AI Financial Advice: Supply, Demand, and Life Cycle Implications*

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

March 20, 2026

Latest Version

Abstract

We develop and implement a novel method to study personal financial advice from Large Language Models (LLMs). Studying this advice is challenging because it depends on the model used (i.e., supply), the questions individuals ask (i.e., demand), and their evolving circumstances. We address these challenges by surveying a representative sample of adults and asking them to write prompts seeking spending and investing advice from an LLM. We then simulate the lifetime paths that result from following this advice under realistic asset and labor market conditions. Applying our method to `GPT-5.2` and `Gemini 3 Flash`, we document three facts about AI-generated financial advice. First, following LLM advice would move most survey respondents closer to the prescriptions of life cycle theory relative to their current behavior, including broader participation in diversified equity funds, equity shares that decline with age, and sizeable saving buffers. Second, replacing individual-written prompts with academic prompts moves LLM advice even closer to life cycle theory, with better consumption smoothing and less reliance on simple heuristics. Third, LLM advice varies systematically with individual characteristics, such as gender and financial literacy. These differences accumulate over the life cycle into wealth differences at retirement of 4-5% between groups and reflect both demand (i.e., systematic variation in the prompts written by different individuals) and supply (i.e., differences in advice for a given prompt). These facts highlight the potential of generative AI to improve financial decision-making, but suggest that its impact is likely heterogeneous across households and depends on how the technology is used.

*First draft: November 2025. We thank James Choi, Antonio Coppola, Darrell Duffie, Fiona Greig, Jonathan Parker, Eric So, Chris Tonetti, and seminar participants at MIT Sloan and Vanguard for helpful comments and feedback. An earlier version of this paper was circulated with the title "How Good is Generative AI Personal Financial Advice?".

†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). Industry surveys suggest that over half of adults in the U.S. and the U.K. have used LLMs for personal 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). This rapid growth suggests that the promise of affordable, widely accessible financial advice may be closer than ever, but also highlights the need to understand what financial advice LLMs provide and why it varies across groups of individuals.

In this paper, we develop and implement a method to describe quantitatively the personal financial advice provided by LLMs. Studying this advice systematically is challenging because it depends on the questions individuals ask (i.e., demand), features of the LLM itself (i.e., supply), and individuals' circumstances, which evolve dynamically based on past decisions and are subject to significant uncertainty. We overcome these three challenges by (i) surveying a nationally representative sample to elicit realistic prompts seeking financial advice from an LLM, (ii) simulating the quantitative implications of the advice received in response to these prompts over the life cycle under realistic asset and labor market conditions, and (iii) evaluating differences in the advice across heterogeneous groups of individuals and manipulating prompts to isolate the role of supply from that of demand.

Applying our method to `GPT-5.2` and `Gemini 3 Flash`, we document three facts about AI-generated financial advice. First, following LLM advice would move most survey respondents closer to the prescriptions of life cycle theory relative to their current behavior, including broader participation in diversified equity funds, equity shares that decline with age, and sizeable saving buffers. Second, replacing individual-written prompts with academic prompts moves LLM advice even closer to life cycle theory, with better consumption smoothing and less reliance on simple heuristics. Third, LLM advice varies systematically with individual characteristics, such as gender and financial literacy. These differences accumulate over the life cycle into wealth differences at retirement of 4-5% between groups and reflect both demand (i.e., variation in the prompts written by different individuals) and supply (i.e., differences in advice for a given prompt). For example, when randomizing gender labels, we find that two-thirds of the gender difference in investment advice is accounted for by men and women writing different prompts (i.e., demand), while one-third stems from the model giving different advice to the same questions when labeled as coming from men versus women (i.e., supply).

Taken together, our results highlight the potential of generative AI to improve financial decision-making, but suggest that its impact likely varies across individuals. While our results apply to specific LLMs at one point in time, several aspects of our contribution are

likely to persist. First, much of the heterogeneity that we document stems from differences in the questions people ask rather than limitations of the models themselves, suggesting that the demand side may remain a binding constraint on the quality of AI financial advice even as models improve. Second, we provide a general and scalable method to study LLM financial advice that can be reapplied as both LLMs and the way in which people interact with them evolve. Third, we propose a set of diagnostic tests grounded in life cycle theory (e.g., consumption smoothing, diversification, portfolio rebalancing) that can be used to evaluate and improve future generations of GenAI personal financial advice tools.

Methodology. Our method for studying LLM financial advice proceeds in 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 that in Choukhmane and de Silva (2026) and features liquid safe and risky assets, income and mortality risk, job-to-job and 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 across four asset classes: fixed income, diversified stock investments, individual stock holdings, and a "residual" risky asset that captures holdings in other cryptocurrencies, metals, or collectibles.

In the second step, we survey a demographically representative sample of 1,000 U.S. adults recruited through Prolific. We ask respondents to write prompts describing their financial situation for an LLM and asking for spending and investing advice. A key challenge in studying LLM advice is that it depends on demand (i.e., the prompt that is provided). Our survey, to the best of our knowledge, provides the first systematic evidence on the questions individuals ask when seeking financial advice from LLMs. In addition to the prompts, we survey individuals about their demographics, time and risk preferences, financial situation, financial literacy, and prior experience seeking financial guidance from AI.

The final step in our method is to simulate the life cycle paths of individuals who follow the advice provided by an LLM in response to our survey prompts. Specifically, at each age, we take a simulated individual characterized by their earnings, asset holdings, and employment status, and randomly draw a prompt from survey respondents in a similar age, income, and employment status bin. We then replace the state variables in the drawn prompt with those of the simulated individual and submit two successive LLM queries: first, we submit the prompt and receive textual financial advice; then, we use the LLM to translate this textual advice into quantitative recommendations for saving, consumption, and asset allocations. We repeat this process at each age over the individual's life cycle

and across many simulated individuals. Importantly, each query to the LLM is independent, with no memory of past responses; the only link across periods is the evolution of state variables that depend on prior choices.

Main findings. Our main results come from applying this method to `GPT-5.2`, "the LLM". We chose this LLM because ChatGPT is the most commonly used model for financial advice in our survey (and others, e.g., Lloyds Banking Group 2025; J.D. Power 2025) and `GPT-5.2` was the frontier model at the time we ran our analysis. We also perform our analysis using `Gemini 3 Flash` and obtain quantitatively similar results, which are discussed in the main text.

After describing the qualitative content of both the prompts written by survey respondents and the textual advice they receive from LLMs, we turn to the quantitative properties of the simulated advice. In doing so, we map the rich textual advice, which often references specific spending categories and commercial products, to the key outcomes in life cycle theory: consumption, saving, and portfolio allocation across broad asset classes.

We document three facts about the quantitative implications of LLM advice over the life cycle. First, following LLM recommendations would move most individuals closer to the broad principles of life cycle theory relative to their current behavior. While many individuals, across all age groups, report having little financial wealth accumulated, following LLM advice would generate sizeable saving buffers for virtually all individuals above age 30. In terms of portfolio choices, following LLM advice would lead over 99% of individuals to participate in the stock market, with equity shares declining with age past 45, patterns more in line with standard normative models than with their actual behavior. Most of the LLM's risky-asset allocation is in diversified equity funds, with low single-digit allocations to less diversified assets such as individual stocks, gold, or cryptocurrency.

Second, we find that the LLM's advice diverges quantitatively from more subtle prescriptions of life cycle theory. LLM recommendations rely heavily on simple saving and withdrawal heuristics and show little adjustment to income or return shocks. For example, consumption is not smoothed over the life cycle, with too little decumulation in retirement, nor in response to job loss, with consumption dropping too sharply even when simulated individuals have liquid wealth. Similarly, recommended asset allocations drift passively with realized returns rather than actively rebalancing. These patterns reflect, in part, the questions people ask. To isolate this channel, we compare our baseline results to those obtained using a structured "academic" prompt that asks the LLM to give regulated professional financial advice, references life cycle planning and modern portfolio theory, and provides explicit information about all of the simulated individual's state variables and

explicit assumptions about the economic environment. We find that using this academic prompt moves the LLM's advice closer to the predictions of life cycle theory by improving consumption smoothing and reducing reliance on heuristics, but does not increase the amount of active portfolio rebalancing.

Third, we find that LLM advice varies systematically across groups of individuals, defined by gender, financial literacy, and past LLM experience. For each characteristic, we split survey respondents into two groups and repeat our simulation, sampling only prompts written by individuals in each group (e.g., simulating life cycle paths using only prompts written by men versus women). Advice varies meaningfully along each dimension: average wealth at retirement is around 5% higher when sampling prompts written by men, individuals with high financial literacy, or those with past LLM experience. These differences arise because the LLM consistently recommends lower equity shares for prompts written by women or individuals with low financial literacy, and lower saving rates for prompts written by individuals with no prior experience using AI for financial advice.

To better understand the variation in LLM advice across individuals, we develop an approach to decompose these differences into demand and supply. Focusing on the heterogeneity by gender, we insert randomized gender labels (e.g., "I am a man") into the subset of prompts that lack explicit gender identifiers. This allows us to decompose gender differences in advice into a demand—men and women ask different questions—and supply—the LLM provides different advice to the same question depending on the stated gender. We find that two-thirds of the gender difference in equity allocation recommendations is driven by demand, while the remaining one-third is driven by supply. These findings echo empirical evidence from human financial advisors, where differences in advice by gender also reflect both demand and supply (Bhattacharya et al. 2024; Bucher-Koenen et al. 2025).

Related literature. An extensive literature in behavioral economics and household finance has documented numerous ways in which individuals make suboptimal financial decisions.[1] While this introduces a role for professional financial advisors, the quality of such advice is often limited due to advisors' high cost, low-quality advice (Linnainmaa et al. 2021; Andries et al. 2025), conflicts of interest (Mullainathan et al. 2012), and misconduct (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

[1]This includes evidence that individuals do not take full advantage of available 401(k) employer matches (Choi et al. 2011; Choukhmane et al. 2025), 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).

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 advice. This paper's main contribution is to introduce and implement a novel framework for quantitatively studying this advice. In doing so, this paper is part of a growing literature that studies 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.[2] 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.[3] 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, the method that we develop to study LLM's personal financial advice makes two contributions. First, we use human-written prompts to ensure that we accurately characterize both the supply *and* demand for LLM financial advice. This approach of "using humans to learn about AI" allows us to decompose the extent to which LLM advice is driven by supply versus demand. Second, we use a life cycle model to simulate LLM advice in a dynamic, stochastic environment. This allows us to characterize the effects of LLM advice over an entire lifetime in a way that can be compared with actual behavior as well as quantitative economic models.

Our study of LLMs as an accessible and cost-effective source of financial advice builds on prior work studying other forms of financial innovation and advice. This literature has shown that automated 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). Relative to other automated forms of financial advice, LLMs have the potential to provide the soft skills and emotional support that advisors can provide and clients value (Greig et al. 2025). More broadly, our framework extends a longstanding

[2]Also related is a part of this literature that studies the properties of agent-based economic models where the agents are LLMs. Two notable examples are Hao and Xie (2025), who show that previous-generation LLMs' solutions to a two-period consumption-saving problem can be quite far from the model's solution even when including the objective function explicitly in the prompt, and Douglas and Verstyuk (2024), who study the behavior of a Krusell and Smith (1998) economy populated with LLM agents.

[3]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).

tradition of using quantitative life cycle models 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.

1 Methodology

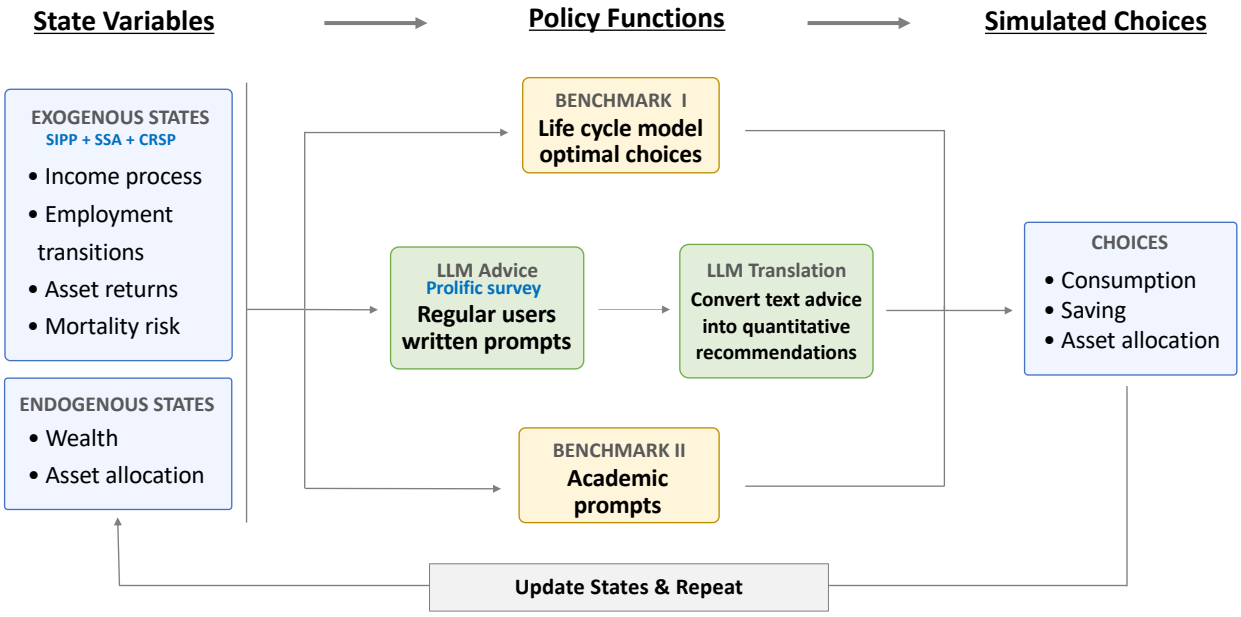
Our methodology, summarized in Figure 1, proceeds in four steps. First, we calibrate a quantitative life cycle model to U.S. income dynamics, employment transitions, asset returns, and tax rules; this model provides both the economic environment for our simulations and a normative benchmark grounded in economic theory (Section 1.1). Second, we conduct a survey to elicit prompts in which real individuals ask an LLM for financial advice (Section 1.2). Third, we simulate the life cycle paths of individuals who follow the LLM's advice by iteratively querying the LLM with these human-written prompts and extracting quantitative consumption, saving, and portfolio recommendations from its qualitative responses (Section 1.3). Fourth, we simulate life cycle paths generated by the LLM in response to structured prompts that provide richer information about individuals and explicit modeling assumptions. These prompts, which we call "academic prompts", are designed to elicit advice more closely aligned with the prescriptions of household finance theory (Section 1.5). The remainder of this section describes each step in turn.

