asset allocation problem, while Fedyk et al. (2025) show that the preferences of LLMs over different asset allocations more closely resemble those of younger high-income individuals, while also being less likely to violate transitivity. Relative to this literature, we develop a methodology to evaluate the quality of the LLMs' responses by combining them with a dynamic life cycle model. This allows us to study behaviors that are more specific to household financial decision-making, such as consumption-smoothing and portfolio rebalancing.

Finally, our study of LLMs as an accessible and cost-effective source of financial advice builds on prior work evaluating other financial innovation and forms of automated advice. This literature has shown that new sources of financial advice, such as robo-advisors, can help individuals make higher-quality financial decisions (D'Acunto et al. 2019; Rossi and Utkus 2020; Reher and Sokolinski 2021). More broadly, our framework extends a longstanding tradition of using quantitative life cycle models as normative benchmarks to evaluate new financial products and policies, with recent applications to adjustable rate mortgages (Guren et al. 2021; Campbell et al. 2021), target-date funds (Duarte et al. 2024), saving nudges (Choukhmane 2025), and income-contingent student loans (de Silva 2025). We contribute by applying this approach to what is perhaps the fastest-growing area of innovation in personal finance: generative AI.

[4]Another strand of this literature studies the conditions under which access to generative AI affects the quality individual's investment decisions (Kim et al. 2024).

1 Normative Benchmark: Life Cycle Model

This section describes the life cycle model of consumption-saving and portfolio choices that we use as a normative benchmark. This model builds on the model of Choukhmane and de Silva (Forthcoming), with the following key differences. First, we introduce transitory income shocks with volatility calibrated to SIPP data. Second, we remove the retirement investment vehicle and only allow the agent to save in saving and investment accounts. Third, the tax and social security schedules as well as the distribution of income have been updated to keep up to date with the LLM's perception of said systems and distributions. Lastly, the income process is recalibrated to match 2025 data.

1.1 Demographics and Preferences

Each period corresponds to one year, and working life starts at $t = 0$ and lasts for T_w periods. Retirement starts at $t = T_w$, and agents can live at most T periods. Before their certain death in period $t = T$, investors face age-dependent mortality risk with survival probability in period $t + 1$ conditional on survival in period t denoted by m_t. We denote an investor's age as $a_t = t + a_0$, where a_0 is the age investors enter working life. Investors have recursive CRRA utility over consumption streams. We denote investors' annualized time discount factor as β and relative risk aversion as γ.

1.2 Labor Market

At any point in time, investors can be in one of four employment states, denoted emp_t: E = employed by the same employer as in the previous period, JJ = employed by a different employer than in the previous period, U = unemployed in the current period, and Ret = retired. The fact that investors face uncertainty about their future employment status, in addition to earnings risk, is an important feature of our model because it introduces deviations in income shocks from normality, which Guvenen et al. (2021) highlight as important empirically.

Employment: $emp_t = E$**.** While working, investors earn an exogenous income w_t. The log income process consists of a deterministic component that is cubic in age, a stochastic component that follows an AR(1) process with normally distributed innovations, and a transitory shock that is normally distributed:

$$\begin{aligned} \ln w_t =& \delta_0 + \delta_1 a_t + \delta_2 a_t^2 + \delta_3 a_t^3 + \eta_t + \varepsilon_t, \quad \eta_t = \rho \eta_{t-1} + \xi_t^E, \\ & \xi_0^E \sim N(0, \sigma_{\xi_0}^2), \quad \xi_t^E \sim N(0, \sigma_\xi^2), \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \qquad \forall t > 0. \end{aligned} \tag{1}$$

Investors' tenure status evolves according to $ten_t = ten_{t-1} + 1$ if they remain employed by the same employer. We assume that the initial distribution of ε_t^E is different in the first period ($t = 0$) to account for heterogeneity in the initial period incomes.

Job transition: $emp_t = JJ$**.** While in the employed state (E), an investor may transition from job-to-job with a probability $\pi^{JJ}(t, ten_t)$ that depends on both her age and tenure at the current job. After a job-to-job transition, income evolves according to:

$$\ln w_t = \delta_0 + \delta_1 a_t + \delta_2 a_t^2 + \delta_3 a_t^3 + \eta_t + \varepsilon_t, \quad \eta_t = \rho\eta_{t-1} + \xi_t^{JJ}, \tag{2}$$

$$\xi_t^{JJ} \sim N(\mu^{JJ}, \sigma_\xi^2), \qquad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{3}$$

This earnings process captures a wage premium associated with switching jobs. Investors' tenure is reset to $ten_t = 0$ following a job-to-job transition.

Unemployment: $emp_t = U$**.** While in the employed state (E), an investor may become unemployed with a probability $\pi^{EU}(t, ten_t)$ that depends on both her age and tenure at her current job. When investors are unemployed, they receive unemployment benefits equal to $ui_t = ui(\eta_t)$, where $ui(\eta_t)$ is described below. If investors become employed at $t+1$ after being unemployed in period t, income at $t+1$ evolves according to

$$\ln w_{t+1} = \delta_0 + \delta_1 a_{t+1} + \delta_2 a_{t+1}^2 + \delta_3 a_{t+1}^3 + \eta_{t+1} + \varepsilon_{t+1}, \tag{4}$$

$$\eta_{t+1} = \rho\eta_t + \xi_{t+1}^U, \quad \xi_{t+1}^U \sim N(\mu^{UE}, \sigma_\xi^2) \quad \varepsilon_{t+1} \sim N(0, \sigma_\varepsilon^2). \tag{5}$$

This earnings process captures the persistent wage reduction associated with experiencing unemployment.

Retirement: $emp_t = Ret$**.** In period $t = T_w$, all investors retire deterministically. During retirement in periods $t \in [T_w, T-1]$, investors earn public pension benefits denoted by ss_t, which are described below.

1.3 Savings Account

Investors start with zero assets at $t = 0$ and cannot borrow. They can accumulate assets inside a liquid taxable savings account, which can be invested in one of two financial assets. First, there is a risk-free bond that has a constant gross return of $R_t^B = R_f$ per year. Second, there is a risky asset that corresponds to a diversified stock market index and pays a stochastic i.i.d. gross return of $R_t^S = R_t$ per year, where

$$\ln R_t^S = \ln R_f + \mu_s + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_s^2). \tag{6}$$

The balance of the savings account, denoted by L_t, then evolves according to:

$$L_{t+1} = \left(L_t + s_t^l\right)\left[1 + (R_{t+1}^{\theta} - 1)(1 - \tau_c)\right], \quad L_0 = 0, \tag{7}$$

where s_t^l is the net savings that the investor places in this account, τ_c is the rate of capital taxation, and R_t^{θ} is the investors' portfolio return that depends on their portfolio share in the risky asset, θ_t.

$$R_{t+1}^{\theta} = R^f + \theta_t \left(R_{t+1}^S - R^f\right). \tag{8}$$

1.4 Government

Unemployment benefits. Investors receive an unemployment benefit of $ui(\eta_t)$ when their employment ends. This benefit depends on the labor productivity, η_t, from the last period in which the agent was employed.

Retirement benefits. After retirement, investors receive Social Security benefits, denoted by $ss_t = ss(ae_{T_w})$. ae_{T_w} is the investor's average lifetime earnings at the time of retirement, which evolves according to:

$$ae_{t+1} = \begin{cases} \frac{w_{t+1} + t * ae_t}{t+1}, & \text{if } t < T_w, \\ ae_{T_w} & \text{else.} \end{cases}$$

Taxation. Investors face a nonlinear income tax schedule $tax_i(\cdot)$, which depends on their taxable income. For employed individuals or those in job transitions, their taxable income is their wages. For unemployed or retired individuals, their taxable income is the benefits they receive from the government.

1.5 Summary of Investors' Problem

Investors face a dynamic optimization problem with 7 state variables: a_t = age, η_t = labor productivity, ε_t = transitory income shock, emp_t = employment status, ten_t = tenure, ae_t = average lifetime income, and L_t = savings. Denote the vector of these state variables as $\mathbf{x}_t$. Investors have 3 controls: c_t = consumption, θ_t = portfolio share, and s_t^l = savings. In choosing these controls, we restrict investors from borrowing and engaging in any margin trading (i.e., short-selling or taking leveraged positions):

$$L_t \geq 0, \quad \theta_t \in [0, 1]. \tag{9}$$

1.6 Calibration

We calibrate our model parameters closely following Choukhmane and de Silva (Forthcoming). We provide a brief overview here and refer the reader to that paper for a more detailed discussion.

Demographics. We set the length of one period in the model to one year and set $a_0 = 22$, $T_w = 43$, and $T = 68$, such that workers are born at 22, retire at 65, and live their final year of life at 89. For each age, we calibrate mortality risk to match the 2015 U.S. Social Security Actuarial Life Tables. We use the equivalence scale estimated in Lusardi et al. (2017) to capture changes in individual composition over the life cycle.

Labor income process. We use data from the Survey of Income and Program Participation (SIPP) to estimate parameters of the labor income process and transition probabilities at the annual frequency. This income process has several components. First, we estimate an earnings process for workers staying in the same job, corresponding to (1), which contains a deterministic and stochastic component. Second, we use data on employment transitions from SIPP to estimate the median salary increase following a job-to-job transition, μ^{JJ}, and the median salary decrease when workers transition back to employment after an unemployment spell, $-\mu^{EU}$. Third, we use SIPP microdata to estimate the three transition probabilities between the three labor market states. Finally, we set the initial unemployment rate equal to 22%, which is the share in SIPP of unemployed individuals at age 22.

Asset returns. We set the net risk-free rate to be constant at 2% to match the average market yield on 1-year Treasury Securities between 2015 and 2025. We set the equity premium to be 6.4%, which is equal to the average inflation-adjusted return on the CRSP Value-Weighted Index between 1925 and 2006 minus our 2% risk-free rate.[5] We set the volatility of log stock returns to 20%, which matches that of the CRSP Value-Weighted Index. We assume that asset returns are uncorrelated with shocks to labor income and employment transition probabilities.

Tax and benefit system. Investors' tax liability, $tax_i(\cdot)$, is calculated according to the 2025 U.S. federal income tax schedule for single filers who claim only the standard deduction. We calculate Social Security benefits according to the 2025 formula with a Supplemental Security Income program floor. Unemployment benefits are computed with a replacement rate of 40%. We set the capital gains tax rate, τ_c, to 21%.

[5] We adjust for inflation using the CPI.

2 Eliciting LLM-Based Financial Advice

In this section, we describe our selection of an LLM to apply our method and the design of prompts used to elicit personalized financial advice. We then simulate the sequence of LLM-based financial decisions for each individual over the life cycle. Finally, we incorporate personal preferences and demographic variability, focusing on race and gender, into our prompts to assess how the LLMs' recommendations differ across preferences and demographic groups.

2.1 Model Selection

We use a state-of-the-art model, `GPT-5 mini`, recently released by OpenAI (OpenAI 2025). This model represents the latest generation of publicly accessible large language models and closely reflects the systems that users engage with through ChatGPT and similar AI interfaces. Additionally, survey evidence suggests that when searching for financial advice from LLMs, individuals' most common choice is ChatGPT (Lloyds Banking Group 2025; J.D. Power 2025). We select `GPT-5 mini` instead of the full `GPT-5` model primarily for reasons of cost efficiency.

2.2 Baseline Prompt Design

We design a detailed prompt that asks an LLM to provide financial advice to a hypothetical individual. The prompt outlines baseline assumptions about the economic environment, including expected asset returns, tax and Social Security rules. These assumptions provide the LLM with enough context to generate coherent and feasible financial advice.[6] To enable personalized recommendations, we include individual-specific state variables, such as age, income, balances in both the individual's savings and investment accounts, employment status, job tenure, and average past income. These variables capture the key determinants of optimal behavior in our life cycle model. However, we do not provide detailed information about the distribution of income and employment risk, such as higher moments of the income process or employment transition probabilities, beyond a statement in our baseline prompt about "normal job market risk". The LLM is then asked to jointly determine the individual's optimal consumption, saving, and investment decisions. When designing this prompt, we aim to strike a balance between providing information in a way that resembles how a typical individual would interact with an LLM, while also ensuring that

[6]To assess the sensitivity of our results to the baseline assumptions, we rerun the analysis omitting all economic assumptions except the portfolio allocation and accumulation rules. The results remain consistent with our main findings. See Section 2.3 for additional discussion.