1.1 Life Cycle Model

We use a quantitative life cycle model of consumption-saving and portfolio choices that builds on Choukhmane and de Silva (2026). The full specification and calibration are in Appendix A.

Each model period corresponds to one year. Investors enter working life at age $a_0 = 22$, retire deterministically at age 65, and face age-dependent mortality risk until a maximum age of 89. They have time-separable CRRA utility over consumption with annualized discount factor β and relative risk aversion γ. The labor market features four employment states: employed, job-to-job transition, unemployed, and retired. While employed, investors earn stochastic income with a deterministic age profile, a persistent AR(1) component, and a transitory shock. Job-to-job transitions carry a wage premium, while unemployment spells entail a persistent wage reduction upon reemployment. The government provides unemployment benefits, Social Security retirement benefits tied to average lifetime earnings,

Figure 1. Overview of Methodology for Simulating LLM Advice



Notes: This figure illustrates the three-step methodology. State variables (left) feed into policy functions (center), which produce simulated choices (right). The life cycle model (Benchmark I) provides optimal policy functions given calibrated state variables. LLM advice uses survey-written prompts to generate textual recommendations, which are then translated into quantitative choices. Academic prompts (Benchmark II) replace survey prompts with researcher-designed prompts. Endogenous state variables are updated each period and the process repeats over the life cycle.

and levies income and capital gains taxes according to the 2025 U.S. federal schedule.

Investors save in a taxable account that can be allocated across four assets: a risk-free bond, a diversified stock market index, individual stocks, and alternative risky investment vehicles like commodities and cryptocurrencies. These last two assets, which we refer to as the non-diversified assets, are designed to capture alternative holdings that the LLM sometimes recommends. Both are calibrated to have the same return process, but independent shock realizations, so they represent distinct risky positions. We construct the return process for these assets so that neither can improve the Sharpe ratio of a portfolio that combines the bond and the diversified stock market index; for more information, see Appendix A.6. In each period, the choice variables in this model are consumption, savings, and the portfolio shares in the four assets. Investors cannot borrow or take leveraged positions. The model has 7 state variables: age, labor productivity, transitory income shock, employment status, tenure, average lifetime income, and liquid savings. We calibrate the model closely following Choukhmane and de Silva (2026), using SIPP data for the income process and employment transitions, Social Security Actuarial Life Tables for mortality risk, and historical CRSP returns for asset return parameters.

1.2 Survey

The second step in our methodology is to conduct a survey to elicit prompts in which individuals ask an LLM for financial advice. We design a survey in Qualtrics and administer it through the Prolific online platform to elicit prompts from a nationally representative sample. Our initial sample consists of 999 complete responses, of which 952 pass Prolific AI authenticity check and a minimum specificity requirement. Appendix B.2 provides more details on this sample selection.

The survey asks each respondent to write three free-text prompts to an LLM financial advisor (see Figure A2 for the exact wording).[4] The first prompt asks respondents to describe their financial situation, including whatever details they think would help the LLM provide useful advice. Respondents are told to assume the LLM has no prior information about them, which encourages them to provide context in a way that mirrors how they would actually interact with a chatbot. The second prompt asks respondents to request advice on how much to spend over the coming year, while the third asks for advice on how to invest their savings between stocks and safer options. For the second and third prompts, respondents are told that the LLM already has the information from their first response.

In addition to the prompts, we survey respondents about the state variables in our life cycle model (including age, income, employment status, and financial wealth), as well as demographic characteristics such as gender, race, financial literacy, and prior experience with AI. Table A1 compares the demographics of our Prolific sample with those of the Current Population Survey (CPS). The sample broadly matches the CPS across age, gender, and income, though it over-represents the unemployed and under-represents the lowest-income earners and respondents older than 70.

Summary statistics for the prompts written by our survey respondents—including word counts, character counts, and the prevalence of numerical and dollar-amount mentions—are reported in Table A2–Table A4. The financial situation prompt is the longest of the three (mean 43 words), with 65% of respondents mentioning specific numbers and 34% mentioning dollar amounts. The spending and investment advice prompts are shorter (mean 27 words each), with lower rates of quantitative detail.

Around half of our respondents (48%) report having used an AI tool for financial advice or information in the past three months, with an additional 38% having considered doing so (Figure A3). This is consistent with recent survey evidence indicating that a large and

[4]One limitation of our methodology is that our survey elicits one-shot prompts rather than interactive conversations. To address this, we ask respondents to write three detailed prompts, each focusing on a distinct aspect of their finances, including both spending and investment decisions.

growing share of consumers are turning to LLMs for personal financial guidance (Lloyds Banking Group 2025; J.D. Power 2025). Among those with prior AI experience, the most common topics are saving (49%) and investing (49%), followed by budgeting (42%) and financial education (35%) (Figure A4). Our survey elicits prompts about spending and investing advice because these are the two most popular topics that individuals discuss with AI financial advisors and because they correspond to the two canonical choice variables in life cycle models—consumption-saving and portfolio allocation—allowing us to compare LLM advice directly against model-based benchmarks.

1.3 Simulating LLM Advice over the Life Cycle

The final step in our methodology is to simulate the life cycle paths of individuals who follow the advice provided by an LLM, as illustrated in Figure 1. Each simulated life path proceeds by iterating through five steps—states, prompt, advice, choices, and shocks—every year from age 22 to 89. Table 1 illustrates these steps with a concrete example for one period.

Step 1: States. At the beginning of each period, an individual is characterized by the state variables of our life cycle model: age, employment status, job tenure, income, and the value of assets held in each of the four asset classes (Table 1, Panel A). For example, a 22-year-old employed individual earning $29,480 with no financial assets.

Step 2: Prompt selection and preparation. Given an individual's state variables, we randomly draw a survey prompt written by an individual with similar characteristics. Prompts are organized into 12 buckets defined by employment status, age, and income, so that the advice given to a simulated individual is drawn from the set of survey respondents of broadly similar characteristics along these three dimensions (Figure B1).[5] This ensures that the tone, concerns, and qualitative descriptions in the prompt are appropriate for the individual's circumstances .

Within a bucket, the simulated individual's specific demographic and financial characteristics may still differ from those of the randomly drawn survey respondent. We therefore identify any references to state variables in the prompt and replace them with the corresponding values from the simulation. For example, Table 1, Panel B shows the raw prompt as originally written by a 23-year-old survey respondent earning approximately $30,000 per year. Panel C shows the same prompt after variable insertion: every mention of a state variable—including age, income, wealth levels, and portfolio allocations—is replaced with

[5] In principle, we could use other state variables as well, but are constrained by the size of our sample.

the simulated individual's current values. This preserves the author's writing style, sentence structure, and qualitative concerns while ensuring that the quantitative details match the simulated individual's situation. The complete set of insertion rules—covering income frequencies, household adjustments, wealth across multiple account types, and lump sum payments—is in Appendix B.

The three survey prompts (financial situation, spending advice, and investment advice) are then concatenated into a single message and sent to the LLM with an instruction to respond in 200 words or fewer.

Step 3: LLM advice. The LLM returns textual advice in response to the prepared prompt (Table 1, Panel D). The advice typically covers multiple topics: building an emergency fund, setting spending targets as shares of income, and recommending specific investment vehicles such as index funds or target-date funds. The qualitative properties of this advice—including which topics the LLM emphasizes and how they compare with what individuals discuss in their prompts—are described in Section 2.

Step 4: Translating advice into quantitative choices. The LLM's textual advice must be converted into the quantitative recommendations for the choice variables of our model: consumption, saving, and the allocation of savings across the four asset classes. We accomplish this by passing the LLM's advice back into an LLM along with deterministic instructions on how to extract numerical recommendations and output them in a structured JSON format. This second LLM call maps qualitative advice (e.g., "save 15–30% of income" and "invest in a total-market index fund") into dollar amounts for consumption and contributions to each asset class (Table 1, Panel E). The extraction follows a priority ordering: annual dollar amounts are used when stated; monthly amounts are annualized; ranges are converted to midpoints; percentages of income are converted to levels. The full translation prompt, including the asset class definitions, extraction priorities, budget constraints, and deterministic rules that ensure all outputs satisfy accounting identities, is in Appendix B.7.

Step 5: Shocks and state transitions. After extracting the quantitative recommendations, we draw realizations of the exogenous shocks in our life cycle model: labor market transitions, persistent and transitory income shocks, and asset returns for each of the four asset classes. The realized shocks, combined with the simulated choices, determine the state variables at the beginning of the next period (Table 1, Panel F). These updated states are then used as the input for the next iteration.

Simulation design. Each query to the LLM is independent, with no memory of past responses; the only link across periods is the evolution of state variables that depend on prior choices. We simulate 1,000 individuals, each starting life at age 22 with no wealth, retiring at age 65, and deterministically dying at age 90. A new prompt is drawn and prepared in every period for every individual.

1.4 Model Selection

We choose an OpenAI model as our primary LLM because survey evidence—both from our own sample (Figure A1) and from other recent surveys (Lloyds Banking Group 2025; J.D. Power 2025)—indicates that ChatGPT is the most commonly used AI tool for financial advice. Among OpenAI's model family, we use `GPT-5.2` (OpenAI 2025), which was the latest available model at the time we ran our analysis. We set its reasoning effort to the low level primarily for reasons of speed and cost efficiency. For the JSON translation task described in Step 4 of Section 1.3, we use `GPT-5 Mini`, a smaller and faster model in the same family, since this task involves deterministic extraction rules rather than open-ended reasoning. As a robustness check, we repeat our main analysis using `Gemini 3 Flash` and report the results throughout the paper. The translation task continues to use `GPT-5 Mini` in this robustness exercise.

1.5 Two Benchmarks

We construct two benchmarks, outlined in Figure 1, to compare against the advice generated by the survey prompts.

Benchmark I: Life cycle model. The first benchmark is constructed by solving the life cycle model described above and simulating optimal behavior under the same exogenous shock realizations used in the LLM simulation. As shown in Figure 1, steps 2 through 4 of the simulation pipeline are replaced by the model's optimal policy functions. Because the aggregate stock index is calibrated to be mean-variance efficient, agents in this benchmark hold neither non-diversified asset class, reducing portfolio choice to a standard allocation between the risk-free bond and the diversified stock index.

Benchmark II: Academic prompts. The second benchmark replaces the human-written survey prompts with an academic prompt that we design to provide the LLM with explicit information about the individual's current states and assumptions about the economic environment. Like the first benchmark, it replaces steps 2 through 4 of the simulation pipeline. The academic prompt provides the LLM with individual-specific state variables—age, income, account balances, employment status, job tenure, and average past income—

Table 1. From States to Choices: Simulation Example for One Period

Panel A: State variables

Age	Employment status	Annual income	Bonds	Diversified stocks	Individual stocks	Alternative risky assets
22	Employed	$29,480	$0	$0	$0	$0

Panel B: Raw survey prompt (as written by respondent)

Financial situation	"I'm 23 years old making 30,000 a year, I dont spend a lot besides normal expenses such as food and gas."
Spending advice	"at 30,000 a year what should my spending look like?"
Investment advice	"i'm looking to start investing to get my money up, preferably slow gain. How can i get into this in a smart and somewhat safe way while making 30,000 a year?"

Panel C: Prepared prompt (after variable insertion)

Financial situation	"I'm **22** years old making **$29,480** a year, I dont spend a lot besides normal expenses such as food and gas."
Spending advice	"at **$29,480** a year what should my spending look like?"
Investment advice	"i'm looking to start investing to get my money up, preferably slow gain. How can i get into this in a smart and somewhat safe way while making **$29,480** a year?"

Panel D: LLM advice

"On $29,480/yr (~$2,457/mo gross), aim for a simple, safe setup: 1) Build a base first — Emergency fund: start with $1,000, then build to 3–6 months of essential expenses in a high-yield savings account (HYSA). — Pay off high-interest debt (esp. credit cards) before investing. 2) Spending targets (monthly take-home varies by state/benefits) — Needs (rent, utilities, food, gas, insurance): 55–70% — Wants: 5–15% — Saving/investing: 15–30% (even $50–$200/mo is fine). If rent is high, keep investing smaller but consistent. 3) 'Slow gain' investing (low-cost, diversified) — If you have a 401(k) match, contribute enough to get the full match. — Otherwise (or in addition), open a Roth IRA (Fidelity/Vanguard/Schwab). — Invest in a total-market index fund/ETF or a Target-Date fund (set-and-forget). Keep fees low."

Panel E: Extracted choices

Consumption	Total saving	Bonds contribution	Diversified stocks contribution	Individual stocks contribution	Alternative risky assets contribution
$22,847	$6,633	+$1,000	+$5,633	$0	$0

Panel F: Updated states (beginning of next period)

Age	Employment status	Annual income	Bonds	Diversified stocks	Individual stocks	Alternative risky assets
23	Employed (no transition)	$36,528 (income shocks)	$1,020	$6,247	$0	$0

Notes: This table illustrates the five steps of the simulation for a single individual in one period. Panel A shows the state variables at the start of the period. Panel B shows the raw survey prompt as originally written by a 23-year-old respondent earning approximately $30,000 per year. Panel C shows the same prompt after variable insertion, where boldface values indicate state variables that were substituted into the respondent's text. Panel D shows the full LLM advice. Panel E shows the quantitative choices extracted by the translation LLM call, which must satisfy the budget constraint: consumption + total saving = post-tax income. Panel F shows the state variables at the start of the next period after labor market, income, and asset return shocks are realized.

along with baseline assumptions about expected asset returns, tax and Social Security rules, and emphasizes the LLM's duty as a professionally-trained financial advisor. The full prompt is presented in Appendix B.6. Table B3 shows how the baseline assumptions of the academic prompt map into the calibration of our life cycle model, and how the individual-specific employment and income statements inserted each period are determined by the simulated individual's state variables.

2 Textual Analysis of Prompts and Advice

We conduct a dictionary-based textual analysis of both the free-text prompts written by respondents and the AI-generated advice they receive.

Methodology. Each response is tokenized, lemmatized, and matched against a dictionary of 27 topic categories (e.g., Debt, Savings, Retirement, Investment Assets, Risk Preference). For each respondent, prompt, and category, we record whether any keyword appears, and report the share of respondents mentioning each category. To compare prompts with advice, we weight by the number of times each respondent's prompt appears in the simulation. Full methodological details—including the text preprocessing pipeline (Figure C1), the complete dictionary (Table C1), and prompt-level summary statistics (Table A2–Table A4)—are in Appendix C.