Table 1. Mapping Between Baseline LLM Prompt Assumptions and Life Cycle Model Calibration

LLM Prompt Assumption	Life Cycle Model Implementation
"Normal life expectancy"	Mortality risk calibrated to SSA life tables
"Normal retirement age"	Exogenous retirement at age 65
"Normal job market risk"	Persistent income risk and employment transitions calibrated to SIPP data, varying by age, income, and tenure
"Current U.S. tax and Social Security rules stay unchanged"	2025 Federal income tax schedule (single filer) and Social Security benefit formula
"Risk-free savings earn 2.0% real return annually"	2% annual risk-free return
"Real stock returns match the 60-year U.S. total stock market historical average"	Log-normal returns: 6.4% equity premium, 20% standard deviation, uncorrelated with income
"I am single with no dependents and have no bequest motive. I do not care about wealth after death."	One individual per household and zero utility from bequests

Notes: This table summarizes the correspondence between the baseline assumptions in the LLM prompts and their calibration in the life cycle model.

the LLM receives enough inputs to produce choices consistent with fundamental economic constraints, such as the budget constraint. The full baseline prompt is presented in Figure A1. Table 1 shows how the baseline assumptions of our prompt correspond to the calibration of our life cycle model.

We use this prompt to simulate individual decisions by sequentially prompting an LLM to provide financial advice to each individual throughout the life cycle. In each period, the individual's personal information is fed into the prompt, and the LLM recommends consumption, saving, and investment decisions. The realizations of the model's exogenous state variables are identical to those in our benchmark life cycle model. The key difference is that we treat the LLM's recommended choices as the individual's policy functions and update the endogenous state variables accordingly. Specifically, at the beginning of each year, individuals receive their annual income and implement the LLM's consumption and portfolio recommendations as their choices. These choices generate asset returns over the year, and the resulting account balances and portfolio shares become inputs for the next period's prompt. For additional details, see Appendix A.

2.3 Additional Prompt Variations

Parsimonious prompt. To examine the robustness of the LLM's recommendations and assess their sensitivity to the baseline economic assumptions, we construct a more parsimonious prompt. This alternative prompt does not include any of the assumptions reported in Table 1, while keeping all other inputs identical to the baseline prompt. See Figure A2 for the full description of this parsimonious prompt.

Time and risk preferences. We examine how LLMs capture individual time and risk preferences. To reflect how economists typically elicit these preferences, we include information on how the individual values resources at different points in time and on their certainty equivalent for a given lottery. Panel B of Table 2 reports the prompt statements used to convey information about time and risk preferences. We construct the risk-preference statement by deriving the certainty equivalent of a lottery under CRRA utility as described in Appendix A.2. Figure A3 and Figure A4 provide examples of the individual-specific prompts that include time- and risk-preference statements, separately.

Demographics. To assess how the models respond to demographic heterogeneity, which prior literature has shown is important (Fedyk et al. 2025; Cen et al. 2025), we consider eight demographic groups defined by the intersection of two genders (male and female) and the four largest racial and ethnic groups in the U.S. (White, Black, Hispanic, and Asian). For each group, we then incorporate the demographic attributes in the prompts and sequentially query the LLM for advice, while keeping the income and employment processes identical across groups. Panel C of Table 2 reports the prompt statements, while Figure A5 gives an example of an individual-specific prompt that includes a demographic statement.

3 Comparing LLM Financial Advice Against the Life Cycle Model

In this section, we begin by describing the financial advice provided by the LLM, specifically how it varies according to an individual's state and preferences. We then compare this advice quantitatively with the optimal policy from our life cycle model.

3.1 LLM Advice Qualitatively Aligns with Standard Life Cycle Theory

Figure 1 shows the median consumption, wealth, and stock share by age that are chosen by LLM-simulated individuals, along with the corresponding 25th-75th percentiles; additional summary statistics are presented in Table A1. The left panel shows that, consistent

Table 2. Summary of Individual-Specific LLM Prompt Variations

Category	LLM Prompt Statement
Panel A: Employment and Income	
Case 1: Employed	
Employment	"I have started a new job this year" "I have been working at the same firm for {tenure} years."
Current Income	"My annual take-home pay income (after taxes) is $X this year"
Past Income	"My average annual income since age 22 is $Y. Going forward, my average income will determine my Social Security benefits in retirement"
Case 2: Unemployed	
Employment	"I am currently unemployed"
Current Income	"My annual unemployment benefit income (after taxes) is $X this year"
Past Income	"My average annual income since age 22 is $Y. Going forward, my average income will determine my Social Security benefits in retirement"
Case 3: Retired (Age $\geq$ 65)	
Employment	"I am retired"
Current Income	"My annual Social Security benefit income (after taxes) is $X this year"
Panel B: Preferences	
Time Preference	"Your advice should reflect my time preference. Roughly, $Z this year feels about the same as $100 next year"
Risk Preference	"Your advice should reflect my risk tolerance. I would be indifferent between winning a guaranteed $W and a 50% chance of winning $G"
Panel C: Demographics	
Gender & Race	"I am a {White/Black/Hispanic/Asian} {man/woman}"

Notes: This table summarizes the individual-specific statements used in the LLM prompts.

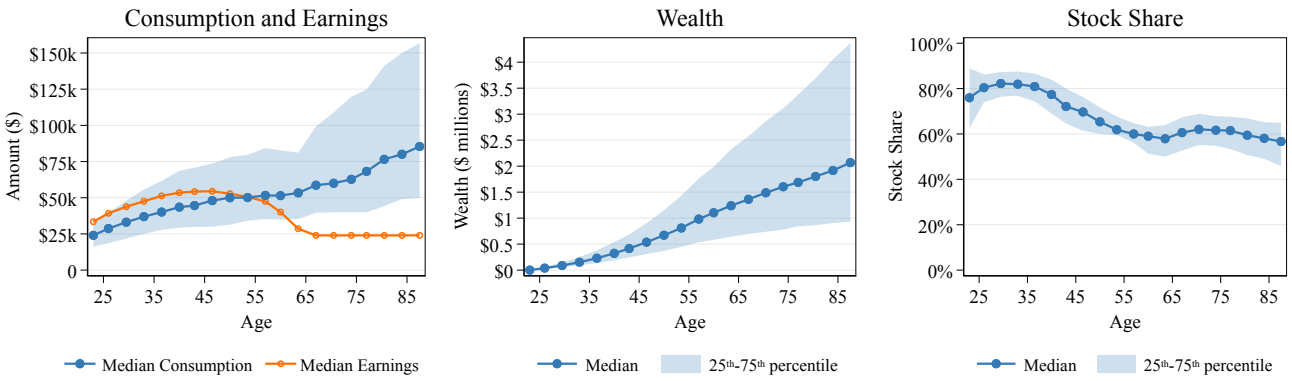
with textbook life cycle theory, the LLM recommends that individuals smooth their income over working life such that consumption is much flatter than income. The LLM-simulated individuals achieve this by saving more during working life and consuming less than their income, while saving less and withdrawing after age 55 and consuming more than their income. This is reflected in their wealth accumulation in the middle panel: individuals following the LLM's advice build up a large stock of over $1 million in assets by the time they retire at age 65.

Turning to portfolio choices, Table A1 shows that the LLM recommends that more than 99% of individuals participate in the stock market. The right panel of Figure 1 shows that the recommended equity shares are relatively high, around 65% on average, and decrease

with age. Qualitatively, these portfolio choices are consistent with textbook theories of portfolio choice with human capital (e.g., Merton 1969; Gomes 2020). While less steep than a typical target-date fund glide path, the LLM's recommended equity allocation in working life is remarkably similar to the average equity shares that retirement savers choose when making active portfolio choices (Choukhmane and de Silva Forthcoming).

Before proceeding to our main findings, we assess the robustness of our results to prompt design. We consider an alternative prompt that excludes all economic environment assumptions listed in Table 1 (see Section 2.3 for details). As shown in Figure A6, our results from using this minimal prompt are quantitatively nearly identical to our baseline. This suggests our findings are robust to small variations in prompt design.

Figure 1. Life Cycle Profiles of LLM-Recommended Consumption, Wealth, and Stock Shares



Notes: This figure plots the life cycle profiles for the simulated individuals who follow the LLM's recommended choices using the baseline prompt, as described in Section 2.2. Ages (22-89) are grouped into 20 equally sized bins. The left panel shows consumption and post-tax salary; the middle panel shows wealth accumulation; and the right panel shows the stock share. Dots denote median values within each bin, and shaded areas represent the interquartile range (25^{th}-75^{th} percentile). All values are in 2025 dollars.

3.2 LLM Advice is Extremely Patient and Moderately Risk Averse

Having characterized the basic properties of the LLM's financial advice, we now turn to a quantitative comparison with our life cycle model. Our approach is to estimate the intertemporal discount factor (β) and coefficient of relative risk aversion (γ) that would rationalize the LLM's recommendations using the simulated method of moments (SMM). Specifically, we estimate the values of β and γ that minimize the difference between model-generated and LLM-generated life cycle profiles. We target two sets of moments: the wealth-to-income ratio during working years and the equity share over the full lifespan (see Appendix B for additional details).

The first row of Table 3 shows the estimated values when using the baseline LLM prompt,

which are $\beta = 1.06$ and $\gamma = 4.65$. This estimated value of risk aversion suggests that the LLM's advice is moderately risk averse, and is within the range of standard estimates. However, the estimated intertemporal discount factor is unrealistically high, well above one. At these estimated values, Figure 2 compares the life cycle profiles of consumption, wealth, and equity shares between the model and the LLM. The results show that, even with this high discount factor, the LLM decumulates assets in retirement at a much slower rate than in our model, resulting in lower consumption. While bequests could in principle explain this, our prompt explicitly states that individuals have no bequest motive and do not care about wealth after death.[7]

Table 3. Simulated Method of Moments Estimation Results

Prompted Values		**Estimated Values**	
β	γ	$\hat{\beta}$	$\hat{\gamma}$
Baseline (No Preference Prompts)			
-	-	1.06	4.7
Time Preference Prompts			
0.9	-	**0.98**	5.5
0.95	-	**1.04**	5
0.99	-	**1.06**	4.7
Risk Preference Prompts			
-	2	1.03	**6.5**
-	5	0.96	**9.5**
-	8	0.92	**11**

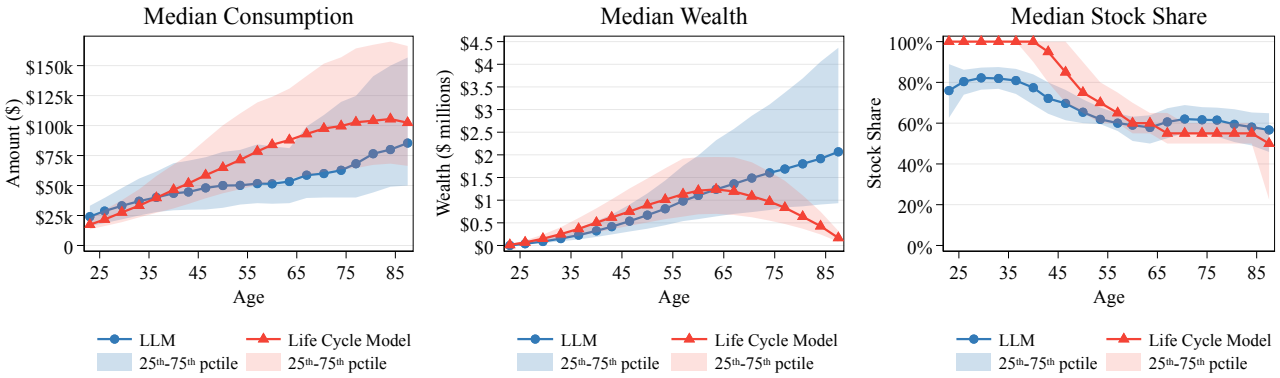
Notes: This table presents the results of our grid search estimation. The left two columns show the β or γ that we input into the LLM model using the prompts shown in figures A3 and A4. The two columns on the right present the results of our grid search estimation, where we minimize the sum of squared proportional errors between a set of LLM simulation moments and their equivalent moments calculated in the life cycle for each set of preference points on a naive grid. The moments we use are the mean wealth to income ratio in each year of working life and the mean equity share for each year in the full life. For a full description of how this estimation works, see section B.

Turning to equity shares in the right panel of Figure 2, we find that the life cycle profile of equity shares in our model is quite close to that recommended by the LLM. A notable exception is the first few years of life, in which the model's equity share is 100% at the mean, which is a well-known feature of portfolio choice models with standard preferences, uncorrelated labor income risk, and no participation costs (Cocco et al. 2005). In contrast, the LLM recommends a lower equity share of around 80% early in life. This finding is consistent with life cycle models that feature a positive correlation between labor income risk and stock returns (Benzoni et al. 2007; Catherine 2022). In these models, this correlation

[7] Table A2 shows that our estimates are almost identical if we instead use the simpler prompt, which (among other assumptions) omits the statement about no bequest motive. See Figure A6.

makes human capital less bond-like, reducing the optimal equity share, especially early in life.

Figure 2. Comparison of Life Cycle Profiles Between LLM and Life Cycle Model



Notes: This figure compares life cycle profiles of consumption, wealth, and stock allocation between LLM-generated and model-based recommendations. Blue dots represent simulated outcomes when individuals follow the LLM's advice (baseline prompt described in Section 2.2). Red triangles represent outcomes from the structural life cycle model (Section 1). Each point shows the median value for one of 20 equal-sized age bins, with shaded areas indicating the interquartile range (25^{th}-75^{th} percentile). All values are in 2025 dollars.

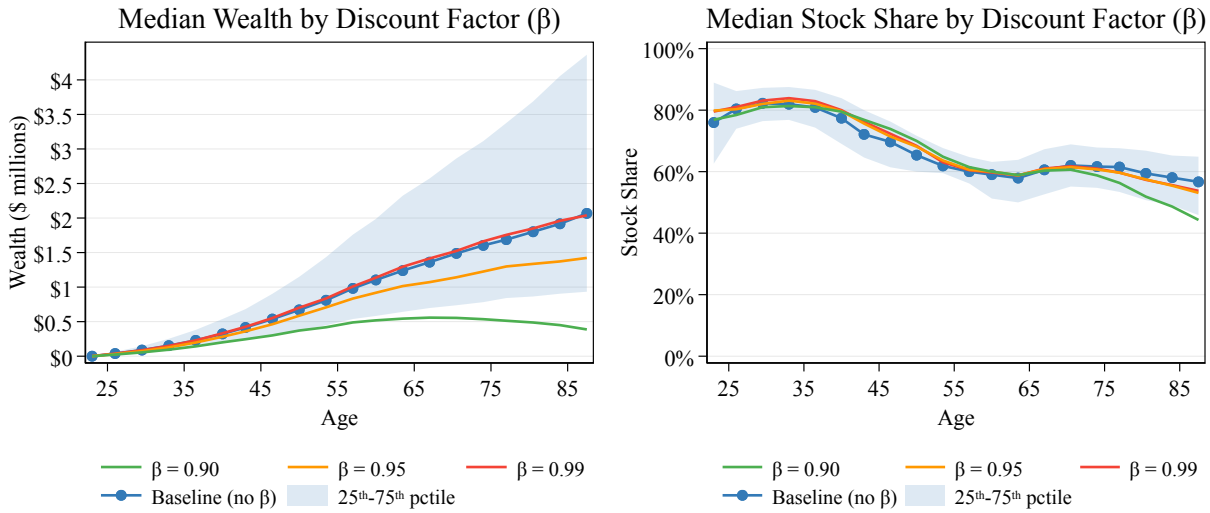
3.3 LLM Advice Can Quantitatively Incorporate Stated Preferences

Next, we ask how the recommendations of the LLM respond to providing time and risk preferences in the prompt, as described in Section 2.3. In Panel A of Figure 3, we show how the life cycle profiles of wealth and stock shares vary as we vary the provided time preference rate. For reference, we also include the baseline profiles from Figure 1 that correspond to the case in which no preferences were provided. Consistent with standard theory, a higher discount factor causes the LLM to recommend more wealth accumulation. Quantitatively, the baseline simulation (without prompting any time preference) produces results most similar to the case where we prompt the LLM with $\beta = 0.99$.

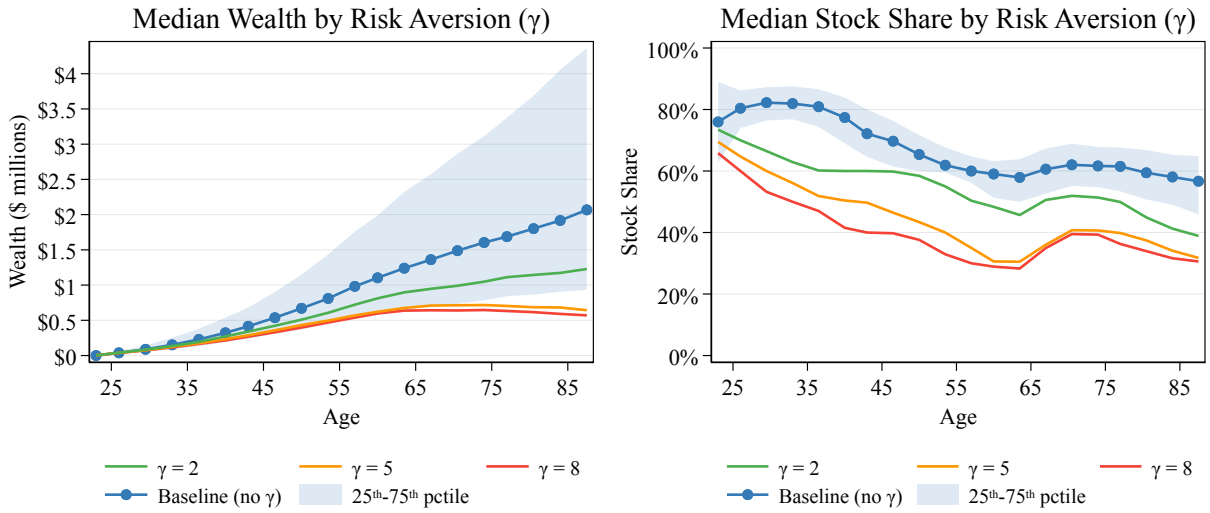
Panel B of Figure 3 shows the same analysis as in Panel A, where we vary risk preferences that are provided in the prompt instead of time preferences. As the provided risk aversion increases, the LLM recommends slightly less wealth accumulation and, more importantly, meaningfully lower equity shares, which is consistent with standard models of portfolio choice. Comparing the cases with risk preference prompts to the baseline shows that providing risk preference information leads to more risk-averse LLM recommendations. Even when prompting with lower values of risk aversion that are more in line with the literature, the LLM recommends lower equity allocations than in the baseline case.

Figure 3. Life Cycle Profiles of LLM Recommendations with Varying Preferences

Panel A: Varying Time Preferences



Panel B: Varying Risk Preferences



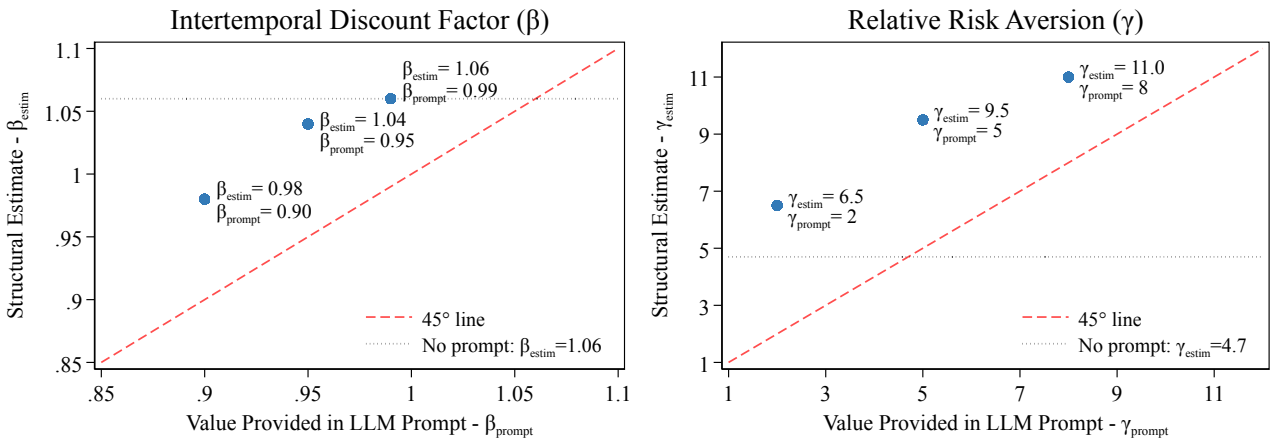
Notes: This figure illustrates preference heterogeneity in life cycle profiles for the simulated individuals who follow the LLM's recommended choices using prompts with specified individual preferences, as described in Section 2.3. The realizations of exogenous shocks are identical across groups except for their prompted preference values. Ages (22-89) are grouped into 20 equally sized bins. Panel A shows median wealth (left) and stock share (right) within each bin under different time preferences, while Panel B shows the corresponding results under different risk preferences. The blue line and shaded area represent the baseline results in which no preferences are specified, as shown in Figure 1. All values are in 2025 dollars.

While Figure 3 shows that the LLM's financial advice can qualitatively respond to changes in time and risk preferences in ways consistent with standard theory, we next assess whether it does so quantitatively. For each of the different values of the provided time preferences and risk aversion, we perform a separate SMM estimation using the same procedure

described previously. The results are shown in the separate rows of Table 3. Figure 4 plots these estimated values against the corresponding values that are provided in the prompt.

The left panel of Figure 4 shows that the LLM's recommendations do a quantitatively good job of incorporating different provided time preferences. The estimated discount factor increases nearly one-to-one with the prompted value, and the relationship is almost parallel to the 45-degree line. This suggests the LLM can effectively tailor its advice to different preferences, a finding we view as promising for personalization. However, the estimates lie consistently above the 45-degree line: the LLM's recommendations require higher patience than prompted to rationalize. For instance, prompting $\beta = 0.99$ yields an estimated $\hat{\beta} = 1.06$. This upward shift reflects the LLM's tendency to recommend very high asset accumulation.

Figure 4. Comparison of Estimated and Prompted Preferences



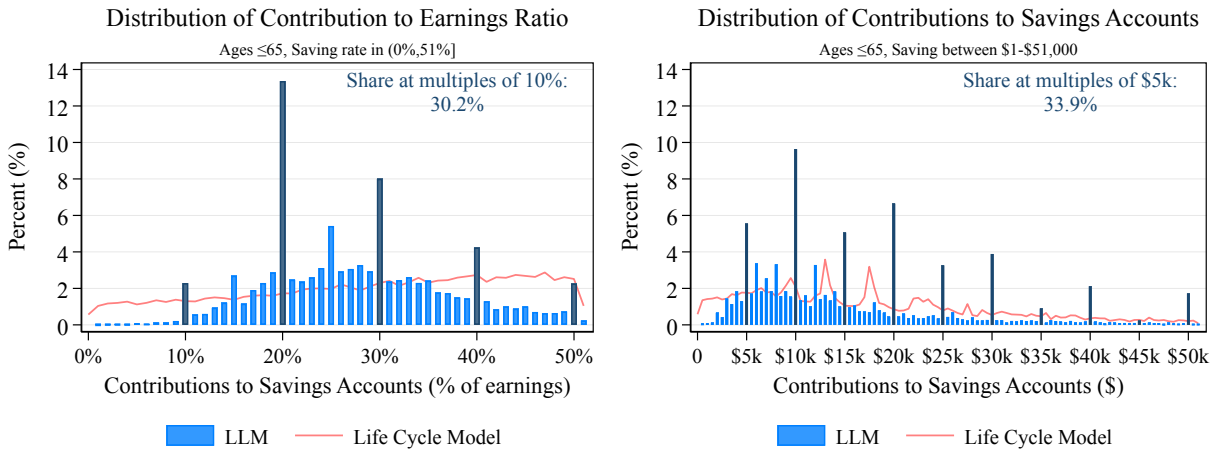
Notes: This figure compares the preference parameters provided in the LLM prompt with those estimated from the simulated life cycle profiles of individuals who follow the LLM's recommended choices. The left panel shows the intertemporal discount factor (β), and the right panel shows the coefficient of relative risk aversion (γ). Each blue dot represents a pair consisting of a prompted preference value (β_{prompt} or γ_{prompt}), and the corresponding simulated method of moments (SMM) estimate (β_{estim} or γ_{estim}) inferred from the LLM advice, as discussed in Section 3.2. The horizontal gray dotted lines indicate the estimated values from the baseline LLM simulation without preference statements in the prompt, as described in Section 2.2. The red dashed line denotes the 45-degree line.