What prompts discuss. Figure 2 provides a visual overview of the language respondents use across all three prompts and the language that appears in AI-generated advice. The most prominent words in the combined prompt word cloud are "money," "save," "invest," "retire," and "budget," reflecting respondents' focus on core personal finance decisions. In the advice word cloud, the most prominent words—"fund," "invest," "save," "stock," and "bond"—reflect the investment-oriented nature of the AI responses, with terms like "essential," "annual," "target," "treasury," and "diversified" highlighting their structured, comprehensive character.

The topics respondents discuss vary across the three prompts they write. When describing their financial situation, Income (35%) and Retirement (30%) are the most common categories. When asking for spending advice, Budgeting is the most frequent (45%), while when asking for investment advice, Investment Assets (53%) and Risk Preference (40%) are most common. The complete set of all 27 categories is shown in Figure A6.

What LLMs recommend. We next compare the topics raised by survey respondents to those present in the AI advice. Figure 3 shows the most prominent categories across both prompts and LLM advice. The top five prompt categories—Budgeting, Investment Assets,

Figure 2. Word Clouds: Respondent Prompts and AI-Generated Advice



Notes: This figure shows word clouds for the pooled text across all three respondent prompts (left) and for AI-generated financial advice (right). Word size is proportional to raw frequency. Stopwords—including scikit-learn's default English stopwords, custom generic terms, and survey-instrument words—and the union of prompt-specific tautological terms across all three prompts (e.g., "spend," "invest," "describe") are excluded from the prompt panel; base stopwords are excluded from the advice panel. Stopwords are removed only for word-cloud display, not for dictionary matching. Text is preprocessed using spaCy tokenization, lemmatization, and domain-specific root merging. Figure A5 shows separate word clouds for each of the three individual prompts.

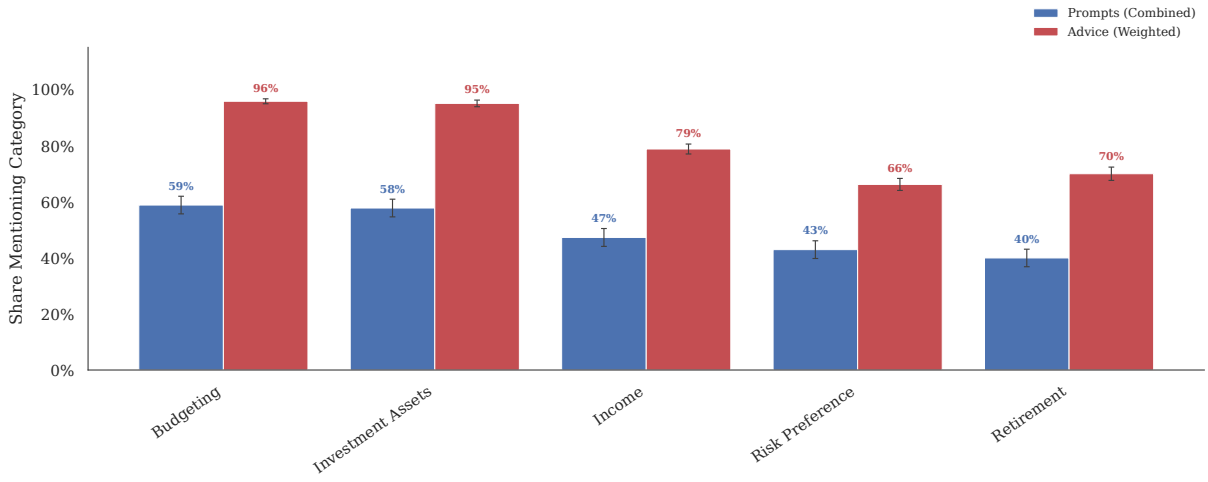
Income, Risk Preference, and Retirement—all appear at substantially higher rates in the advice, confirming that AI responses address the main financial topics respondents raise in their prompts.

More striking is what the AI introduces unprompted. Liquidity appears in 83% of advice responses despite only 6% of respondents mentioning it explicitly, reflecting the models' high propensity to recommend building an emergency fund. Savings shows a similar pattern, rising from 20% prevalence in survey prompts to 76% prevalence in advice. Despite these differences, topic rankings remain broadly correlated: the Spearman rank correlation between prompt and advice prevalence across all 27 categories is $\rho = 0.71$ ($p < 0.001$), indicating that AI advice amplifies respondents' concerns rather than ignoring them. The full comparison across all 27 categories is in Figure A8.

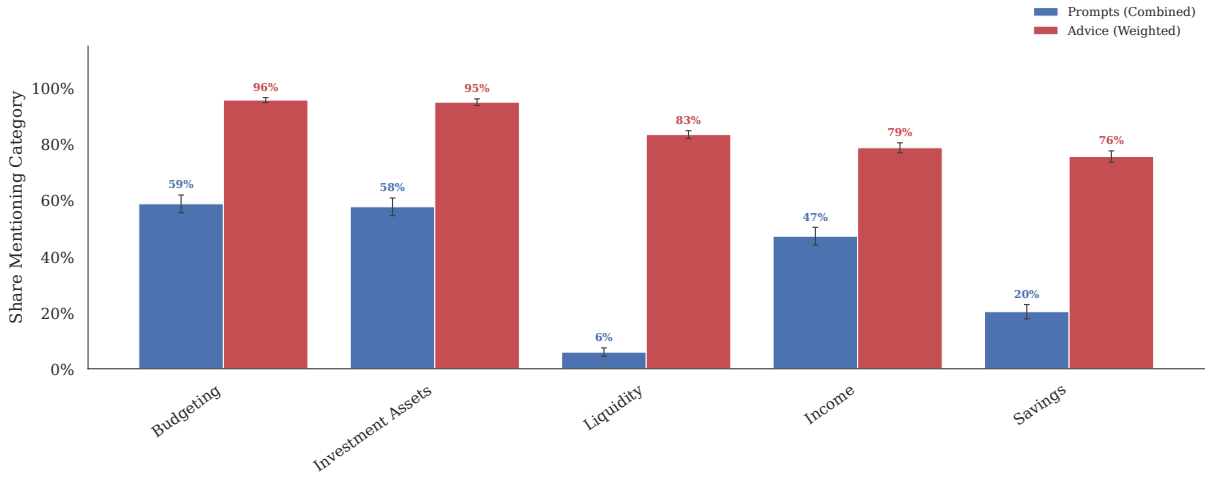
Figure 4 shows that a similar pattern emerges for specific asset types and financial products. Stocks (Aggregate) are the most common asset class in both prompts (49%) and advice (85%), but LLM advice also systematically mentions conservative, diversified alternatives that respondents rarely mention. For example, High-Yield Savings accounts appear in 59% of advice responses (vs. 3.6% of prompts), Bonds in 54% (vs. 10%), and Treasury Bills/Notes/Bonds in 42% (vs. 1%). At the product level, fewer than 3% of respondents name any specific ticker, yet AI advice routinely recommends branded products: Vanguard and its investment products appears in 6% of responses, iShares in 2.9%, and cryptocurrency tokens such as Bitcoin or Ethereum in 2.0%. Together, these patterns show

Figure 3. Top 5 Dictionary Categories: Prompts vs. Advice

Panel A: Ranked by Prompt Prevalence



Panel B: Ranked by Advice Prevalence



Notes: Top 5 dictionary categories ranked by combined-prompt prevalence (Panel A) and by advice prevalence (Panel B), for $n = 944$ matched respondent-advice pairs. Blue bars show the share of respondents mentioning the category in any of the three prompts. Red bars show the respondent-weighted mean mention rate in AI-generated advice. Error bars show 95% confidence intervals.

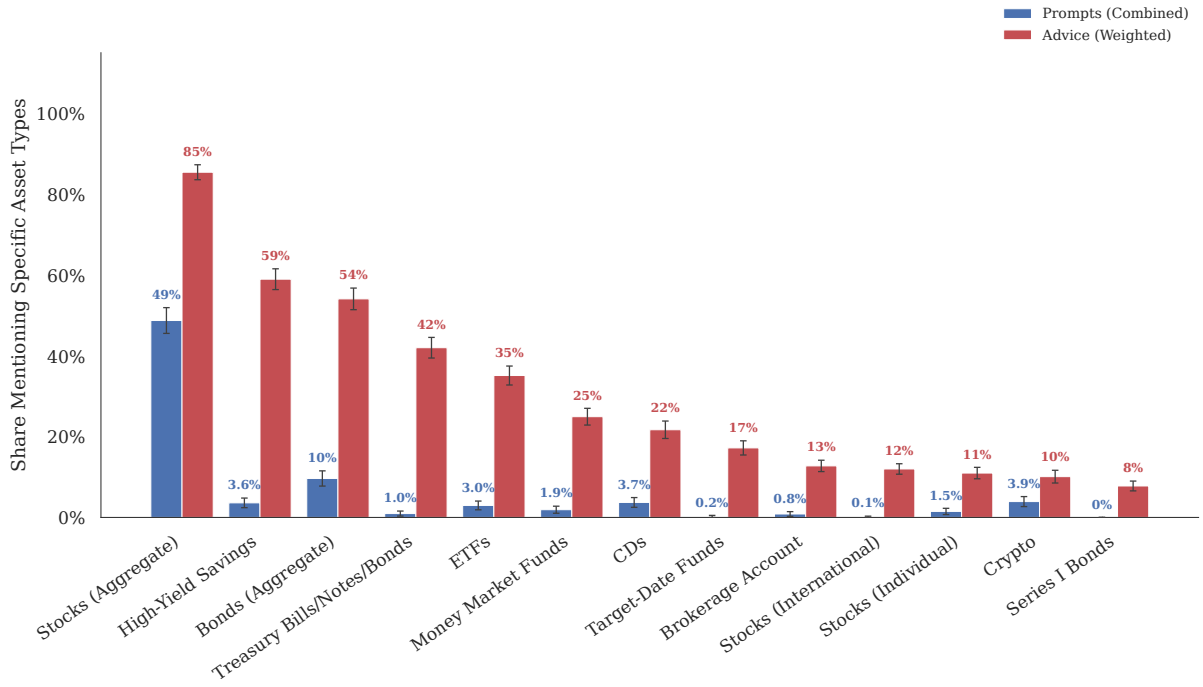
that AI advice is both broader and more specific than the prompts it responds to: it covers more topics and names specific products, even when respondents ask about neither.

3 LLM Advice Over the Life Cycle

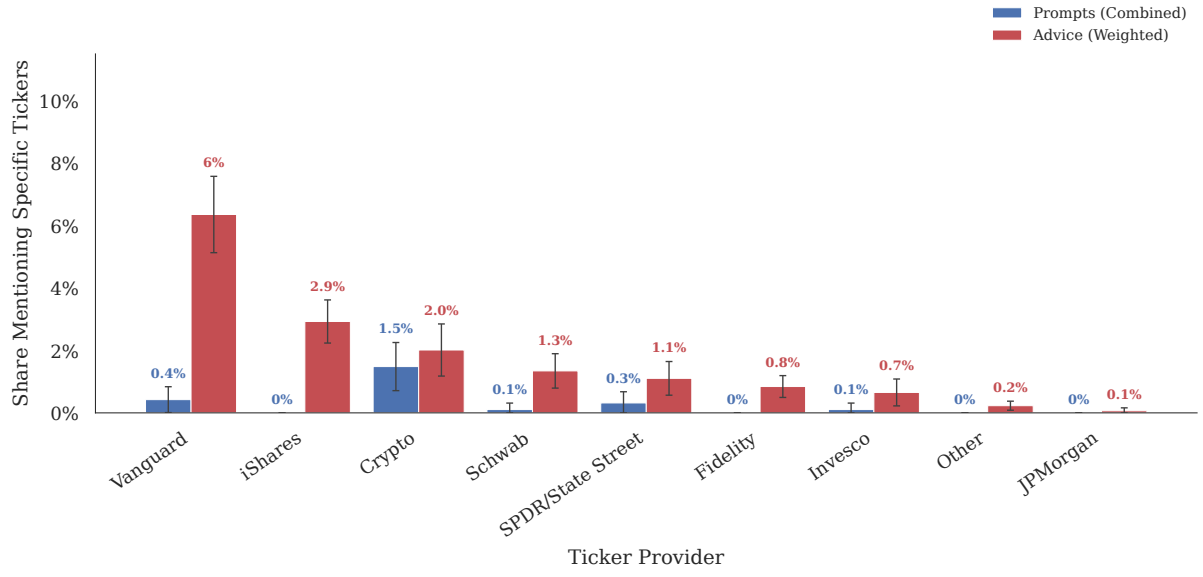
In this section, we describe how LLM recommendations vary over the life cycle, and compare them to survey respondents' actual financial behavior and to the prescriptions of a standard life cycle model. Our baseline results are generated using `GPT-5.2` (with robustness in the appendix using `Gemini 3 Flash`).

Figure 4. Investment Assets and Products: Prompts vs. Advice

Panel A: Investment Asset Types



Panel B: Specific Products by Provider



Notes: Grouped bar charts comparing prompt and advice prevalence (n = 944 matched pairs). Panel A shows specific investment asset types with at least 5% prevalence in either prompts or advice, ordered by descending advice prevalence. Panel B shows specific financial products (tickers and fund names) aggregated by issuing provider, ordered by descending advice prevalence. Blue bars show the share of respondents mentioning the category in any of the three prompts. Red bars show the respondent-weighted mean mention rate in AI-generated advice. Error bars show 95% confidence intervals.

3.1 Consumption, Saving, and Portfolio Choice

Figure 5 plots the median consumption, wealth, and stock share by age for individuals following the LLM's advice, along with 25thâĂŞ75th percentile bands. Table A5 presents additional summary statistics.

Consumption and saving. The left panel shows that, consistent with textbook life cycle theory, the LLM recommends that individuals smooth their income over working life such that consumption is flatter than income. Individuals following the LLM's advice achieve this by saving more during working life and consuming less than their income, while withdrawing and consuming (slightly) more than their income during retirement. This is reflected in their wealth accumulation: Table A5 shows that 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.

Portfolio allocation. The right panel of Figure 5 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 similar to the average equity shares that retirement savers choose when making active portfolio choices (Choukhmane and de Silva 2026). The LLM also recommends small allocations to non-diversified assets—individual stocks and other risky assets such as crypto, gold, and commodities. As shown in Figure A10, participation in these asset classes increases with age, reaching approximately 40% for individual stocks and 75% for other risky assets by retirement. However, conditional portfolio shares remain small, averaging 2–3% of financial wealth, so that the vast majority of the LLM's recommended risky asset allocation is in diversified equities.