The right panel of Figure 4 repeats the same analysis for the prompts that vary risk aversion instead of time preferences. Here we again find that the LLM quantitatively incorporates stated preferences: estimated risk aversion increases nearly one-to-one with prompted risk aversion, with points roughly parallel to the 45-degree line. However, the estimates again lie consistently above the 45-degree line. Even when prompting low risk aversion ($\gamma = 2$), the estimated coefficient is ($\hat{\gamma} = 6.5$). This indicates that the LLM's advice is systematically more risk-averse than the prompted risk preferences.

3.4 LLM Advice Often Relies on Simple Heuristics

Given the differences between the LLM's financial advice and our life cycle model in Figure 2, we now try to understand these differences by looking at the individual-level recommendations from the LLM. In Figure 5, we plot the distribution of savings rates, defined as contributions to savings accounts divided by wage and benefit income, in the left panel and the distribution of raw savings amounts in the right panel. In both panels, we focus on savings amounts pre-retirement, when individuals are disproportionately saving as opposed to withdrawing. The blue bars are a histogram of the LLM recommendations, while the light red line is the corresponding distribution in our model. Both panels show that the LLM disproportionately recommends savings amounts and rates that are round numbers, with 30% of savings rates being in multiples of 10% and 34% of savings amounts being in multiples of $5K. Figure A7 shows that the prevalence of both of these heuristics seems to be pretty consistent across age and income. By construction, these types of heuristics are not present in the choices made by our life cycle model.

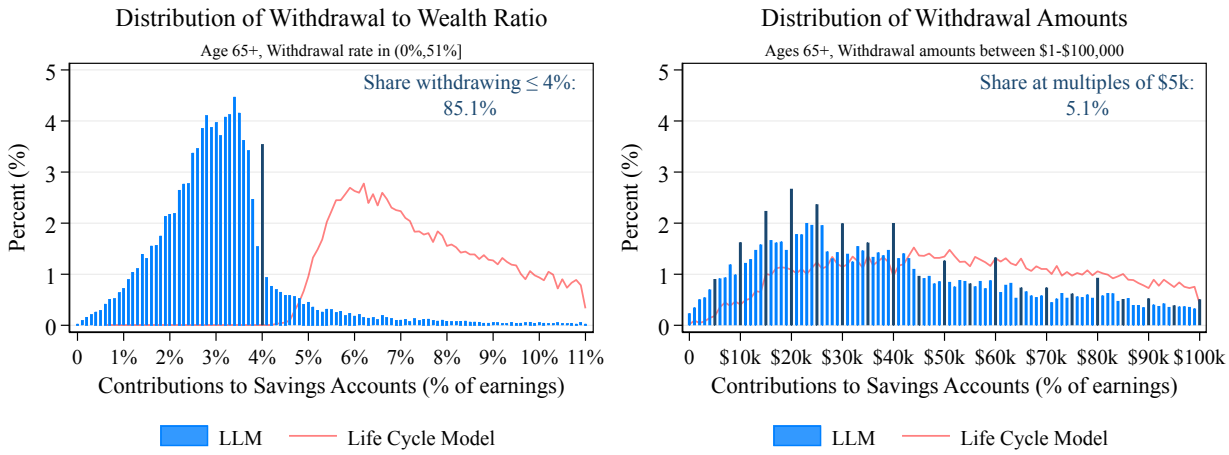
Figure 5. Saving Heuristics in LLM Advice



Notes: This figure plots the distributions of savings rates and savings amounts in both the LLM advice and the life cycle model for the simulated individuals during the pre-retirement period. The LLM advice is generated using the baseline prompt described in Section 2.2. The left panel compares the distributions of savings rates, defined as the ratio of savings amounts to income, between 0 and 51%. The right panel compares the distributions of savings amounts between $1 and $51,000. Blue histograms represent the LLM advice, and dark blue bars highlight heuristic savings, with spikes in savings rates at multiples of 10% (left) and in savings amounts at multiples of $5,000 (right). The red lines represent the corresponding distributions from the life cycle model described in Section 1. All values are in 2025 dollars.

Figure 6 repeats the same analysis as in Figure 5 for withdrawal amounts during retirement. We find that over 85% of the LLM's recommended withdrawals respect the 4% withdrawal rule, which is a rule that was popularized by Cooley et al. (1998) and is recommended by many financial advisors. In contrast, almost no individuals in our life cycle model choose withdrawal rates below 4%, consistent with the fact that the LLM's advice is overly conservative. Additionally, as in the case of savings amounts, the LLM disproportionately recommends withdrawal amounts that are in round increments, but this behavior is far less pronounced.

Figure 6. Withdrawal Heuristics in LLM Advice

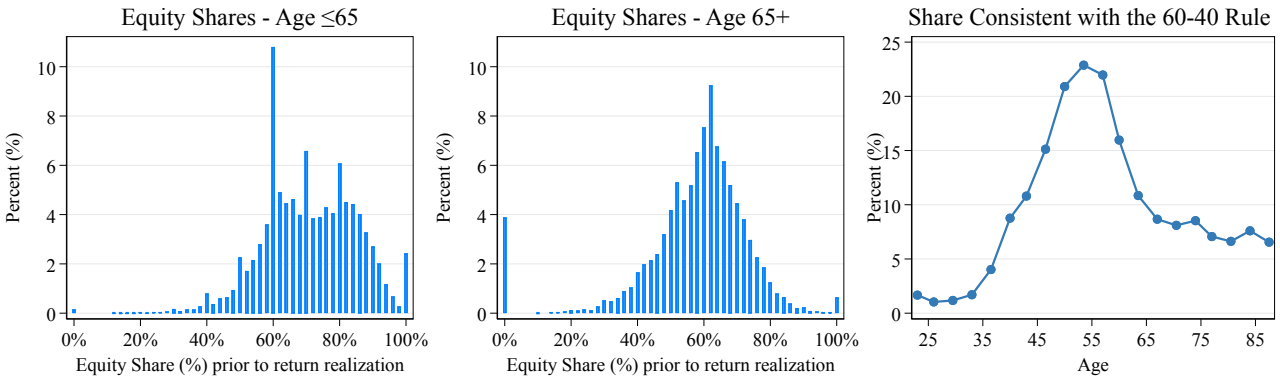


Notes: This figure plots the distributions of withdrawal-to-wealth ratios and withdrawal amounts in both the LLM advice and the life cycle model for the simulated individuals during the retirement period. The LLM advice is generated using the baseline prompt described in Section 2.2. The left panel compares the distributions of the withdrawal-to-wealth ratios between 0 and 51%. The right panel compares the distributions of withdrawal amounts between $1 and $100,000. Blue histograms represent the LLM advice, and dark blue bars highlight heuristic withdrawals, with spikes in withdrawal-to-wealth ratios at 4% (left) and in withdrawal amounts at multiples of $5,000 (right). The red lines represent the corresponding distributions from the life cycle model described in Section 1. All values are in 2025 dollars.

Turning to portfolio choices, Figure 7 shows the distribution of equity shares recommended by the LLM before and after retirement. During working life, the LLM recommends a participation rate of essentially 100%, but a disproportionate fraction of the recommendations are for an equity share of 60% consistent with the popular "60-40" stock-bond allocation rule.[8] This 60-40 allocation is most common in the middle of individuals' working lives, where the share choosing this exact allocation increases to almost 25%. During retirement, the LLM starts recommending non-participation occasionally—around 4% of the time—and centers its allocations around the 60-40 rule. However, the fraction of individuals following this exact rule is lower than during working life.

[8] Figure A8 examines the prevalence of two other common heuristics: the (120-age) and (110-age) rules.

Figure 7. Investing Heuristics in LLM Advice



Notes: This figure shows heuristic patterns in equity shares recommended by the LLM for the simulated individuals. The LLM advice is generated using the baseline prompt described in Section 2.2. The left panel plots the distributions of equity shares before return realization during the pre-retirement period, while the middle panel plots the distributions during the retirement period. The right panel shows the share of LLM-recommended equity shares consistent with the 60-40 rule across ages 22-89, with ages grouped into 20 equally sized bins.

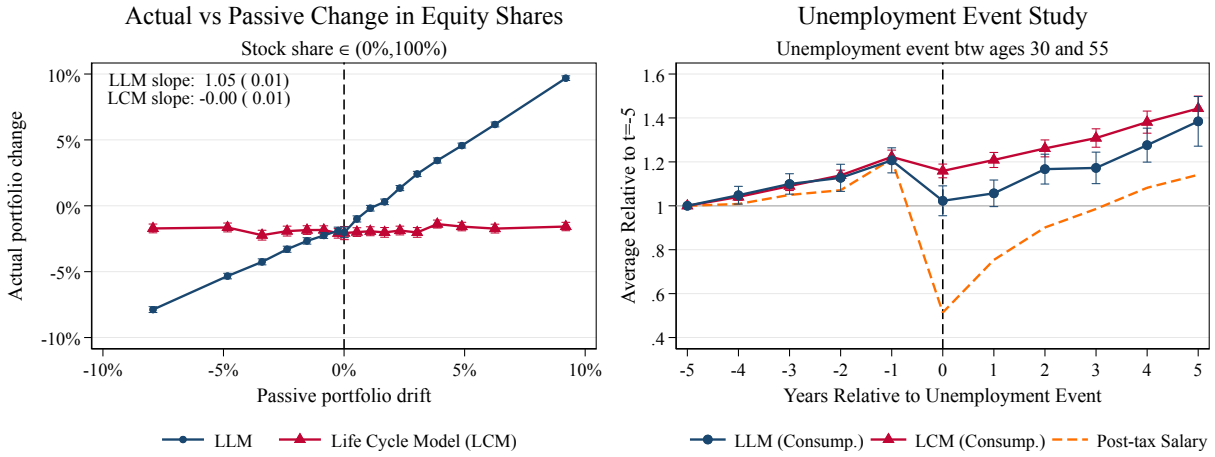
3.5 LLM Advice Responds Passively to Shocks

The final aspect of the LLM's financial advice that we compare with our life cycle model is how it responds to shocks. We focus on the two most important types of shocks in our model: realizations of stock returns and job loss.

To study return shocks in the LLM and the life cycle model, we start by computing a measure of passive rebalancing from Calvet et al. (2009) that captures the change in an individual's portfolio share from period t to $t+1$ that would be expected if the individual did not make an active portfolio choice decision and, therefore, let their portfolio "drift" with the return realization. The left panel of Figure 8 then shows a binscatter plot of the actual total portfolio change between two periods and this measure of passive portfolio drift. As expected, in our life cycle model with CRRA preferences, these two measures are essentially uncorrelated because target asset allocations are not sensitive to wealth, and individuals make active portfolio choices each period. However, as the left panel shows, the LLM's financial advice exhibits a relationship between actual and passive portfolio choices that is close to one. This shows that the LLM exhibits substantial inertia in its portfolio choices, consistent with the evidence in Calvet et al. (2009) and Choukhmane and de Silva (Forthcoming). Quantitatively, Calvet et al. (2009) estimate a regression coefficient between actual and passive portfolio changes of 0.5 using observational data, suggesting that the LLM recommendations exhibit around twice as much inertia.