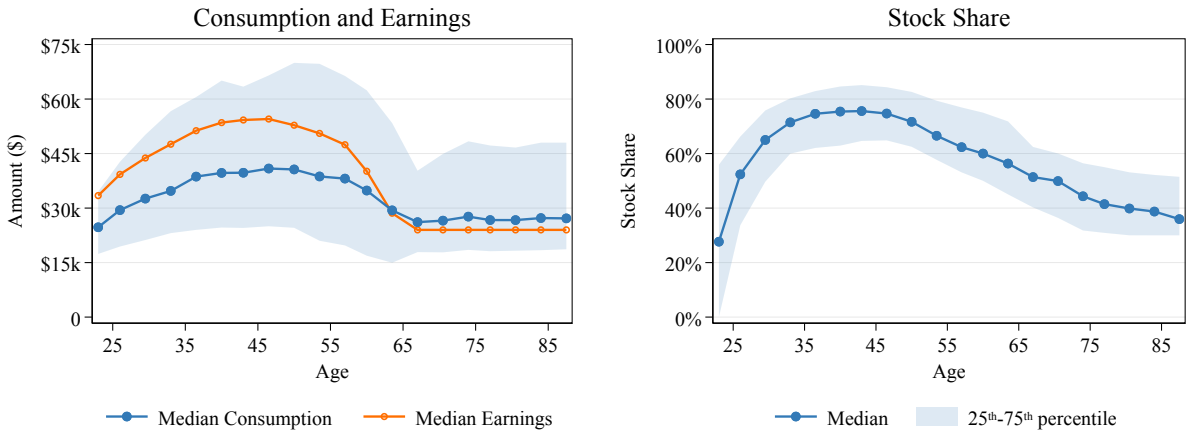
3.2 LLM Advice Would Move Individuals Toward Life Cycle Theory

Relative to their current behavior, following LLM recommendations would move most of our survey respondents toward life cycle theory, with broader stock market participation, age-declining equity shares, and sizeable saving buffers. Figure 6 plots respondents' current financial situation (in red) alongside the counterfactual implied by following LLM recommendations throughout the life cycle (in blue).[6] Three differences stand out.

First, consistent with the empirical evidence on limited stock market participation, one-third of respondents to our survey report not holding any equities, and their average equity

[6]We find quantitatively similar results when using `Gemini 3 Flash`, which are shown in Figure A11.

Figure 5. Life Cycle Profiles of LLM-Recommended Consumption, Wealth, and Stock Shares



Notes: This figure plots the life cycle profiles for individuals following the LLM's recommended choices using the survey prompts, as described in Section 1.5. Ages (22-89) are grouped into 20 equally sized bins. The left panel shows consumption and post-tax salary; 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.

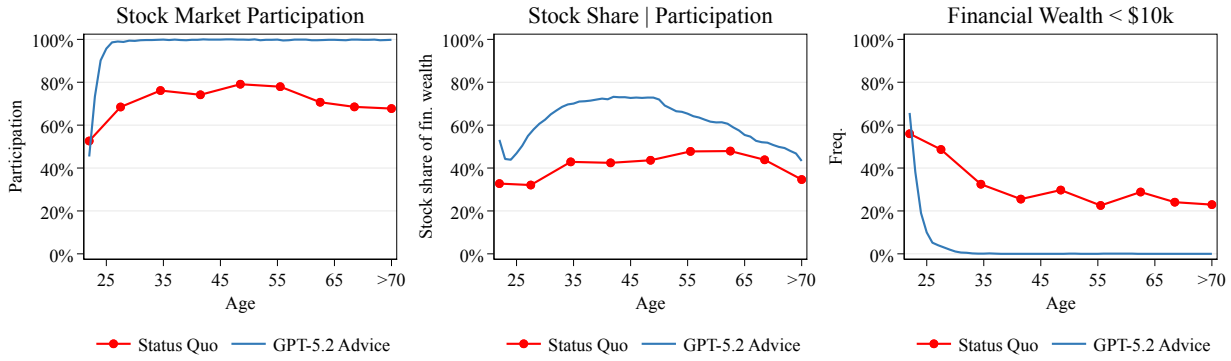
share conditional on participation is around 40%. In contrast to their current behavior, following the LLM recommendation would lead to near-universal participation with equity shares 20 to 30 percentage points higher. The one exception is young adults, where the LLM recommends lower initial participation, reflecting its tendency to prioritize building an emergency fund with safe assets before investing in equities.

Second, respondents' equity shares are roughly flat (if not increasing) by age, whereas the LLM recommends a declining age profile past age 45, more consistent with life cycle theory. As we discuss in the next section, however, this decline is quantitatively more modest than our life cycle model predicts.

Third, more than 20% of survey respondents hold less than $10,000 in financial wealth across all age groups. Under LLM recommendations, virtually all individuals build a substantial savings buffer by age 30.

Together, these patterns indicate that following LLM advice would move most respondents meaningfully closer to the broad prescriptions of life cycle theory. An important caveat is that this comparison assumes individuals would follow the LLM advice and act accordingly. Whether and how LLM advice affects actual financial decisions, and how it compares to other forms of financial advice, are important questions for future work. We next examine how the LLM's recommendations compare to our life cycle model quantitatively.

Figure 6. Current Behavior vs. LLM-Recommended Behavior



Notes: This figure compares survey respondents' current financial behavior with the LLM's recommendations. Current behavior is reported by individuals in our Prolific survey. LLM recommendations are generated using survey prompts as described in Section 1.5. Ages (22-89) are grouped into 20 equally sized bins. All values are in 2025 dollars.

3.3 Summarizing LLM Advice with Structural Preference Parameters

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).[7] 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 D for additional details).

The first row of Table 2 shows the estimated values when using survey prompts, which are $\hat{\beta} = 1.044$ and $\hat{\gamma} = 5.2$. 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 very high, above one. This reflects the fact that, as shown in the left panel of Figure 5, the LLM recommends consuming less than income throughout working life, resulting in a significant amount of wealth accumulation. While bequests could in principle explain this wealth accumulation, almost no individuals in our survey explicitly mention bequests in their prompts. An alternative interpretation of the high estimated $\hat{\beta}$ is that it reflects a form of paternalism or sophisticated self-control on the part of the LLM: by recommending higher savings rates than individuals would choose under standard preferences, the LLM may be acting as a commitment device that helps individuals avoid undersaving, even at the cost of implying an unrealistically patient

[7]This approach is similar to Cook et al. (2025), who estimate the structural preference parameters from LLM's behavior in dictator games and job search scenarios.

Table 2. Simulated Method of Moments Estimation Results

Prompt Type	LLM Model	$\hat{\beta}$	$\hat{\gamma}$
Survey	`GPT-5.2`	1.044	5.2
Survey	`Gemini 3 Flash`	1.054	5.1
Academic	`GPT-5.2`	0.99	4.7

Notes: This table presents the results of our simulated method of moments estimation. The two columns on the right present the estimated preference parameters $\hat{\beta}$ (intertemporal discount factor) and $\hat{\gamma}$ (coefficient of relative risk aversion) that minimize the sum of squared proportional errors between a set of LLM advice moments and their equivalent moments calculated in the life cycle model. The moments we use are the 25th, 50th, and 75th percentiles of wealth-to-income ratio in each year of working life and the 25th, 50th, and 75th percentile of equity share for each year in the full life. For a full description of the estimation procedure, see Appendix D. Survey prompts use the individual-written prompts from our Prolific survey. Academic prompts use the researcher-designed prompt described in Section 1.5.

discount factor.

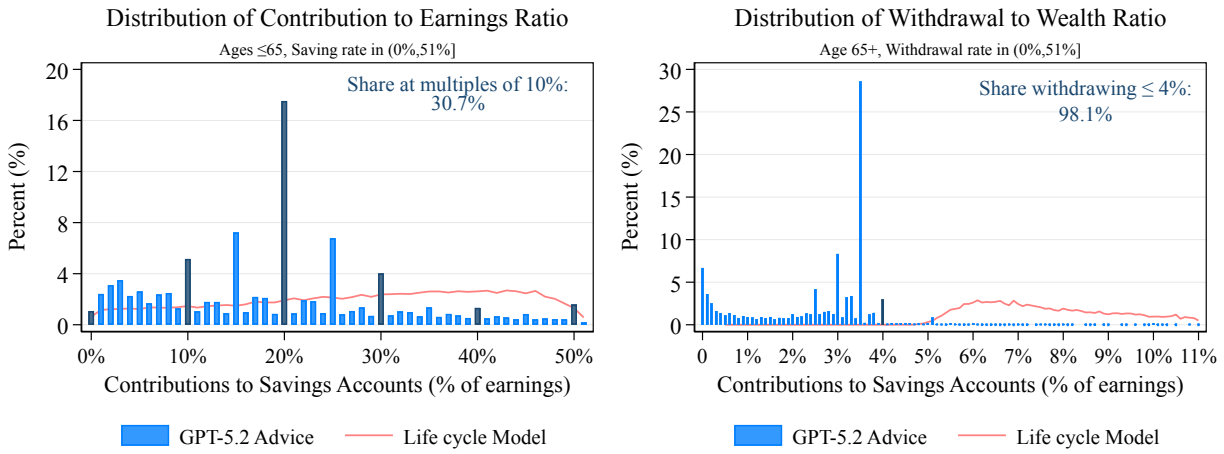
Figure A14 shows the fit of the estimated life cycle model on the targeted moments. With the estimated $\hat{\beta}$ above one, the model does a good job of quantitatively replicating wealth accumulation during working life. Turning to equity shares, the model does a good job of replicating the LLM's advice after age 45. However, between age 25 and 45, the model cannot generate an equity share as low or as upward-sloping as the LLM recommends. This is to be expected: in standard portfolio choice models with uncorrelated labor income risk and no participation costs (Cocco et al. 2005), very high equity shares are optimal early in life. One way to rationalize the differences early in life is through a positive correlation between labor income risk and stock returns (Benzoni et al. 2007; Catherine 2022), which makes human capital less bond-like and reduces the optimal equity share. However, doing so would also require including participation costs, given that a large part of the increasing equity share is driven by the extensive margin, as shown in Figure 6.

The second row of Table 2 also shows estimates using `Gemini 3 Flash`. As illustrated in the estimates and the model fit in Figure A14, our results are quantitatively very similar. This suggests that our inferences are not specific to one LLM, but rather characteristic of a particular generation of models.

3.4 LLM Advice Relies on Simple Savings and Withdrawal Heuristics

Given the differences between the LLM's financial advice and our life cycle model, we now try to understand these differences by looking at the individual-level recommendations from the LLM. Figure 7 shows the distributions of savings rates, savings amounts, and withdrawal heuristics recommended by the LLM when using survey prompts. 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

Figure 7. Saving and Withdrawal Heuristics in LLM Advice: Survey Prompts



Notes: This figure plots the distributions of savings rates, savings amounts, and withdrawal heuristics in both the LLM advice and the life cycle model. The LLM advice is generated using the survey prompts as described in Section 1.5. Blue histograms represent the LLM advice, and dark blue bars highlight heuristic values, with spikes in savings rates at multiples of 10%, savings amounts at multiples of $5,000, and withdrawal-to-wealth ratios at 4%. The red lines represent the corresponding distributions from the life cycle model described in Section 1.1. All values are in 2025 dollars.

multiples of $5K. By construction, these types of heuristics are not present in the choices made by our life cycle model.

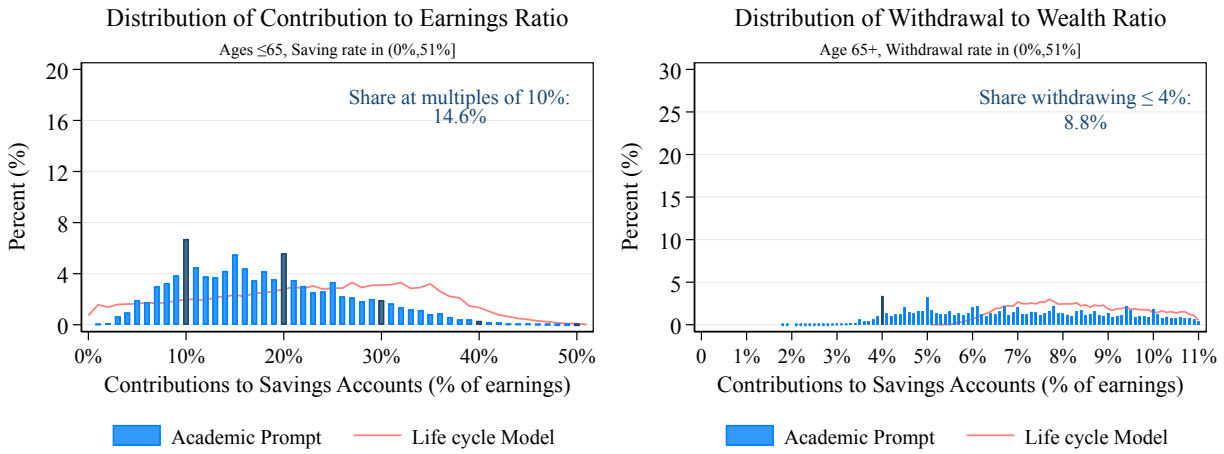
Turning to withdrawals, we find that over 85% of the LLM's recommended withdrawals respect the 4% withdrawal rule, which was popularized by Cooley et al. (1998) and is recommended by many financial advisors. In contrast, almost no individuals in our life cycle model choose withdrawal rates below 4%, consistent with the fact that the LLM's advice is overly conservative. Figure A15 shows that we also find similar heuristic patterns with `Gemini 3 Flash`.

3.5 Academic Prompts Move Advice Closer to Life Cycle Theory

We next examine whether replacing individual-written survey prompts with structured academic prompts can bring the LLM advice more in line with life cycle theory. As described in Section 1.5, the academic prompt provides the LLM with explicit information about all of individuals' current states, clarifies assumptions about the economic environment, and emphasizes the LLM's duty as a professionally-trained financial advisor.