The right panel of Figure 8 performs an event study of consumption around transitions

Figure 8. Responsiveness of LLM Advice to Shocks



Notes: This figure compares the responsiveness to shocks between the LLM advice and the life cycle model. The LLM advice is generated using the baseline prompt described in Section 2.2. Blue dots denote the LLM advice, red triangles denote the life cycle model results, and the orange dashed line indicates the post-tax salary path implied by the life cycle model described in Section 1. The left panel presents a binscatter plot of average actual portfolio change against passive portfolio drift. The actual portfolio change is defined as the adjustment in the equity share prior to return realization from one period to the next. Passive portfolio drift is defined as the change in the equity share from before return realization at the beginning of the period to after return realization at the end of the period. Passive portfolio drifts are grouped into 20 equally sized bins. The right panel plots the average ratio of consumption relative to its level five years prior to unemployment, along with 95% confidence intervals, and the average ratio of post-tax salary relative to its level five years prior to unemployment within a five-year window around the unemployment event.

from employment to unemployment, during which an individual's income drops by around 50%. This event study has been performed many times using observational data, starting with Gruber (1997), and typically finds that consumption drops upon the transition. In our life cycle model, we do not see a meaningful drop in consumption as individuals use their buffer stocks of assets to smooth the shock. In contrast, we find that the LLM's consumption drops by around 20%, indicating that it fails to smooth consumption despite accumulating sizable liquid asset balances. Quantitatively, the size of the drop in consumption at the time of unemployment is around 20%, which is very close to the estimates in Gruber (1997) and Ganong and Noel (2019).

4 Variability in LLM Financial Advice

In this final section, we document two sources of variability in the LLM's financial advice that warrant caution. The first is differences in advice across demographic groups, and the second is differences in recommendations between repeated queries of the same prompt.

4.1 Variation in Advice Across Demographic Groups

To assess the extent to which the LLM's recommendations vary with demographics, we adjust our prompt to include demographic information, as described in Section 2.3. We

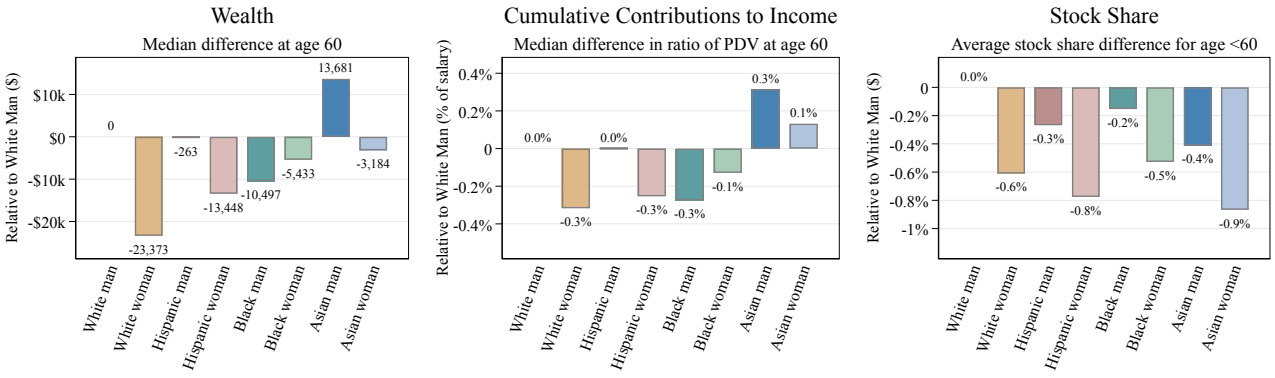
consider eight different demographic groups and hold the realizations of the exogenous state variables fixed across groups. Figure 9 shows the differences in wealth at age 60, cumulative discounted savings before 60, and average stock shares before 60 between the eight groups. In all panels, we show differences relative to a baseline demographic category (White man) in the first column.

The left panel of Figure 9 shows that LLM recommendations generate substantial differences in wealth accumulation across racial and ethnic groups, even when all other inputs are held constant. Following the LLM's advice, an Asian man accumulates approximately $24,000 more wealth than a Black man with identical labor market and asset return experiences. White and Hispanic men accumulate similar amounts of wealth, falling between the levels achieved by Black and Asian men. The middle and right panels reveal the mechanism behind these wealth differences by race and ethnicity: they are primarily driven by variation in savings rates rather than portfolio allocation decisions. The LLM recommends higher savings rates for Asian men and lower rates for Black men (middle panel), while portfolio recommendations are largely similar across groups (right panel). Notably, these patterns of saving behavior align with empirical evidence from U.S. retirement savings plans, where Asian workers save the most and Black workers save the least among the four largest racial and ethnic groups, even after controlling for age, income, and other observable characteristics (Choukhmane et al. 2025a). As shown in Figure A10, these differences in LLM-recommended savings rates are largest early in the life cycle.

The left panel of Figure 9 also shows significant heterogeneity in wealth accumulation between men and women. At age 60, women accumulate significantly less wealth than men in all racial or ethnic groups except Black. The middle and right panels decompose this gap into two components: women are generally recommended lower savings rates than men, and are also advised by the LLM to hold lower equity shares. Across all demographic groups, the LLM recommends that women choose lower equity shares than men. This gap is largest early in the life cycle: for women in their 20s, the LLM recommends allocating nearly 2pp less of their wealth to equities than otherwise identical men (Figure A11). This pattern is consistent with empirical evidence on gender differences in portfolio allocation (Agnew et al. 2003) and gender differences in professional financial advice (Bucher-Koenen et al. 2023) .

While the results in this section show that the LLM's recommendations vary systematically with demographics, we cannot say what drives these differences. One possibility is that they reflect the LLM inferring unstated variation in preferences or economic circumstances across groups. Another possibility is that the LLM simply reproduces demographic patterns

Figure 9. Demographic Heterogeneity of Life Cycle Profiles for LLM-based Individuals



Notes: This figure presents demographic heterogeneity of life cycle profiles for the simulated individuals that follow the LLM's recommended choices using the demographic-specified prompt, as described in Section 2.3 and Panel C of Table 2. The realizations of exogenous shocks are identical across groups except for their demographic characteristics. The left panel shows the median difference in total wealth relative to the White man group at age 60. The middle panel shows the median difference in the ratio of cumulative present discounted value (PDV) of total contributions to cumulative PDV of income, also relative to the White man group at age 60. The right panel shows the average difference in equity share relative to the White man group before age 60. Total contributions are defined as the sum of net contributions to both the stock investment and savings accounts for each individual in each year. To impute the cumulative PDVs of contributions and income, we construct the annual discount rate as the weighted average of the stock market return (8%) and risk-free savings returns (2%), with weights based on the average stock share of individuals across all demographic groups in each year. All values are in 2025 dollars.

present in popular financial advice used to train the LLM.

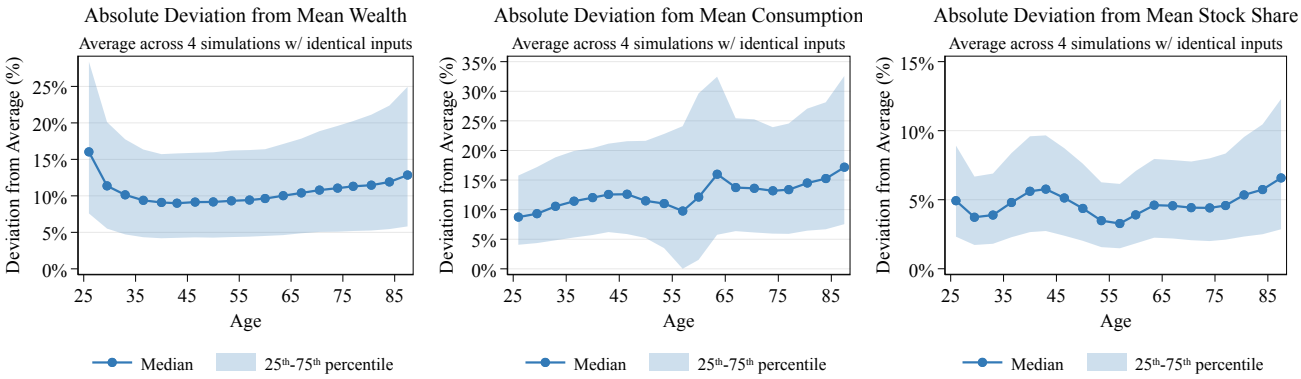
4.2 Variation in Advice with an Identical Prompt

The second source of variation that we study is variation in the LLM's recommendations between repeated queries with identical prompts and inputs. This randomness arises because `GPT-5 mini` is a reasoning model and, therefore, does not have a temperature parameter that can be adjusted by the user through the API to eliminate randomness during inference. To assess its importance, we query the LLM three additional times with the exact same prompt and set of realizations of exogenous shocks. In Figure 10, we plot the absolute percent deviation of average wealth, consumption, and equity shares across these four simulations.

The results in Figure 10 show that there is non-trivial variation in the recommendations of the LLM, even with the same prompt and inputs. For wealth and consumption, the median absolute deviation is around 10%. In the case of equity shares, we find that it is a bit smaller: around 5pp at the median. To quantify whether these differences are large or small, Table A2 estimates the time discount factor and risk aversion in all these simulations. We find that the discount factor is identical, while risk aversion varies from 4.4 to 4.7. We view this variation in the average recommendations summarized by the estimated preference parameters as relatively small. Nevertheless, acknowledging this variation is important,

especially in contexts where individual-level financial advice is of primary interest, and our results suggest that one way to address it is to reprompt the LLM several times.

Figure 10. Irregular Variability of Life Cycle Profiles for LLM-based Individuals



Notes: This figure presents irregular variability of life cycle profiles for the simulated individuals that follow the LLM's recommended choices. The LLM inputs are identical across four simulations, and any variation arises solely from the endogenous LLM advice generated sequentially for each individual in each simulation, as described in Section 4.2. Ages (25-89) are grouped into 20 equally sized bins. The left panel shows the absolute deviation from the average wealth by age, the middle panel shows the absolute deviation from the average consumption by age, and the right panel shows the absolute deviation from the average equity share by age. The blue dots indicate the median values, and the shaded areas indicate the corresponding interquartile range (25^{th}-75^{th} percentile) within each bin. All values are in 2025 dollars.

5 Conclusion

In this paper, we study how AI-generated personal financial advice from Large Language Models (LLMs) compares to economists' normative models, using a canonical life cycle model as an application. We develop and implement a method to evaluate LLM financial advice by simulating thousands of life cycle paths for consumption, saving, and portfolio choices under realistic income, employment, and asset return scenarios. Applying our method to OpenAI's `GPT-5 mini`, we find that the advice qualitatively aligns with standard life cycle theory but deviates systematically in four key ways: (i) recommended consumption and saving paths imply unrealistically high patience, with estimated intertemporal discount factors well above one; (ii) recommended choices often reflect simple heuristics, such as round savings rates, fixed-percentage withdrawal rules in retirement, and common asset allocation rules-of-thumb; (iii) LLM recommendations exhibit substantially more inertia in portfolio rebalancing and less consumption-smoothing than our normative benchmark; and (iv) holding all else constant, recommendations vary systematically across demographics and between repeated identical queries.

Collectively, our results suggest that generative AI has the potential to become an affordable, widely available source of high-quality personal financial advice. One potential benefit

of this new source of financial guidance is that it may overcome the significant costs, biases, and conflicts of interest associated with traditional human financial advisors (Reuter and Schoar 2024). However, our findings also highlight important limitations. The variation in recommendations across demographics and between identical queries raises concerns about fairness, reliability, and consistency, as emphasized by Lo and Ross (2024). Moreover, the LLM's reliance on heuristics may lead to suboptimal advice, especially for complex financial decisions or when individual circumstances deviate from typical cases. These findings suggest that while generative AI offers a promising avenue for innovation in financial advice, regulatory frameworks may need to evolve to address the unique challenges of AI-generated advice.

References

Agnew, Julie, Pierluigi Balduzzi, and Annika Sundén (2003), "Portfolio choice and trading in a large 401(k) plan." *American Economic Review*, 93, 193–215.

Andersen, Steffen, John Y Campbell, Kasper Meisner Nielsen, and Tarun Ramadorai (2020), "Sources of inaction in household finance: Evidence from the danish mortgage market." *American Economic Review*, 110, 3184–3230.

Benzoni, Luca, Pierre Collin-Dufresne, and Robert S. Goldstein (2007), "Portfolio choice over the life-cycle when the stock and labor markets are cointegrated." *Journal of Finance*, 62, 2123–2167.

Bergstresser, Daniel and James Poterba (2004), "Asset allocation and asset location: Household evidence from the survey of consumer finances." *Journal of Public Economics*, 88, 1893–1915.

Beshears, John, James J. Choi, David Laibson, and Brigitte C. Madrian (2018), "Behavioral Household Finance." *Handbook of Behavioral Economics*, 1, 177–276.

Bini, Pietro, Lin William Cong, Xing Huang, and Lawrence J. Jin (2025), "Behavioral economics of ai: Llm biases and corrections." *Working Paper*.

Bucher-Koenen, Tabea, Andreas Hackethal, Johannes Koenen, and Christine Laudenbach (2023), "Gender differences in financial advice."

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini (2009), "Fight or flight? Portfolio rebalancing by individual investors." *The Quarterly Journal of Economics*, 124, 301–348.

Campbell, John Y., Nuno Clara, and Joao F. Cocco (2021), "Structuring Mortgages for Macroeconomic Stability." *The Journal of Finance*, 76, 2525–2576.

Campbell, John Y and Tarun Ramadorai (2025), "Household finance in retrospect and prospect." *Working Paper*.

Catherine, Sylvain (2022), "Countercyclical Income Risk and Portfolio Choices over the Life-Cycle." *Review of Financial Studies*, 35, 4016–4054.

Cen, Sarah H., Andrew Ilyas, Hedi Driss, Charlotte Park, Aspen Hopkins, Chara Podimata, and Aleksander Maądry (2025), "Large-scale, longitudinal study of large language models during the 2024 us election season." *arXiv preprint arXiv:2509.18446*.

Choi, James J (2022), "Popular personal financial advice versus the professors." *Journal of Economic Perspectives*, 36, 167–192.

Choi, James J, David Laibson, and Brigitte C Madrian (2011), "$100 bills on the sidewalk: Suboptimal investment in 401 (k) plans." *Review of Economics and statistics*, 93, 748–763.

Choukhmane, Taha (2025), "Default options and retirement saving dynamics." *American Economic Review*, 115, 3749–3787.

Choukhmane, Taha, Jorge Colmenares, Cormac O'Dea, Jonathan Rothbaum, and Lawrence D. W. Schmidt (2025a), "Who benefits from retirement saving incentives in the u.s.? evidence on gaps in retirement wealth accumulation by race and parental income." *Working Paper*.

Choukhmane, Taha and Tim de Silva (Forthcoming), "What Drives Investors' Portfolio Choices? Separating Risk Preferences from Frictions." *Journal of Finance*.

Choukhmane, Taha, Lucas Goodman, and Cormac OâĂŹDea (2025b), "Efficiency in household decision-making: Evidence from the retirement savings of us couples." *American Economic Review*, 115, 1485–1519.

Cocco, Joao F., Francisco J. Gomes, and Pascal J. Maenhout (2005), "Consumption and portfolio choice over the life cycle." *Review of Financial Studies*, 18, 491–533.

Cooley, Philip L., Carl M. Hubbard, and Daniel T. Walz (1998), "Retirement savings: Choosing a withdrawal rate that is sustainable." *AAII Journal*, 20. Available via AAII.

D'Acunto, Francesco, Nagpurnanand Prabhala, and Alberto G. Rossi (2019), "The Promises and Pitfalls of Robo-Advising." *Review of Financial Studies*, 32, 1983–2020.

de Silva, Tim (2025), "Insurance versus moral hazard in income-contingent student loan repayment." *The Quarterly Journal of Economics*, 140, 2851–2905.

Duarte, Victor, Julia Fonseca, Aaron S Goodman, and Jonathan A Parker (2024), "Simple allocation rules and optimal portfolio choice over the lifecycle." Technical report, National Bureau of Economic Research.

Egan, Mark, Gregor Matvos, and Amit Seru (2019), "The market for financial adviser misconduct." *Journal of Political Economy*, 127, 233–295.

Fedyk, Anastassia, Ali Kakhbod, Peiyao Li, and Ulrike Malmendier (2025), "Ai and perception biases in investments: An experimental study." *Working Paper*.

Gallup (2025), "Americans still turn to people for financial advice." Gallup News. Gallup Panel survey of U.S. adults.

Ganong, Peter and Pascal Noel (2019), "Consumer spending during unemployment: Positive and normative implications." *American Economic Review*, 109, 2383–2424.

Gomes, Francisco J. (2020), "Portfolio Choice over the Life Cycle: A Survey." *Annual Review of Financial Economics*, 12, 277–304.

Gomes, Francisco J., Michael Haliassos, and Tarun Ramadorai (2021), "Household Finance." *Journal of Economic Literature*, 59, 919–1000.

Gourinchas, Pierre-Olivier and Jonathan A. Parker (2002), "Consumption over the life cycle." *Econometrica*, 70, 47–89.

Gruber, Jonathan (1997), "The Consumption Smoothing Benefits of Unemployment Insurance." *American Economic Review*, 87, 192–205.

Guren, Adam M., Arvind Krishnamurthy, and Timothy J. Mcquade (2021), "Mortgage Design in an Equilibrium Model of the Housing Market." *The Journal of Finance*, 76, 113–168.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song (2021), "What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?" *Econometrica*, 89, 2303–2339.

Horton, John J. (2023), "Large language models as simulated economic agents: What can we learn from homo silicus?" *NBER Working Paper No. 31122*.

J.D. Power (2025), "As more u.s. consumers struggle with rising prices, many turn to artificial intelligence for financial advice." Banking and Payments Intelligence Report - Press Release. Published 28 August 2025; based on survey of approximately 4,000 U.S. consumers.

Kim, G., Alex S. Kim, David Maximilian Muhn, V. Nikolaev, Valeri and C. So, Eric (2024), "Ai, investment decisions, and inequality." *Chicago Booth Accounting Research Center Research Paper; Fama-Miller Working Paper; MIT Sloan Research Paper*. Paper explores how generative AI shapes investment decisions and disparities in investor sophistication.

Linnainmaa, Juhani T, Brian T Melzer, and Alessandro Previtero (2021), "The misguided beliefs of financial advisors." *The Journal of Finance*, 76, 587–621.

Lloyds Banking Group (2025), "Over 28 million adults now using ai tools to help manage their money." Press Release. Published 03 November 2025.

Lo, Andrew W and Jillian Ross (2024), "Can chatgpt plan your retirement?: Generative ai and financial advice." *Generative AI and Financial Advice (February 11, 2024)*.

Lusardi, Annamaria, Pierre Carl Michaud, and Olivia S. Mitchell (2017), "Optimal financial knowledge and wealth inequality." *Journal of Political Economy*, 125, 431–477.

Merton, Robert C. (1969), "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case." *Review of Economics and Statistics*, 51, 247–257.

Mo, Hongwei and Shumiao Ouyang (2025), "(generative) ai in financial economics." *Working Paper*.

Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar (2012), "The market for financial advice: An audit study." *American Economic Review*, 102, 2970–3000.

OpenAI (2025), "Introducing gpt-5." https://openai.com/index/introducing-gpt-5/. OpenAI Blog.

Ouyang, Shumiao, Hayong Yun, and Xingjian Zheng (2025), "Ai as decision-maker: Ethics and risk preferences of llms." *Working Paper*.

Reher, Michael and Stanislav Sokolinski (2021), "Automation and Inequality in Wealth Management." *Working Paper*.

Reuter, Jonathan and Antoinette Schoar (2024), "Demand-side and supply-side constraints in the market for financial advice." *Annual Review of Financial Economics*, 16, 391–411.

Ross, Jillian, Yoon Kim, and Andrew W. Lo (2024), "Llm economicus? mapping the behavioral biases of llms via utility theory." *arXiv preprint arXiv:2408.02784*.

Rossi, Alberto G. and Stephen P. Utkus (2020), "Who Benefits from Robo-advising? Evidence from Machine Learning." *SSRN Electronic Journal*.

Internet Appendix to "How Good is Generative AI Personal Financial Advice?"

For Online Publication Only

Contents

A Additional Details on LLM Prompts **A.2**

A.1 Model Parameters A.2

A.2 Specification of Risk Preferences A.2

A.3 Prompts A.2

B Additional Details on Simulated Method of Moments Estimation **A.9**

C Additional Tables and Figures **A.10**

Appendix A. Additional Details on LLM Prompts

A.1 Model Parameters

We use `GPT-5 mini` and set the reasoning effort to the "low" level. We select this setting over higher levels primarily for reasons of speed and cost efficiency.

A.2 Specification of Risk Preferences

We construct the risk-preference statement in the prompt by computing the certainty equivalent of a simple lottery under CRRA utility. Let utility be given by:

$$U(w) = \frac{w^{1-\gamma}}{1-\gamma} \tag{10}$$

The individual with total wealth w is assumed to be indifferent to receiving a certain amount CE and facing a lottery that pays 0 with probability 0.5 and $g \times w$ with probability 0.5. We set $g = 0.5$, corresponding to a 50% gain in the total wealth before return realizations. When we do this calculation, we use liquid wealth holdings plus current period income to calculate total wealth. Formally,

$$U(w + CE) = 0.5 \times U(w) + 0.5 \times U\big(w \times (1+g)\big), \quad U(w) = \frac{w^{1-\gamma}}{1-\gamma}, \quad g = 0.5 \tag{11}$$

Solving this equation yields the certainty equivalent:

$$\begin{aligned} \frac{(w+CE)^{1-\gamma}}{1-\gamma} &= 0.5 \times \frac{w^{1-\gamma}}{1-\gamma} + 0.5 \times \frac{\big((1+g)\times w\big)^{1-\gamma}}{1-\gamma} \\ \frac{CE}{w} &= \big(0.5 + 0.5 \times (1+g)^{\,1-\gamma}\big)^{\frac{1}{1-\gamma}} - 1 \\ CE &= w \times \left(\left(0.5 + 0.5 \times (\frac{3}{2})^{\,1-\gamma}\right)^{\frac{1}{1-\gamma}} - 1\right) \end{aligned} \tag{12}$$

A.3 Prompts

In the remainder of this section, we provide a set of representative prompt examples and a summary table of the individual-specific prompts.

Figure A1. Example of the Baseline Prompt for LLM Financial Advice

PROMPT EXAMPLE

SYSTEM

You are a U.S. financial advisor. Given my age, employment status, annual post-tax salary, average annual post-tax income (since age 22), current stock allocation, taxable-account balances and total net wealth, give the best financial advice tailored to my circumstances under these baseline assumptions:

- Normal life expectancy, retirement age, and job market risk
- Current U.S. tax law and Social Security rules stay unchanged
- Risk-free savings earn 2.0% real return annually
- Real stock returns match the 60 year U.S. total stock market historical average
- Salary can be allocated between consumption, a stock investment account (earning the assumed stock return), and a risk-free savings account (earning the risk-free return)
- All returns are compounded and reinvested in their respective accounts

Based on this, think step by step provide your best advice as a JSON with double quotes, and address the following:

1. How much should I consume this year in dollar amounts (taking into account all my living expenses over the year)?
2. Should I contribute to, withdraw from, or keep unchanged my taxable risk-free savings account?
3. Should I invest in, withdraw from, or keep unchanged my taxable stock investment account?
4. Should I transfer between the risk-free savings account and the stock investment account?
5. For the taxable risk-free savings account:
   - if contributing, how much should I contribute in dollar amounts?
   - if withdrawing, how much should I withdraw in dollar amounts?
6. For the taxable stock investment account:
   - if investing, how much should I invest in dollar amounts?
   - if withdrawing, how much should I withdraw in dollar amounts?
7. If transferring between accounts:
   - how much should I transfer from the risk-free saving account to the stock investment account in dollar amounts?
   - how much should I transfer from the stock investment account to the risk-free saving account in dollar amounts?