The academic prompt improves consumption smoothing over the life cycle and following income shocks. The left panel of Figure 9 shows that consumption profiles no longer track income over the life cycle, with significant decumulation in retirement. As shown in Figure 8, this difference is driven by reduced reliance on simple heuristics: as shown in Figure 7, the

Figure 8. Saving and Withdrawal Heuristics in LLM Advice: Academic Prompts



Notes: This figure plots the distributions of savings rates, savings amounts, and withdrawal heuristics in both the LLM advice and the life cycle model. The LLM advice is generated using the academic prompt as described in Section 1.5. Blue histograms represent the LLM advice, and dark blue bars highlight heuristic values. The red lines represent the corresponding distributions from the life cycle model described in Section 1.1. All values are in 2025 dollars.

academic prompts produce fewer round-number saving rates and less adherence to strict safe withdrawal rules in retirement relative to the survey prompts (Figure 7). Quantitatively, these differences translate into more reasonable estimated intertemporal discount factor of $\hat{\beta} = 0.99$ (third row of Table 2). Portfolio choices, by contrast, are broadly similar (Figure A16) and yield a similar estimate of relative risk aversion ($\hat{\gamma} = 4.7$).

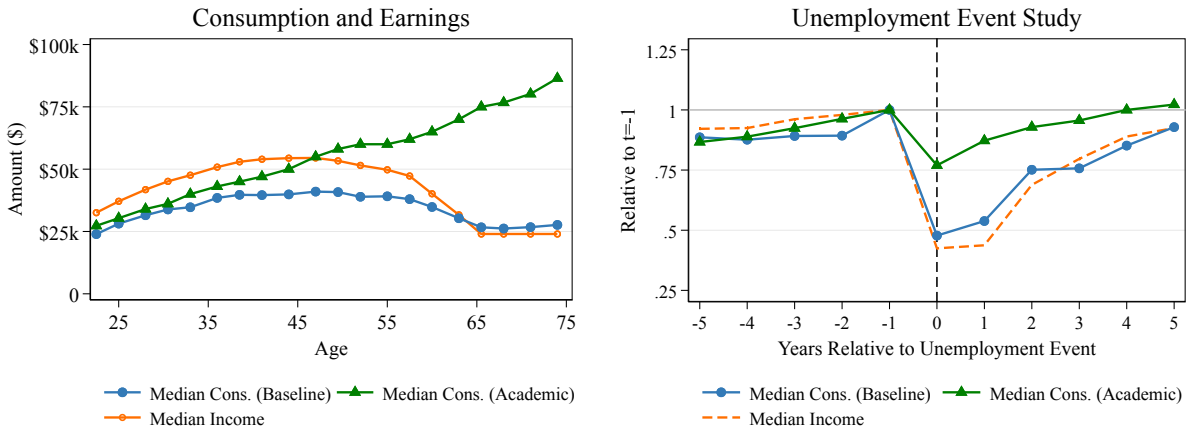
The academic prompt also produces better consumption smoothing in response to income shocks. The right panel of Figure 9 presents an event study of consumption around transitions from employment to unemployment, during which income falls by roughly 50%. Under the survey prompt, consumption drops by around 50% upon job loss, consistent with the empirical evidence in Gruber (1997) and Ganong and Noel (2019), indicating that the LLM fails to smooth consumption despite holding sizable liquid asset balances.[8] Under the academic prompt, the consumption drop is smaller and recovers more quickly, consistent with greater emphasis on maintaining emergency savings buffers rather than following round-number rules. In our life cycle model, consumption does not fall meaningfully as individuals optimally draw on their buffer stocks to absorb the shock (Figure A18).

3.6 LLM Advice Responds Passively to Return Shocks

The final aspect of the LLM's financial advice that we compare with our life cycle model is how it responds to stock return shocks. Following Calvet et al. (2009), we compute a

[8] `Gemini 3 Flash` exhibits similar patterns, as shown in Figure A17.

Figure 9. Consumption Smoothing: Survey vs. Academic Prompts

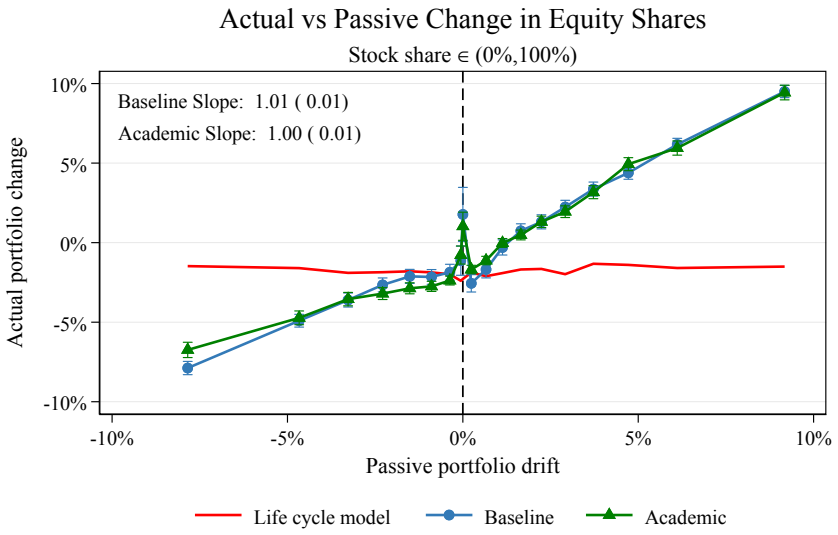


Notes: This figure compares consumption profiles between the survey and academic prompts. The left panel plots median consumption and median income over the life cycle for each prompt type. The right panel plots the median ratio of consumption relative to its level one year prior to unemployment within a five-year window around the unemployment event, along with median income. All values are in 2025 dollars.

measure of passive portfolio drift that captures the change in an individual's equity 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. Figure 10 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. In contrast, the LLM's financial advice exhibits a relationship between actual and passive portfolio changes that is close to one for both the survey and academic prompts (regression slopes of 1.01 and 1.00, respectively). 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 (2026). Notably, the academic prompt does not reduce this inertia: the regression slopes are nearly identical across the two prompt types, indicating that while academic prompts improve consumption smoothing and reduce reliance on savings heuristics, they do not affect portfolio rebalancing behavior. 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 regardless of prompt design. One likely explanation for this passive rebalancing behavior is that the LLM's recommendations typically pertain to how individuals allocate new savings rather than how they reallocate existing wealth, which naturally generates inertia in the overall portfolio.

Figure 10. Passive Portfolio Drift in LLM Advice



Notes: This figure compares actual portfolio changes against passive portfolio drift for the LLM advice (using both survey and academic prompts) and the life cycle model. 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. Error bars show 95% confidence intervals. The regression slopes for the survey and academic prompts are reported in the top-left corner. The sample is restricted to observations with stock shares strictly between 0% and 100%.

4 Heterogeneity in LLM Financial Advice

The results in the previous section characterize the average properties of LLM financial advice. However, LLM advice may vary across individuals, not just because of different economic circumstances—which was studied above—but also because of individual characteristics such as gender, financial literacy, and prior AI experience. In this section, we study these sources of heterogeneity. For each characteristic, we split the survey respondents into two groups based on their characteristic, simulate full life cycles using only prompts from each group, and compare outcomes. The realizations of exogenous shocks are identical across groups, so any differences in outcomes are driven entirely by differences in the prompts written by each group and the LLM's responses to them.

4.1 Decomposition of Wealth Differences across Demographics

Figure 11 summarizes the differences in life cycle outcomes across three demographic dimensions. Each panel reports four outcome measures: wealth at age 60 (in levels), log wealth at age 60, the log future value of net saving through retirement (computed using the risk-free rate), and the average risky asset share over the life cycle. The future value of net saving isolates the saving channel from portfolio allocation by compounding all net

contributions at the risk-free rate, so differences in this measure reflect only the amount saved, not the returns earned on those savings.

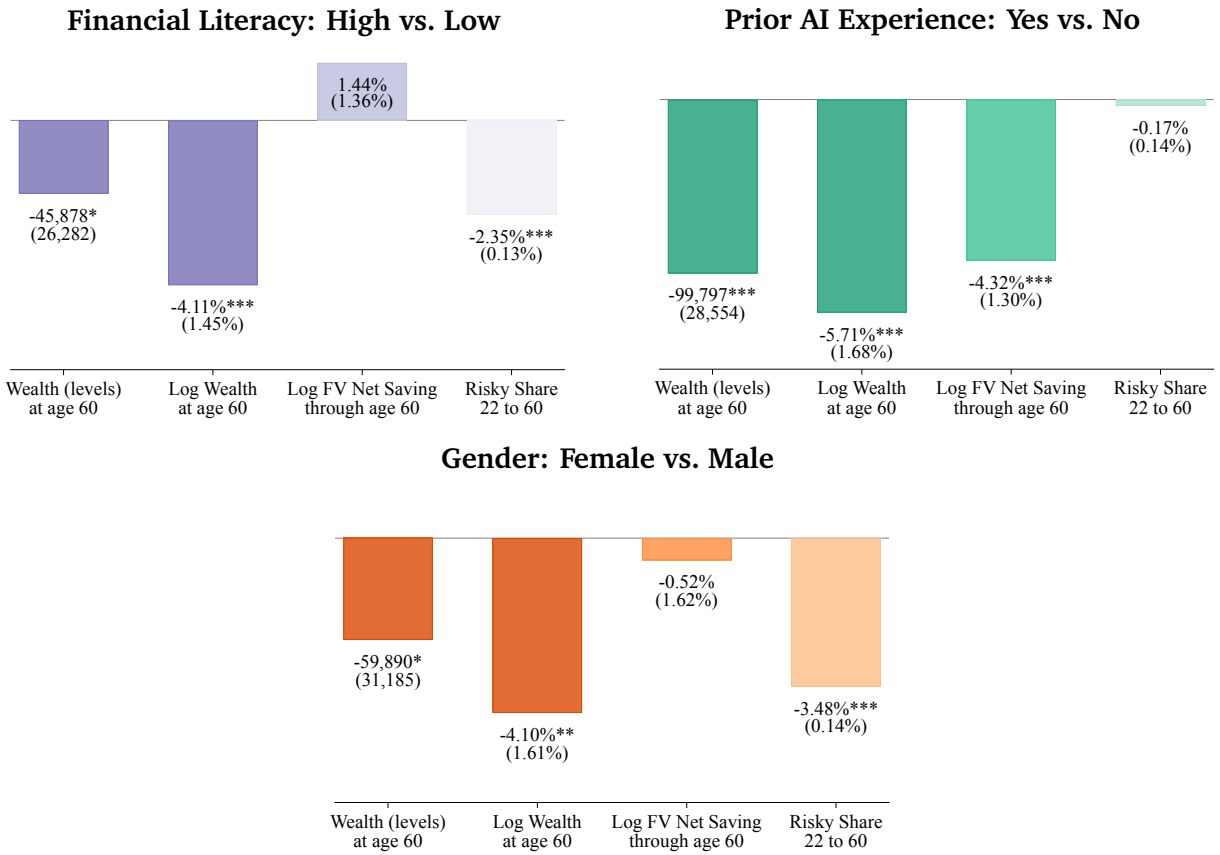
Financial literacy. The top-left panel of Figure 11 splits respondents into high- and low-literacy groups based on a Big 5 financial literacy test administered in our survey. Individuals with low financial literacy receive advice that leads to $45,878 (4.11%) lower wealth at age 60. The decomposition across the remaining columns reveals that this difference is driven primarily by equity shares: while the future value of net saving is actually *higher* for individuals with low financial literacy (1.44%), the average risky asset share recommended to low-literacy individuals is 2.35pp lower. This indicates that the wealth differences reflects lower equity allocations rather than lower cumulative savings. This difference in investment advice is reflected in the distribution of words mentioned in the advice: Figure A20 shows that the LLM's advice tends to discuss investment-related terms, such as equity and portfolio, more frequently in response to prompts written by high financial literacy respondents. The figure also shows that these differences are present in the prompts people write; see Figure A19 for an example of prompts that illustrate these differences.

Prior AI experience. The top-right panel of Figure 11 examines whether LLM advice varies with individuals' prior experience using AI for financial advice. We find that individuals who have not previously used AI for financial advice receive recommendations that lead to $99,797 (5.71%) lower wealth at age 60. In contrast to the financial literacy gap, this difference is driven almost entirely by lower cumulative savings: the future value of net saving is 4.32% lower in the non-AI users simulation, while the difference in average risky asset shares is not significant.[9] Figure A21 in the appendix shows example prompts from each group, and Figure A22 shows the corresponding differences in the LLM's textual advice.

Gender. The bottom panel of Figure 11 compares advice given by the LLM to men and women. Women accumulate $59,890 (4.10%) less wealth at age 60 than men. The difference is driven entirely by differences in recommended equity shares: the average risky asset share recommended to women is 3.48pp lower than for men, while the future value of net saving is nearly identical (−0.52%). Figure A23 shows that this difference in equity shares is present across all age and income groups, and that men are also recommended higher rates of active portfolio rebalancing. Notably, these gender differences in equity shares are consistent with those observed in observational data (Agnew et al. 2003) and in professional financial advice (Bucher-Koenen et al. 2025).

[9] In principle, AI use and financial literacy could be correlated. However, the difference in the underlying mechanisms driving the wealth difference suggests that these are, in fact, two distinct results.

Figure 11. Heterogeneity in LLM Financial Advice



Notes: This figure compares life cycle outcomes when simulating LLM advice using prompts written by different groups of individuals. All exogenous shocks are held fixed across groups. The top-left panel compares prompts written by individuals with high versus low financial literacy (as measured by a Big 5 financial literacy test), the top-right panel compares prompts written by individuals who have versus have not previously used AI for financial advice, and the bottom panel compares prompts written by men versus women. Within each panel, the four bars show (from left to right) the difference in wealth at age 60 (in levels), the difference in log wealth at age 60, the difference in the log future value of net saving through retirement (computed at the risk-free rate), and the difference in the average risky asset share over the life cycle. All values are in 2025 dollars. Standard errors in parentheses.