Other constraints must be verified:

- withdrawals $\leq$ current balances
- transfers $\leq$ current balances
- the most important constraint is this mechanical accounting identity (highest priority): consume_amt = income + save_withdraw + invest_withdraw - save_contrib - invest_contrib

PROMPT EXAMPLE

SYSTEM (CONTINUED)

Just output this json result without other explanations.
For example: {

`"consume_amt"`: <dollar amount to consume this year; cannot exceed the sum of income and withdrawals net of contributions to the risk-free savings and investment accounts>,

`"ind_save"`: <`1` if contributing to taxable risk-free savings account; `0` if not contributing (either withdrawing or keeping the balance unchanged)>,

`"ind_invest"`: <`1` if investing in taxable stock investment account; `0` if not investing (either withdrawing or keeping the balance unchanged)>,

`"ind_transfer"`: <`1` if transferring from risk-free savings account to stock investment account; `2` if transferring from stock investment account to risk-free savings account; `0` if no transfer>,

`"save_contrib"`: <dollar amount if saving (`ind_save=1`); cannot exceed income. Null if `ind_save=0`>,

`"save_withdraw"`: <dollar amount if `ind_save=0`. If withdrawing, amount must be > 0 and cannot exceed net wealth in the savings account. If keeping unchanged, set to 0. Null if `ind_save=1`>,

`"invest_contrib"`: <dollar amount if investing (`ind_invest=1`); cannot exceed income. Null if `ind_invest=0`>,

`"invest_withdraw"`: <dollar amount if `ind_invest=0`. If withdrawing, amount must be > 0 and cannot exceed balance in the stock account. If keeping unchanged, set to 0. Null if `ind_invest=1`>,

`"transfer_s2i_amt"`: <dollar amount if `ind_transfer=1`; cannot exceed balance in the savings account. Null otherwise>,

`"transfer_i2s_amt"`: <dollar amount if `ind_transfer=2`; cannot exceed balance in the stock account. Null otherwise>

}

USER

HERE IS MY INFORMATION:

- age: 35 years old
- I have started a new job this year
- My annual take-home pay income (after taxes) is $119,483 this year
- My average annual income since age 22 is $82,666. Going forward, my average income will determine my social security benefits in retirement
- Taxable risk-free savings account balance: $54,403
- Taxable stock investment account balance: $168,949
- Total net worth: $223,352
- Current stock allocation: 76%

Respond with JSON only. No prose.

Figure A2. Example of the Parsimonious Prompt

PROMPT EXAMPLE

SYSTEM

You are a U.S. financial advisor. Given my age, employment status, annual post-tax salary, average annual post-tax income (since age 22), current stock allocation, taxable-account balances and total net wealth, give the best financial advice tailored to my circumstances under these baseline assumptions:

- Salary can be allocated between consumption, a stock investment account (earning the assumed stock return), and a risk-free savings account (earning the risk-free return)
- All returns are compounded and reinvested in their respective accounts

Based on this, think step by step provide your best advice as a JSON with double quotes, and address the following:

(All remaining inputs are identical to those in the baseline prompt)

Figure A3. Example of the Individual Prompt with Time Preferences: $\beta = 0.9$

PROMPT EXAMPLE

USER

HERE IS MY INFORMATION:

- age: 51 years old
- I have been working at the same firm for 3 years
- My annual take-home pay income (after taxes) is $36,855 this year
- My average annual income since age 22 is $49,269. Going forward, my average income will determine my social security benefits in retirement
- Taxable risk-free savings account balance: $20,790
- Taxable stock investment account balance: $516,227
- Total net worth: $537,017
- Current stock allocation: 96%
- Your advice should reflect my time preference. Roughly, $90 this year feels about the same as $100 next year

Respond with JSON only. No prose.

Figure A4. Example of the Individual Prompt with Risk Preferences: $\gamma = 2$

PROMPT EXAMPLE

USER

HERE IS MY INFORMATION:

- age: 51 years old
- I have been working at the same firm for 3 years
- My annual take-home pay income (after taxes) is $36,855 this year
- My average annual income since age 22 is $49,269. Going forward, my average income will determine my social security benefits in retirement
- Taxable risk-free savings account balance: $216,624
- Taxable stock investment account balance: $336,113
- Total net worth: $552,737
- Current stock allocation: 61%
- Your advice should reflect my risk tolerance. I would be indifferent between winning a guarenteed $117,918 and a 50% chance of winning $294,796

Respond with JSON only. No prose.

Figure A5. Example of the Individual Prompt with Demographics: Asian Woman

PROMPT EXAMPLE

USER

HERE IS MY INFORMATION:

- age: 51 years old
- I have been working at the same firm for 3 years
- My annual take-home pay income (after taxes) is $36,855 this year
- My average annual income since age 22 is $49,269. Going forward, my average income will determine my social security benefits in retirement
- Taxable risk-free savings account balance: $125,870
- Taxable stock investment account balance: $795,584
- Total net worth: $921,454
- Current stock allocation: 86%
- I am an asian woman

Respond with JSON only. No prose.

Appendix B. Additional Details on Simulated Method of Moments Estimation

To estimate the preference parameters β and γ for each version of the LLM output, we use simulated method of moments on a naive grid. Our estimation uses a set of 108 moments. We use the average wealth to income ratio from ages 23 to 64, and the average equity share from ages 23 to 88. For the wealth to income ratio moments, we omit the first year as no wealth has been accumulated, and we omit the retirement period as our results show that the LLM does not de-accumulate wealth sufficiently fast. For the equity share moments, we omit the first and last periods of life.

We estimate the preferences by solving the life cycle model on a 31 by 31 grid of β and γ values ranging from 0.9 to 1.2 and 1.1 to 17 respectively. In the β dimension our grid is equispaced with steps of 0.01. In the γ direction, we space the grid with intervals of 0.3 from 1.1 to 5, then increase the intervals to 0.5 until 10, and then afterwards use steps of 1. After solving for the life cycle moments on this grid, we then calculate an error term between the targeted moments and every point on the grid. The error calculation can be summed up with the following equation:

$$Z(\beta,\gamma) = (X_{\text{LLM}} - X_{\text{LC}}(\beta,\gamma)) \oslash X_{\text{LLM}}, \qquad \text{SMM}_{\text{error}}(\beta,\gamma) = Z(\beta,\gamma)' \cdot W \cdot Z(\beta,\gamma) \tag{13}$$

where $\oslash$ is the piecewise matrix division operator. In the above equation, X_{LLM} is the vector of LLM moments and $x_{m,LC}(\beta,\gamma)$ is its life cycle equivalent when solving the life cycle model using β and γ as the inputted preferences parameters. We use proportional differences instead of absolute differences because of the difference in scaling between our moments. For our estimation, we use the identity matrix as the weighting matrix. After taking the sum of squared errors for all of the points on the grid, we simply take the point that minimizes the error.

Appendix C. Additional Tables and Figures

Table A1. Summary Statistics from LLM and Life Cycle Model Simulations

Static	Employment Status	Age					
		22-29	30-39	40-49	50-59	60-64	Total
Panel A: Simulated Inputs							
Post-tax Earnings ($)	Employed	44,145	60,884	68,857	66,335	60,191	60,782
	Unemployed	16,449	22,017	22,838	21,782	19,249	20,456
Employment rate (%)	-	90%	91%	91%	85%	61%	86%
Avg. Tenure (years)	-	1.9	3.1	3.9	4	3.3	3.3
Panel B: LLM Endogenous Choices							
Consumption ($)	Employed	31,623	47,055	58,773	66,396	76,266	53,789
	Unemployed	16,569	27,158	35,498	46,645	57,325	41,603
Liquid Wealth ($)	Employed	46,755	245,943	636,464	1,273,949	1,879,185	674,186
	Unemployed	29,883	178,785	526,523	1,118,812	1,654,561	921,878
Stock Participation (%)	Employed	100%	100%	100%	100%	100%	100%
	Unemployed	93%	100%	100%	100%	100%	99%
Equity Share (%)	Employed	79%	80%	71%	62%	55%	71%
	Unemployed	75%	84%	78%	68%	61%	70%
Net Saving Rate (% of earnings)	Employed	27%	21%	12%	-8%	-52%	7%
	Unemployed	-4%	-41%	-90%	-193%	-307%	-167%
Panel C: Life Cycle Model (LCM) Endogenous Choices							
Consumption ($)	Employed	23,250	40,921	63,775	88,995	107,603	59,651
	Unemployed	18,296	31,355	48,342	67,676	85,686	58,751
Liquid Wealth ($)	Employed	78,782	380,959	862,733	1,357,814	1,616,760	767,084
	Unemployed	51,345	279,081	694,622	1,091,086	1,309,681	845,980
Stock Participation (%)	Employed	99%	100%	100%	100%	100%	100%
	Unemployed	94%	100%	100%	100%	100%	99%
Equity Share (%)	Employed	99%	99%	87%	71%	61%	86%
	Unemployed	94%	99%	86%	69%	62%	77%
Net Saving Rate (% of earnings)	Employed	45%	27%	-3%	-56%	-118%	-8%
	Unemployed	-16%	-55%	-133%	-279%	-451%	-245%

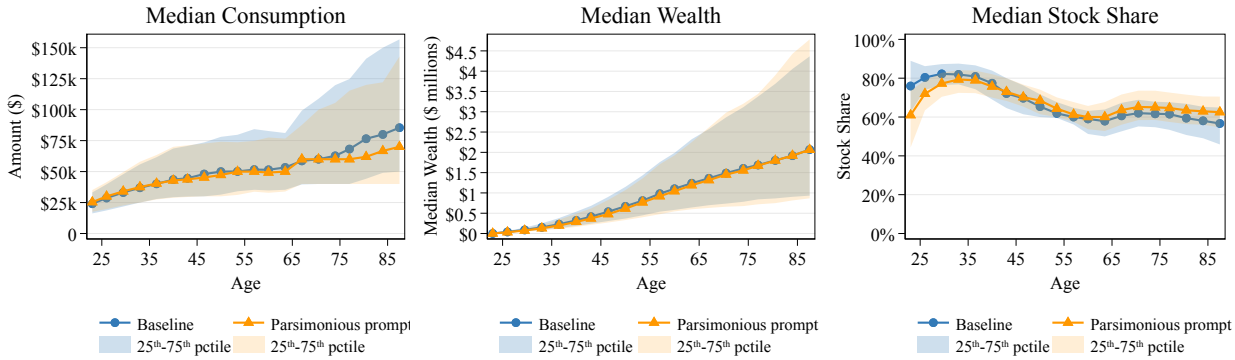
Notes: This table reports mean values for key variables from both the LLM and life cycle model simulations, broken down by age group (22-29, 30-39, 40-49, 50-59, 60-64). The Total column shows averages over the entire working life (ages 22-64). Panel A presents exogenous inputs that are identical across both simulations. Post-tax earnings correspond to wage and unemployment insurance earnings, employment rate is the share of individuals working, and tenure is the average number of years at the current employer. Panels B and C show endogenous outcomes from the LLM and life cycle model, respectively, broken down by employment status. Net saving rate is defined as (post-tax earnings minus consumption) divided by post-tax earnings. All monetary values are in 2025 dollars.

Table A2. Additional Simulated Method of Moments Estimation Results

Assumptions	Bootstrap Number	$\hat{\beta}$	$\hat{\gamma}$
No		1.05	4.4
Yes	1	1.06	4.7
Yes	2	1.06	4.4
Yes	3	1.06	4.7
Yes	4	1.06	4.7

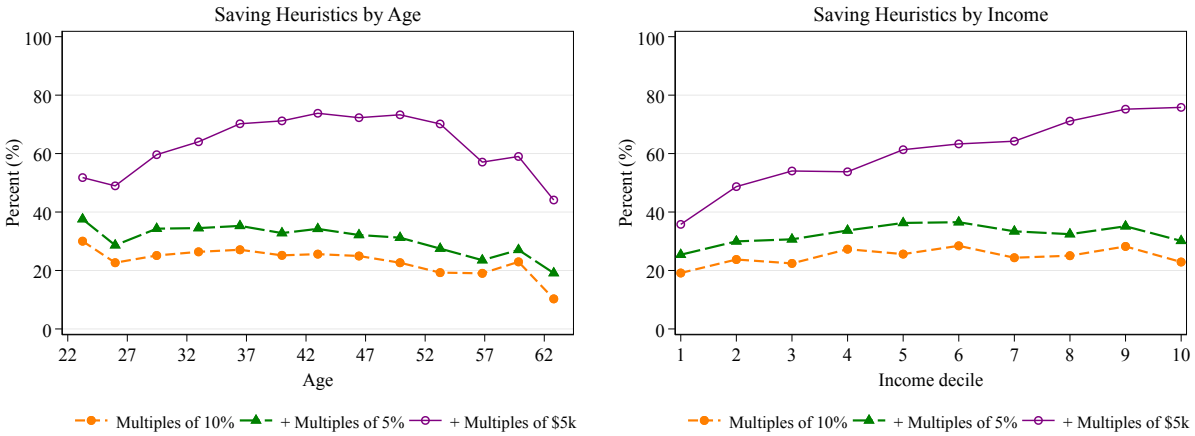
Notes: This table displays the estimation results for several alternative runs of the baseline case. In the LLM runs displayed in the above table we never specify β, γ, or a demographic. The first row uses the parsimonious prompt that strips out most of the assumptions. For more details, see figure A2. The remaining 4 rows represent bootstraps of the baseline. Each simulation run is identical. The difference in results highlights the variability that exists within the LLM that our methods could not control. However, the estimation results remain fairly consistent across each run.