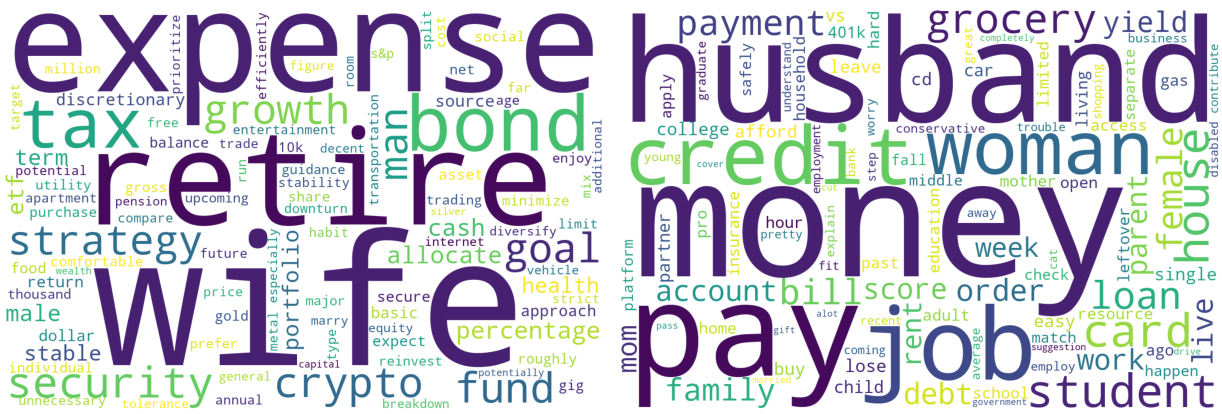
4.2 Decomposing the Gender Difference into Demand and Supply

The heterogeneity results above raise a natural question: to what extent do the differences in LLM advice reflect differences in demand (i.e., variation in the prompts written by different individuals) versus supply (i.e., differences in the LLM's advice for a given prompt)? We focus this decomposition on gender for three reasons. First, there is an extensive literature studying gender differences in traditional financial advice (Bucher-Koenen et al. 2025), making it a natural case to study. Second, separately identifying the roles of demand and supply is difficult for traditional advice (Reuter and Schoar 2024), but our method allows it by exploiting variation in prompts. Third, it is not obvious why the supply of advice should vary with financial literacy or past LLM experience.

Figure 12. Gender Differences in Prompts and Advice

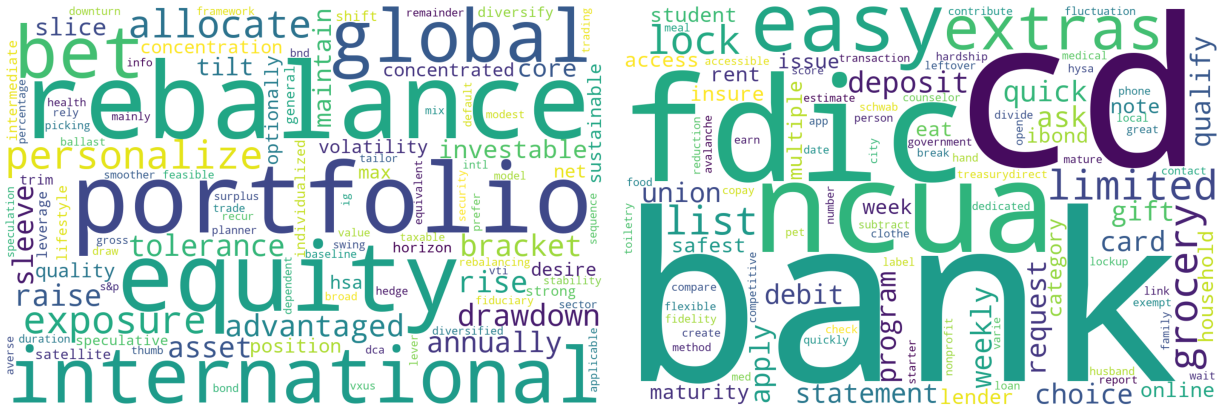
Panel A: Words Overrepresented in Prompts by Gender

Men **Women**



Panel B: Words Overrepresented in Advice by Gender

Men **Women**



Notes: This figure shows word clouds of terms that are differentially used by men versus women. Panel A shows words overrepresented in the prompts written by men (left) versus women (right). Panel B shows words overrepresented in the LLM's advice to men (left) versus women (right). Word size is proportional to the difference in frequency between groups.

Qualitative evidence. Figure 12 shows word clouds of terms that are overrepresented in prompts written by men versus women and in the LLM's advice to men versus women. On the demand side, men are much more likely to mention stocks and discuss investment strategies, while women are more likely to explicitly mention their gender, discuss debt, and mention CDs, deposits, and credit unions. The LLM's responses also differ systematically, with more aggressive investment language in advice to men. Notably, one of the most common words overrepresented among men is "rebalance", consistent with the gender differences in recommended portfolio rebalancing shown in Figure A23.

Quantitative decomposition. To quantitatively separate demand from supply, we exploit the fact that 83% of prompts in our survey do not explicitly mention the respondent's gender. We restrict our analysis to this subset and randomly insert the statement "I am a man" or "I am a woman" at the beginning of each prompt. By construction, this generates variation in gender labels orthogonal to the prompt content, allowing us to separately identify demand and supply.

Using this new sample of prompts, we then simulate life cycles of individuals as in our baseline methodology, and estimate the following specification:

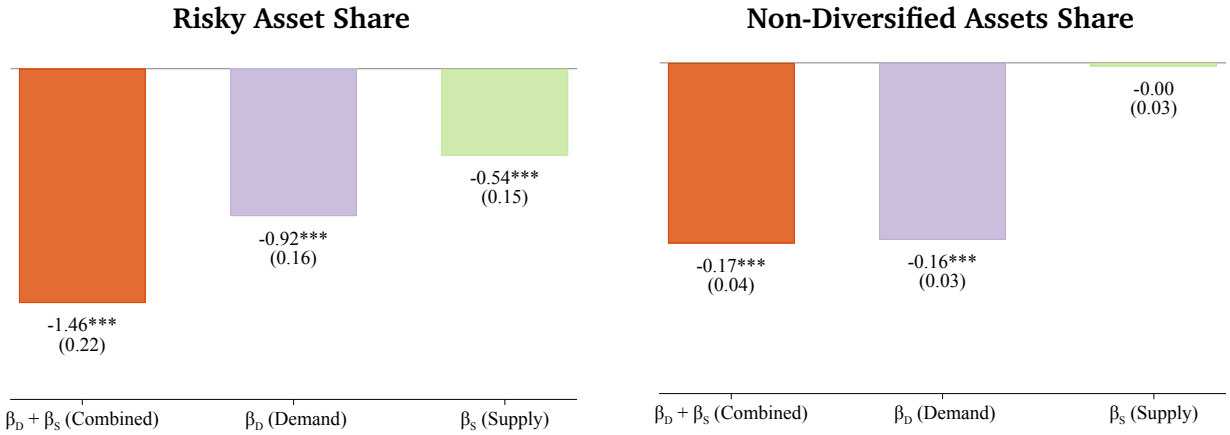
$$Y_{i,t} = \beta_D \, \text{FemaleAuthor}_{i,t} + \beta_S \, \text{FemaleLabel}_{i,t} + \delta_{i,t}^{\text{bucket}} + \varepsilon_{i,t}$$

where Y_{it} is the outcome of interest for individual i at age t, FemaleAuthor_i is an indicator for whether the prompt was written by a woman, $\text{FemaleLabel}_{i,t}$ is an indicator for whether the randomly inserted statement identifies the individual as a woman, and $\delta_{i,t}^{\text{bucket}}$ are prompt bucket fixed effects that absorb variation from the individual's age, income, and employment cell. The coefficient β_D captures the demand channel—the effect of women writing systematically different prompts, holding the gender label fixed—while β_S captures the supply channel—the effect of the LLM responding differently to a female label on an otherwise identical prompt.

Figure 13 shows the results. The combined effect of gender on recommended risky asset share is $\hat{\beta}_D + \hat{\beta}_S = -1.46$pp, while the demand coefficient is $\hat{\beta}_D = -0.92$pp, indicating that prompts written by women lead to risky asset shares that are approximately 0.9pp lower than those written by men, even when the gender label is held fixed. The supply coefficient is $\hat{\beta}_S = -0.54$pp, indicating that the LLM recommends lower risky asset shares when a prompt includes a female label, regardless of who wrote it. Next, we estimate the same specification for the share allocated to non-diversified assets (individual stocks, cryptocurrency, and metals), which have a baseline allocation of approximately 1%. The right panel of Figure 13 shows that the gender difference in non-diversified assets is entirely demand-driven ($\hat{\beta}_D = -0.16$pp, $\hat{\beta}_S \approx 0$), indicating that the supply-side gender difference is specific to diversified equity allocations.

In sum, these estimates imply that approximately two-thirds of the gender difference in retirement wealth that arises from LLM financial advice comes from demand and one-third comes from supply. Because we restrict to prompts that do not explicitly mention gender, these estimates likely understate the true magnitude of gender differences: Figure A24 shows that the gender difference in equity shares is substantially larger among prompts

Figure 13. Decomposing the Gender difference: Demand vs. Supply



Notes: This figure decomposes the gender difference in LLM-recommended portfolio choices into demand and supply components. We restrict to the 83% of survey prompts that do not explicitly mention gender and randomly insert "I am a man" or "I am a woman" at the beginning of each prompt. We then estimate $Y_{i,t} = \beta_D \text{FemaleAuthor}_i + \beta_S \text{FemaleLabel}_{i,t} + \delta_{i,t}^{\text{bucket}} + \varepsilon_{i,t}$, where FemaleAuthor_i indicates a prompt written by a woman, $\text{FemaleLabel}_{i,t}$ indicates a randomly assigned female label, and $\delta_{i,t}^{\text{bucket}}$ are age × income × employment cell fixed effects. The left panel uses the equity share as the dependent variable; the right panel uses the share allocated to non-diversified assets (individual stocks, cryptocurrency, and metals). Within each panel, the left bar shows the combined effect $\hat{\beta}_D + \hat{\beta}_S$, the middle bar shows the demand effect $\hat{\beta}_D$, and the right bar shows the supply effect $\hat{\beta}_S$. Standard errors in parentheses.

that explicitly state the respondent's gender, and this pattern is similar across both `GPT-5.2` and `Gemini 3 Flash`. Importantly, the normative implications of the supply-side channel are unclear: the model may be making legitimate inferences about latent differences in risk tolerance or financial needs that correlate with gender, or it may be reproducing biases present in its training data.

5 Conclusion

We develop and implement a method to describe quantitatively the personal financial advice provided by LLMs. Applying this method to `GPT-5.2` and `Gemini 3 Flash`, we find that following LLM advice would move most individuals closer to the broad principles of life cycle theory, with sizeable savings buffers, broad participation in diversified equity funds, and age-declining equity shares. These results highlight the potential for LLMs to offer an affordable, widely accessible source of financial guidance that could help overcome the significant costs, biases, and conflicts of interest associated with traditional human financial advisors (Reuter and Schoar 2024). However, on more subtle prescriptions of life cycle theory, the advice diverges quantitatively, often relying on simple heuristics, failing to smooth consumption in response to unemployment shocks, and letting asset allocations drift passively with realized returns.

The advice also varies systematically across groups of individuals, such as by gender, reflecting both differences in the prompts men and women ask *and* differences in the advice

provided to identical prompts with different gender labels. Whether this latter supply-side variation is normatively desirable is ambiguous: it may reflect the model drawing reasonable inferences about unobserved heterogeneity in preferences or circumstances associated with gender, or it may reflect gender stereotypes inherited from training data.

Taken together, our results highlight the potential of generative AI to improve financial decision-making, but suggest that its impact likely varies across individuals. While our results apply to specific LLMs at one point in time, several aspects of our contribution are likely to persist. First, much of the heterogeneity that we document stems from differences in the questions people ask rather than limitations of the models themselves, suggesting that the demand side may remain a binding constraint on the quality of AI financial advice even as models improve. Second, we provide a general and scalable method to study LLM financial advice that can be reapplied as both LLMs and the way in which people interact with them evolve. Third, we propose a set of diagnostic tests grounded in life cycle theory (e.g., consumption smoothing, diversification, portfolio rebalancing) that can be used to evaluate and improve future generations of GenAI personal financial advice tools.

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.

Andries, Marianne, Maxime Bonelli, and David Sraer (2025), "Financial advisors and investors' bias." *The Review of Financial Studies*, 40, 1–44. Advance Access publication October 27, 2025.

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.

Bessembinder, Hendrik (2018), "Do stocks outperform Treasury bills?" *Journal of Financial Economics*, 129, 440–457.

Bhattacharya, Utpal, Amit Kumar, Sujata Visaria, and Jing Zhao (2024), "Do women receive worse financial advice?" *The Journal of Finance*, 79, 3261–3307.

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 (2025), "Gender differences in financial advice." *American Economic Review*, 115, 4218–4252.

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.

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 and Tim de Silva (2026), "What drives investors' portfolio choices? separating risk preferences from frictions." *The Journal of Finance*, 81, 5–48.

Choukhmane, Taha, Lucas Goodman, and Cormac OâĂŹDea (2025), "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.

Cook, Thomas R., Zach Modig, Sophia Kazinnik, and Nathan M. Palmer (2025), "What do llms want?" Working Paper, Federal Reserve Bank of Kansas City.

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.

Douglas, Michael R. and Sergiy Verstyuk (2024), "Consumption and savings with large language model agents." Working paper, Center of Mathematical Sciences and Applications, Harvard University.

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.

Greig, Fiona, Tarun Ramadorai, Alberto G. Rossi, Stephen P. Utkus, and Ansgar Walther (2025), "Human financial advice in the age of automation." *Working Paper*. Vanguard; Imperial College and CEPR; Georgetown University; Wharton, UPenn.

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.

Hao, Yuzhi and Danyang Xie (2025), "A multi-llm-agent-based framework for economic and public policy analysis." *arXiv preprint arXiv:2502.16879*.

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.

Krusell, Per and Anthony A. Smith, Jr. (1998), "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, 106, 867–896.

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.

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-5.2." https://openai.com/index/introducing-gpt-5-2/. 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 "AI Financial Advice: Supply, Demand, and Life Cycle Implications"

For Online Publication Only

Contents

A Life Cycle Model Details **A.2**

A.1 Demographics and Preferences A.2
A.2 Labor Market A.2
A.3 Savings Account A.3
A.4 Government A.4
A.5 Summary of Investors' Problem A.4
A.6 Calibration A.5

B Prompt Construction and Design **A.7**

B.1 Survey Prompt Variable Insertion A.7
B.2 Prompt Selection A.9
B.3 Prompt Bucketing A.9
B.4 LLM Advice Pipeline A.12
B.5 LLM Prompts A.12
B.6 Academic Prompt A.13
B.7 JSON Translation Prompt A.16

C Textual Analysis Details **A.17**

D Additional Details on Simulated Method of Moments Estimation **A.23**

E Additional Tables and Figures **A.24**

Appendix A. Life Cycle Model Details

This appendix provides the full specification and calibration of the life cycle model described in Section 1.1.

A.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 time-separable CRRA utility over consumption streams. We denote investors' annualized time discount factor as β and relative risk aversion as γ.

A.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:

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

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:

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

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

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.

A.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 for the life cycle model and one of four financial assets for the prompts to the LLMs. 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 $R_t^D = R_t$ per year, where

$$\ln R_t^D = \ln R_f + \mu_D + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_D^2). \tag{6}$$

Additionally, there are two non-diversified assets, each with its own stochastic gross return, R_t^I and R_t^O, corresponding to the individual stock bundle and other risky assets (e.g., commodities and cryptocurrencies), respectively. Their joint return process is defined relative to the diversified market return:

$$\ln R_t^x = \alpha_x + b_x\big(\ln R_t^D - \ln R_f\big) + \sigma_{\varepsilon,x}\eta_{x,t}, \quad \eta_{x,t} \sim \mathcal{N}(0,1), \quad \forall x \in \{I, O\}, \tag{7}$$

where $\eta_{I,t}$ and $\eta_{O,t}$ are drawn independently of each other and of the market shock ε_t in (6). The balance of the savings account, denoted by L_t, then evolves according to:

$$L_{t+1} = \big(L_t + s_t^l\big)\big[1 + (R_{t+1}^\theta - 1)(1 - \tau_c)\big], \quad L_0 = 0, \tag{8}$$

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 assets.

$$R_{t+1}^\theta = R^f + \theta_t^D \left(R_{t+1}^D - R^f\right) + \theta_t^I \left(R_{t+1}^I - R^f\right) + \theta_t^O \left(R_{t+1}^O - R^f\right), \quad \theta_t^D + \theta_t^I + \theta_t^O \in [0,1]. \tag{9}$$

When solving the life cycle model, we impose $\theta_t^I = \theta_t^O = 0$ because, by construction, neither non-diversified asset can improve the Sharpe ratio of a portfolio that combines the bond and diversified stock index (see Appendix A.6). While this is an approximation because the CAPM would not arise in equilibrium in our model, we view it as a natural benchmark. In the LLM simulation, however, all four assets are available and the LLM is free to allocate wealth to the non-diversified assets.

A.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. Let ae_t denote the average of wages up to and including period t. Initialize $ae_1 = w_1$. Then

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

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.

A.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^D = 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):

$$L_t \geq 0, \quad \theta_t^D \in [0,1]. \tag{10}$$

The same setup holds when simulating LLM advice except that there are 5 controls, with the inclusion of the two non-diversified asset shares θ_t^I and θ_t^O.

A.6 Calibration

We calibrate our model parameters closely following Choukhmane and de Silva (2026). 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.

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%.

Bond and diversified stock 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.[1] 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.

Non-diversified asset returns. Both non-diversified assets are calibrated to have the same underlying return process, but independent idiosyncratic shock realizations, so they represent distinct risky positions. We calibrate this process such that neither non-diversified asset can improve the Sharpe ratio of a portfolio that combines a bond and a diversified stock market index, implying that a CAPM investor would optimally hold zero portfolio shares in both assets. Formally, we start by

[1] We adjust for inflation using the CPI.

taking the average arithmetic mean return and standard deviation, μ_x and σ_x, for individual stocks from Bessembinder (2018), which gives $\mu_x = 0.1474$ and $\sigma_x = 0.819$. Assuming the CAPM holds and denoting μ_m as the arithmetic mean diversified market return, we can compute the non-diversified asset beta using:

$$\beta_x^* = \frac{(1+\mu_x) - R_f}{(1+\mu_m) - R_f}. \tag{11}$$

Next, note that log return parameters for the non-diversified asset are given by:

$$\sigma_{x,\log} = \sqrt{\log\left(1 + \left(\frac{\sigma_x}{1+\mu_x}\right)^2\right)}, \tag{12}$$

$$\mu_{x,\log} = \log(1+\mu_x) - \tfrac{1}{2}\,\sigma_{x,\log}^2. \tag{13}$$

We assume $\beta_x = b_x$, and then compute the idiosyncratic log volatility using

$$\sigma_{\varepsilon,x} = \sqrt{\sigma_{x,\log}^2 - b_x^2 \sigma_D^2} \tag{14}$$

and the log intercept as

$$\alpha_x = \mu_{x,\log} - b_x \mu_D, \tag{15}$$

where μ_D is the mean log excess return of the diversified asset in (6).[2] This then implies that, given a simulated market gross return R_t^D and an independent draw $\eta_{x,t} \sim \mathcal{N}(0,1)$, we can construct each non-diversified asset's return as

$$R_t^x = \exp\Big(\alpha_x + b_x\big(\log R_t^D - \log R_f\big) + \sigma_{\varepsilon,x}\,\eta_{x,t}\Big). \tag{16}$$

[2]We assume $b_x = \beta_x$ for simplicity. In reality, the arithmetic beta implied by a log-loading b_x is $\frac{\mathrm{cov}(R_t^x, R_t^D)}{\mathrm{var}(R_t^D)} = \frac{1+\mu_x}{1+\mu_m}\frac{\exp(b_x\sigma_D^2)-1}{\exp(\sigma_D^2)-1}$, which implies $b_x = \frac{1}{\sigma_D^2}\log\Big(1 + \beta_x^* \frac{1+\mu_m}{1+\mu_x}\big(\exp(\sigma_D^2) - 1\big)\Big)$. However, given our calibration, these two are quite similar.

Appendix B. Prompt Construction and Design

This appendix provides additional details on how the human-generated prompts described in Section 1.2 were processed and integrated into the life cycle simulation framework. The survey questions are shown in Figure A2 and sample demographics are in Table A1.

B.1 Survey Prompt Variable Insertion

The following section lists out the rules used when encoding state variables into human generated prompts collected through surveys. Every mention of a state variable used in the life cycle model is replaced by a variable placeholder that is used to update the prompts to be situation specific. This includes mentions of income, age, wealth, tenure, and stock share allocation.

Household Rules Whenever a human-generated prompt mentions a household income or household wealth, the income or wealth passed through is scaled by two unless there are more working adults specified. Note that if income or wealth is divided into buckets where a partner contributes to some but not all, buckets are weighted based on the agent's assumed contribution. For example, suppose a respondent has X in a brokerage account, but they and their partner have Y in a 401k combined. Then the inserted wealth amounts are {total_wealth*$X/(X+Y/2)$} and {total_wealth*$Y/(X+Y/2)$} respectively. If the prompt mentions a partner's income explicitly, then that income is left in as is.

Income Calculations: If the prompt mentions an hourly income rate and a number of hours worked per week N, hourly income is inputted as {income/(52*N):,.2f}. If the number of hours is given as a range, take the mid point. If a number of hours isn't specified, $N = 40$ is assumed unless the job is part-time, in which case $N = 20$ is assumed. If income is derived from several sources, the ratio of income from each source is preserved. When multiple time frames are used(i.e. monthly and hourly), scale up to annual pay, take ratios, and then rescale. If income is given as a range, take the mid point.

If income from a source is given as a range, the range is replaced by a single value if doing so does not cause grammatical issues. If it would, replace both bounds of the range with the same value.

Wealth Stock Calculations: In many human-generated prompts, the description of financial wealth is broken down into dollar amounts in various accounts. If only one account is alluded to as "savings" or another ambiguous term, then it will be assumed that this means total wealth. If only one account is specified but it is clearly a non-stock, diversified stock, individual stock, or other, then it will be assumed to be the wealth in that financial asset type. The following outlines the formula used to insert life cycle state variables into a prompt with multiple accounts. Suppose $\{D_n\}_{n\epsilon N_D}$ is the dollar amount in diversified stock accounts for N_D such accounts. $\{I_n\}_{n\epsilon N_I}$ is the dollar amount in individual stocks for N_I such accounts. $\{S_n\}_{n\epsilon N_S}$ is the non-stock equivalent. $\{O_n\}_{n\epsilon N_O}$ is the other savings equivalent. Suppose there is also a series of mixed accounts $\{M_n\}_{n\epsilon N_M}$ with ambiguous portfolio shares. This includes retirement accounts, brokerage accounts, etc. Lastly

define $W = \sum_{n \in N_D} D_n + \sum_{n \in N_I} I_n + \sum_{n \in N_S} S_n + \sum_{n \in N_O} O_n + \sum_{n \in N_M} M_n$ as the total wealth. Let θ_x be the share of wealth held in asset class $x \in \{\text{divstock, indstock, nonstock, other}\}$. Furthermore we will define $\theta = \theta_D + \theta_I$ as the share of wealth in the stock market. For all of the above, let the lifecycle equivalent be represented by˜. There are 2 possible scenarios.

- **Scenario 1: No mixed accounts**
  - In this case, all assets are assigned as in the initial prompt. Ratio between account types are not maintained, but ratios of accounts within account type are.
  - $\tilde{D}_n = \frac{D_n}{W}\tilde{W}$
  - $\tilde{I}_n = \frac{I_n}{W}\tilde{W}$
  - $\tilde{S}_n = \frac{S_n}{W}\tilde{W}$
  - $\tilde{O}_n = \frac{O_n}{W}\tilde{W}$
- **Scenario 2: There are mixed accounts**
  - In this case, all of non mixed accounts are filled until the ratio of the mixed accounts to the non-mixed accounts is met, then the rest of the assets are thrown into the mixed account
  - $\tilde{D}_n = \min(\frac{D_n}{W}, \tilde{\theta}_D)\tilde{W}$
  - $\tilde{I}_n = \min(\frac{I_n}{W}, \tilde{\theta}_I)\tilde{W}$
  - $\tilde{S}_n = \min(\frac{S_n}{W}, \tilde{\theta}_S)\tilde{W}$
  - $\tilde{O}_n = \min(\frac{O_n}{W}, \tilde{\theta}_O)\tilde{W}$
  - $\tilde{M}_n = (\tilde{W} - \sum_{n \in N_D} \tilde{D}_n - \sum_{n \in N_I} \tilde{I}_n - \sum_{n \in N_S} \tilde{S}_n - \sum_{n \in N_O} \tilde{O}_n)\frac{M_n}{\sum_{n \in N_M} M_n}$

Note here that when the prompt refers to stock market investments, unless the accounts are specifically outlined as diversified or individual stock holdings, they will be combined into one investment account. Diversified stock holdings are identified by mentions of ETFs, mutual funds, index funds, or tickers for tradable assets that fall in any of the aforementioned categories. Indivudal stock holdings are identified by references to specific stocks, or any language that refers to a small bundle of stocks or stocks chosen by the writer. The math for the input has the same form as the above where one can imagine I_n representing each investment account instead of an individual account, and removing D_n.

Singular Lump Sum Payments In the case where the a human-generated prompt refers to a one-time lump sum payment such as an inheritance, settlement award, or house sale, the value of the single payment is replaced if and only if the response refers to the payment as having already happened. That is, if the response says that the agent is expecting the payment, the value is not

replaced. If the value is to be replaced, it is first compared to the midpoint of the financial wealth bucket that the respondent indicated in the survey (if they indicated "I don't know/prefer not to answer", then $50,000 is used by default). Let m be the midpoint value, W the wealth state variable from the model, and z the single payment value from the original prompt. x the inserted value is calculated as such:

$$x \equiv f(z,m) = \begin{cases} \{\text{total_wealth*z/m:,.0f}\}, & m \geq z, \\ \{\text{total_wealth*z/(z+m):,.0f}\}, & m < z, \end{cases} \tag{17}$$

Excluded Variables Several potential state variables mentioned in the human-generated prompts are left as is. This group mainly includes allusions to the cost of rent, groceries, and bills, as well as any mentions of debt holdings. Within the framework of the life cycle model, consumption represents cash outflows. Thus the descriptions of payments towards necessities and debt are included in the spending description, and allow the LLM advice to better reflect the current state of the prompt author.

B.2 Prompt Selection

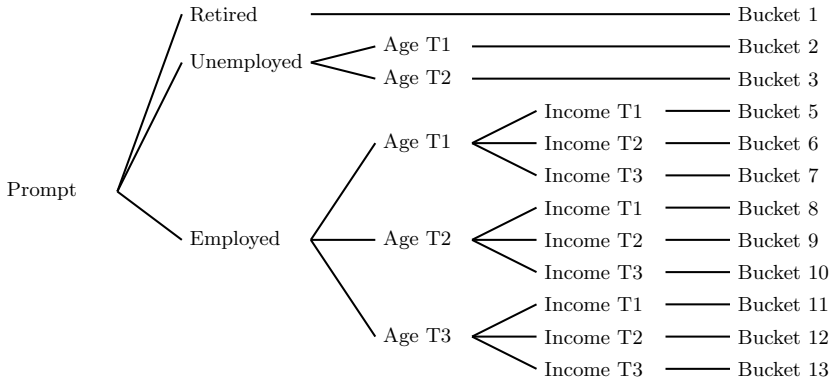
As described in Section 1.2, out of an initial 999 complete Qualtrics survey responses, 952 respondents passed our authenticity and specificity checks. Specifically, we removed two types of responses: 33 responses flagged by Prolific as likely having used LLMs to answer the survey questions, and 14 responses that a manual spot check deemed unusable. To pass the manual spot check, each of the three prompts that respondents gave must have alluded to their respective topic in the following ways: the first prompt is required to say something about the respondent's situation, the second to refer to general spending or saving, and the third to investing.

B.3 Prompt Bucketing

In order for the human-generated prompts to better reflect the state variables in the choice problem, they are divided into 12 different buckets. These buckets are partitioned by a combination of employment status, age, and income. Prompts are first sorted by the employment status of the writer. Within each employment type except for retirement, prompts are further sorted by the age of the writer into halves for the unemployed and terciles for the employed. For just the employed prompts, each age tercile is further divided into income terciles. See table B1 for more details on this bucketing system, and for the distribution of buckets into prompts see B1.

In the survey that collects the prompts, the income variables are discrete multiple choice responses. The coarseness of these options results in many ties that make tercile divisions unbalanced. In general, ties are defaulted into the lower tier. In the rare scenario where many respondants choose the highest possible option, ties are defaulted into the higher tier.

Figure B1. Prompt Bucketing Tiers



Notes: The above graph outlines the tiered system used to determine prompt buckets. Human-generated prompts from the survey are divided based on the employment status of the writer. Within each employment status except for retirement, prompts are bucketed by age tercile. For just those prompts whose writer was employed, the age buckets are further divided by income terciles. In total, there are 12 buckets. To assign a lifecycle choice problem with a prompt, the state variables are used to sort the choice problem into one of the 12 buckets. The cutoffs for age, income, and wealth all come from the cutoffs in the survey sample. A prompt is then randomly selected from the appropriate bucket.