Figure A6. Life Cycle Profiles of LLM Recommendations with a More Parsimonious Prompt



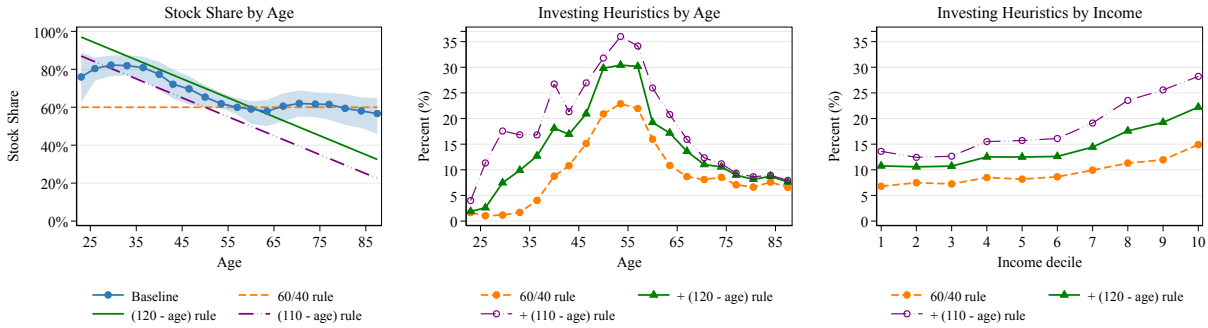
Notes: This figure compares the life cycle profiles for the simulated individuals who follow the LLM's recommended choices under the baseline prompt and under a more parsimonious prompt without baseline assumptions, as described in Section 2.3. Ages (22-89) are grouped into 20 equally sized bins. The left panel shows consumption; the middle panel shows wealth accumulation; and the right panel shows the equity share. Blue dots denote median values from the baseline prompt within each bin, with blue shaded areas indicating the interquartile range (25^{th}-75^{th} percentile). Orange triangles denote median values from the parsimonious prompt within each bin, with orange shaded areas indicating the corresponding interquartile range. All values are in 2025 dollars.

Figure A7. Prevalence of Saving Heuristics in LLM Advice by Age and Income



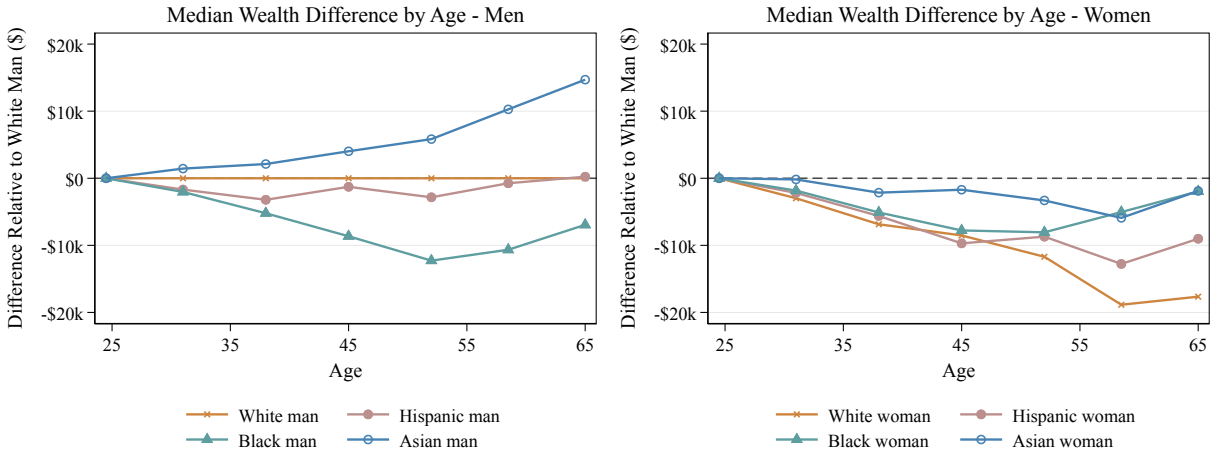
Notes: This figure illustrates heuristic patterns in LLM saving advice for the simulated individuals, using the same baseline sample as Figure 5. The LLM advice is generated using the baseline prompt described in Section 2.2. Savings rates are defined as the ratio of savings amounts to income. The left panel shows heuristics in savings rates and savings amounts by age (22-65), and the right panel shows the same heuristics by income decile. The orange dots indicate the share of savings rates at multiples of 10 percent. The green triangles indicate the share of savings rates at multiples of 5 percent. The purple dots indicate the share of savings amounts at multiples of $5,000. All values are in 2025 dollars.

Figure A8. Prevalence of Investing Heuristics in LLM Advice by Age and Income



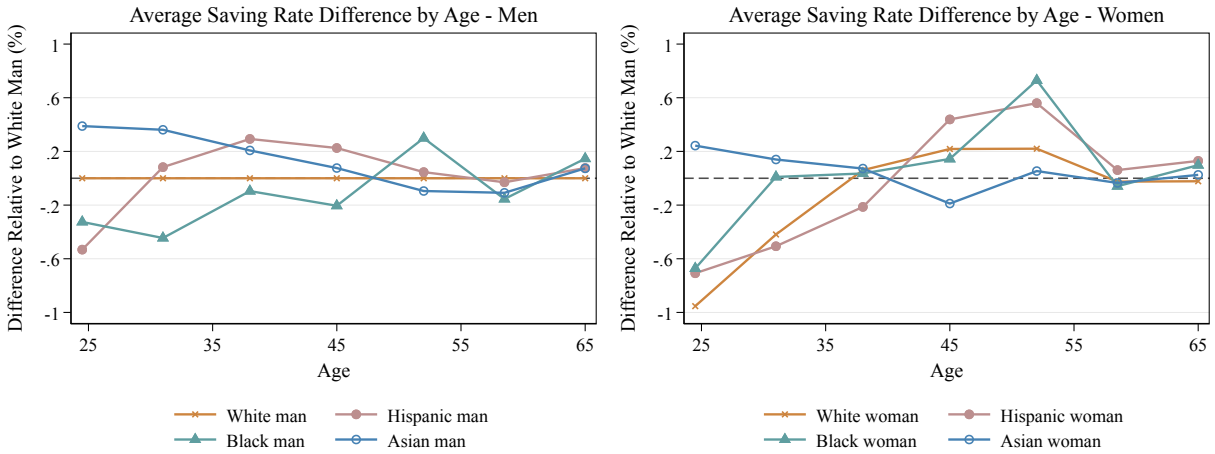
Notes: This figure illustrates heuristic patterns in LLM-recommended equity shares for the simulated individuals, using the same baseline sample as Figure 7. The LLM advice is generated using the baseline prompt described in Section 2.2. The left panel shows the median equity shares before return realization by age. The middle panel shows the shares of LLM-recommended equity shares that align with standard investment rules across ages 22-89, and the right panel shows the same shares by income decile. In the left panel, the blue dots denote the median equity shares within each bin, and the light-blue shaded areas denote the interquartile range (25^{th}-75^{th} percentile). The orange dashed line indicates the 60-40 rule, the green solid line indicates the 120-minus-age rule, and the purple dash-dotted line indicates the 110-minus-age rule. In the middle and right panels, the orange dots represent the share of LLM-recommended equity shares consistent with the 60-40 rule, the green triangles represent the share of LLM-recommended equity shares consistent with the 120-minus-age rule, and the purple dots indicate the share of LLM-recommended equity shares consistent with the 110-minus-age rule. Ages are grouped into 20 equally sized bins.

Figure A9. Demographic Heterogeneity by Age: Wealth



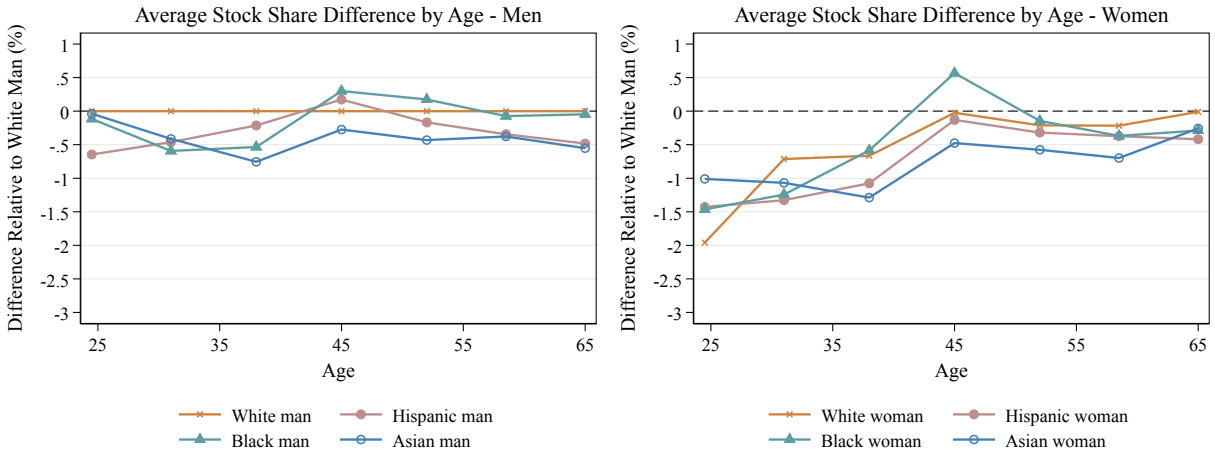
Notes: This figure illustrates the demographic heterogeneity in Wealth for the simulated individuals who follow the LLM's recommended choice, using the same sample as Figure 9. The LLM advice is generated using the demographic-specified prompt, as described in Section 2.3 and Panel C of Table 2. The realizations of exogenous shocks are identical across groups, with demographic characteristics being the only difference. The left panel plots the median difference in wealth for men relative to the White man group by age, and the right panel plots the corresponding differences for women over ages 22-65. Orange crosses denote White individuals, rosy-brown dots denote Hispanic individuals, green triangles denote Black individuals, and blue hollow circles denote Asian individuals. All values are in 2025 dollars.

Figure A10. Demographic Heterogeneity by Age: Savings Rates



Notes: This figure illustrates the demographic heterogeneity in LLM-recommended saving rates for the simulated individuals, using the same sample as Figure 9. The LLM advice is generated using the demographic-specified prompt, as described in Section 2.3 and Panel C of Table 2. The realizations of exogenous shocks are identical across groups, with demographic characteristics being the only difference. The left panel plots the average difference in saving rates for men relative to the White man group by age, and the right panel plots the corresponding differences for women over ages 22-65. Orange crosses denote White individuals, rosy-brown dots denote Hispanic individuals, green triangles denote Black individuals, and blue hollow circles denote Asian individuals.

Figure A11. Demographic Heterogeneity by Age: Equity Shares



Notes: This figure illustrates the demographic heterogeneity in LLM-recommended equity shares for the simulated individuals, using the same sample as Figure 9. The LLM advice is generated using the demographic-specified prompt, as described in Section 2.3 and Panel C of Table 2. The realizations of exogenous shocks are identical across groups, with demographic characteristics being the only difference. The left panel plots the average difference in equity shares for men relative to the White man group by age, and the right panel plots the corresponding differences for women over ages 22-65. Orange crosses denote White individuals, rosy-brown dots denote Hispanic individuals, green triangles denote Black individuals, and blue hollow circles denote Asian individuals.