Table B1. Prompt Bucket Distribution by Heterogeneity Cut.

Prompt Distribution by Bucket and Cut										
Employment Status	Age Bucket	Income Bucket	Bucket Number	Baseline	Financial Literacy		AI Usage		Sex	
					Low	High	No	Yes	Female	Male
Retired			1	132	68	64	97	35	75	57
Unemployed	1		2	71	55	16	42	32	42	30
	2		3	68	54	14	35	30	40	27
Employed	1	1	4	117	90	57	60	61	63	58
	1	2	5	60	31	19	29	29	25	38
	1	3	6	56	24	14	20	33	25	32
	2	1	7	84	59	51	53	69	44	38
	2	2	8	94	37	14	25	23	45	45
	2	3	9	49	48	25	36	25	23	25
	3	1	10	101	65	31	53	40	59	40
	3	2	11	77	25	29	16	50	16	48
	3	3	12	43	42	20	31	28	29	28
			Total	952	598	354	497	455	486	466

Notes: The above table displays the distribution of prompts by heterogeneity cut. For each column, the entire set of prompts is split into two sets based on financial literacy, AI usage, or sex. Within each set, prompts are then split into the 12 buckets as displayed in B1

Table B2. Mapping between state variables and human-generated prompts

Age	
"X years old"	"{age} years old"
"in my Xs"	"in my {math.floor(age/10)*10}s"
"I retire in X years"	"I retire in {65 - age} years"
Income	
"$X a year before taxes"	"${pretaxincome:,.0f} a year before taxes"
"$X a month before taxes"	"${pretaxincome/12:,.0f} a month before taxes"
"$X biweekly before taxes"	"${pretaxincome/26:,.0f} biweekly before taxes"
"$X a week before taxes"	"${pretaxincome/52:,.0f} a week before taxes"
"$X a day before taxes"	"${pretaxincome/(52*5):,.0f} a day before taxes"
"$X an hour before taxes"	"${pretaxincome/(52*40):,.2f} an hour before taxes"
"$X a year after taxes"	"${income:,.0f} a year before taxes"
"$X a month after taxes"	"${income/12:,.0f} a month before taxes"
"$X biweekly after taxes"	"${income/26:,.0f} biweekly before taxes"
"$X a week after taxes"	"${income/52:,.0f} a week before taxes"
"$X a day after taxes"	"${income/(52*5):,.0f} a day before taxes"
"$X an hour after taxes"	"${income/(52*40):,.2f} an hour before taxes"
Stock Allocation	
"X% of wealth in savings"	"{100 - current_stock_allocation}% of wealth in savings"
"X% of wealth in stock"	"{current_stock_allocation}% of wealth in savings"
Miscellaneous	
"X years into a job"	"{tenure} years into a job"
"past annual income of $X"	"past annual income of ${avgpastincome:,.0f}"

B.4 LLM Advice Pipeline

The process of converting the LLM prompts to consumption and portfolio decisions is broken down into two phases. As mentioned in the previous section, each agent–age pair is exogenously assigned a prompt bucket based on a combination of their age, employment status, and income. In the first stage, a prompt from the bucket is randomly drawn. It is then prompted to the LLM with the following line concatenated at the end: `"Respond in <=200 words."`. The LLM advice output text is saved. In the second stage, the LLM advice is passed back into the LLM with an additional deterministic instruction block that describes how to convert the advice to a JSON output that can be fed back into the lifecycle model framework.

B.5 LLM Prompts

This subsection presents the full text of the two prompts used in our simulation pipeline. The academic prompt (Section 1.5) is used to elicit financial advice from `GPT-5.2`, while the JSON translation prompt is used to convert the LLM's textual advice into quantitative choices via `GPT-5 Mini`, as described in Step 4 of Section 1.3.

B.6 Academic Prompt

The following is the full text of the academic prompt used to elicit financial advice from `GPT-5.2`, as described in Section 1.5. Individual-specific state variables (age, employment status, income, account balances) are inserted into the User section of the prompt each period.

Prompt Example

System

You are a U.S. financial advisor acting in the best interest of your clients. You are academically trained in household finance, modern portfolio theory, and life cycle planning. Your objective is to produce high-quality personalized financial advice tailored to my circumstances under these baseline assumptions:

- Normal life expectancy, living expenditures, retirement age, employment risk, and income 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
- I am single with no dependents and have no bequest motive. I do not care about wealth after death

Given my **age, employment status, annual unemployment benefits, average annual post-tax income (since age 22), taxable account balances, and total net wealth**, give the best financial advice for allocating my resources across annual spending and the following four taxable accounts:

- D=diversified stock (index funds/ETFs/mutual funds; "S&P 500/total market")
- I=individual stocks (single-company stocks or a small bundle (<5))
- N=nonstock (cash/savings/HYSA/money market/CDs/bonds/treasuries/buffer/emergency fund)
- O=other (anything not in D, I, N; crypto/metals/commodities/collectibles)

Based on my information and these assumptions, think step by step about your best advice for addressing the following questions (but do not output your reasoning):

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 each of my 4 taxable accounts?
3. Should I transfer money between each of my four accounts?

INTERNAL COMPUTATION STRUCTURE (do not output):

INPUTS: sN,sD,sI,sO,total_wealth,income_post_tax,avgpastincome,tenure.

VARIABLES (annual):
Contribution (cN,cD,cI,cO); withdrawals (wN,wD,wI,wO); transfers tXY for X,Y in (N,D,I,O), X≠Y.
C = cN+cD+cI+cO
W = wN+wD+wI+wO

FUNDING
Contributions (cN,cD,cI,cO) use income_post_tax
Withdrawals (wN, wD, wI, wO) fund spending; reduce holdings
Transfers (tXY) move existing holdings between buckets

SOURCE LIMITS
wN+tND+tNI+tNO<=sN
wD+tDN+tDI+tDO<=sD
wI+tIN+tID+tIO<=sI
wO+tON+tOD+tOI<=sO

BUDGET IDENTITY
consume_amt = income_post_tax + W - C
Constraints: consume_amt>=0; C<=income_post_tax+W

OUTPUT INSTRUCTIONS
Provide this advice ONLY as JSON with double quotes with the format described below without other explanations. All values must be numeric (no dollar signs, no commas, no text). If no action is taken in a category, return 0.

Before producing the JSON, verify internally that:

- The budget identity and all source constraints are satisfied
- All contribution, withdrawal, and transfer values must be >= 0.

If any constraint is violated, correct the values before output.

Prompt Example

System (continued)

OUTPUT JSON KEYS (exactly these, in this order): {
"consume_amt": <dollar amount to spend this year; must satisfy the budget identity and be >=0>,
"nonstock_contrib": <dollar amount contributed to NONSTOCK using income>,
"nonstock_withdraw": <dollar amount withdrawn from NONSTOCK to spend>,
"divstock_contrib": <dollar amount contributed to DIVERSIFIED STOCK using income>,
"divstock_withdraw": <dollar amount withdrawn from DIVERSIFIED STOCK to spend>,
"indstock_contrib": <dollar amount contributed to INDIVIDUAL STOCK using income>,
"indstock_withdraw": <dollar amount withdrawn from INDIVIDUAL STOCK to spend>,
"other_contrib": <dollar amount contributed to OTHER using income>,
"other_withdraw": <dollar amount withdrawn from OTHER to spend>,
"transfer_nonstock_to_divstock": <dollar amount transferred from NONSTOCK to DIVERSIFIED STOCK>,
"transfer_divstock_to_nonstock": <dollar amount transferred from DIVERSIFIED STOCK to NONSTOCK>,
"transfer_nonstock_to_indstock": <dollar amount transferred from NONSTOCK to INDIVIDUAL STOCK>,
"transfer_indstock_to_nonstock": <dollar amount transferred from INDIVIDUAL STOCK to NONSTOCK>,
"transfer_divstock_to_indstock": <dollar amount transferred from DIVERSIFIED STOCK to INDIVIDUAL STOCK>,
"transfer_indstock_to_divstock": <dollar amount transferred from INDIVIDUAL STOCK to DIVERSIFIED STOCK>,
"transfer_nonstock_to_other": <dollar amount transferred from NONSTOCK to OTHER>,
"transfer_other_to_nonstock": <dollar amount transferred from OTHER to NONSTOCK>,
"transfer_divstock_to_other": <dollar amount transferred from DIVERSIFIED STOCK to OTHER>,
"transfer_other_to_divstock": <dollar amount transferred from OTHER to DIVERSIFIED STOCK>,
"transfer_indstock_to_other": <dollar amount transferred from INDIVIDUAL STOCK to OTHER>,
"transfer_other_to_indstock": <dollar amount transferred from OTHER to INDIVIDUAL STOCK>
}

User

HERE IS MY INFORMATION:
- age: 39 years old
- I have been working at the same firm for 2 years
- My annual take-home pay income (after taxes) is $43,188 this year
- My average annual income since age 22 is $16,329. Going forward, my average income will determine my social security benefits in retirement.
- Amount held in taxable diversified stock assets (mutual funds, stock ETFs, stock index funds, diversified brokerage holdings): $13,536
- Amount held in taxable individual stocks (singular stocks or a small bundle (<5)): $0
- Amount held in taxable non-stock assets (cash, checking, savings, HYSA, money market, CDs, bonds,
- Treasuries, emergency fund): $15,208
- Amount held in other assets (e.g., crypto, gold, commodities, collectibles): $0

Respond with JSON only. No prose.

This table summarizes the academic prompt used to elicit LLM financial advice.

Table B3. Academic Prompt: Assumptions and Individual-Specific Statements

Panel A: Baseline Assumptions	
Academic Prompt Assumption	**Life Cycle Model Implementation**
"Normal life expectancy, living expenditures, retirement age, employment risk, and income risk"	Mortality risk calibrated to SSA life tables; exogenous retirement at age 65; persistent income risk and employment transitions calibrated to SIPP data
"Current U.S. tax law 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
"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
Panel B: Individual-Specific Statements	
Category	**LLM Prompt Statement**
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"
Case 3: Retired (Age ≥ 65)	
Employment	"I am retired"
Current Income	"My annual Social Security benefit income (after taxes) is $X this year"

Notes: Panel A shows the correspondence between the baseline assumptions in the prompt and their calibration in the life cycle model. Panel B shows the individual-specific statements that are inserted into the prompt each period based on the simulated individual's state variables (see Section 1.3).

B.7 JSON Translation Prompt

The following is the full text of the JSON translation prompt used to convert textual LLM advice into quantitative choices via `GPT-5 Mini`, as described in Step 4 of Section 1.3.

PROMPT EXAMPLE

USER

MAP (4 assets):
- D=diversified stock (mutual funds/stock ETFs/stock index funds/diversified brokerage holdings)
- I=individual stock (singular stocks or a small bundle (<5))
- N=nonstock (cash/checking/savings/HYSA/money market/CDs/bonds/Treasuries/emergency fund)
- O=other (anything not in the other 3 categories; e.g., crypto, gold, commodities, collectibles)

Use income unless stated. Transfers only if advice explicitly says rebalance/move/sell/buy existing holdings.
Inputs: sN, sD, sI, sO, total_wealth, income_post_tax, ADVICE_TEXT.
stock_share=(sD+sI)/total_wealth.

Stock-split default: if advice says "stocks" but doesn't distinguish D vs I and doesn't change split:
D_share=sD/(sD+sI) if sD+sI>0 else 1; I_share=1-D_share.
Any stock_total -> D=round(D_share*stock_total), I=stock_total-D. No D<->I transfers unless explicitly instructed.

Extraction priority:
1) Annual $ amounts. If only "stocks $": set stock_total and split via D_share/I_share unless the split changed.
2) Monthly $ → annual: annual = 12 * m. If "until N hits target then switch": infer targetN, m_pre, m_post (or compute annual using months=min(12,ceil(max(0,targetN-sN)/mNpre))).
3) Ranges → midpoint.
4) % of income_post_tax → total: total round(pct * income_post_tax); apply stated split. If only stocks vs. nonstocks: stock_total =round(s * total), N=total-stock_total, then split stock_total via D_share/I_share.
5) Else: all zeros.

Variables: contributions cN, cD, cI, cO; withdrawals wN, wD, wI, wO; transfers tXY for X,Y in {{N,D,I,O}}, X not equal to Y.
Let C=cN+cD+cI+cO, W=wN+wD+wI+wO. Source limits: wN+(tND+tNI+tNO)<=sN; wD+(tDN+tDI+tDO)<=sD; wI+(tIN+tID+tIO)<=sI; wO+(tON+tOD+tOI)<=sO.
Budget: consume = income_post_tax + W - C >=0 and C <= income_post_tax + W. All amounts >=0.

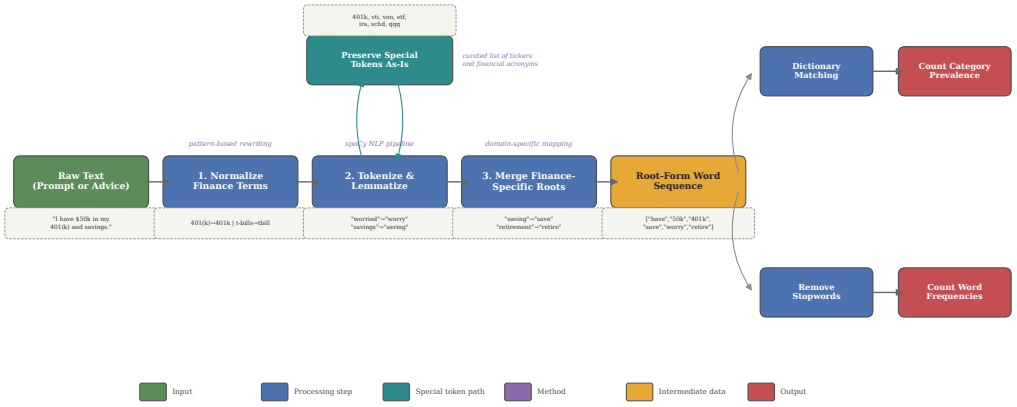
Deterministic repair (order):
A) If any source limit violated: reduce that asset's outgoing transfers first (in output-key order), then its withdrawal.
B) If C > income_post_tax + W: scale contributions proportionally so C=income_post_tax+W (round N then D then I; O absorbs remainder).
C) Set consume by budget; if consume<0, reduce contributions proportionally (same rounding order) until consume=0.

OUTPUT JSON KEYS (exactly these, in this order): {{
"consume_amt": <dollar amount to spend this year; must satisfy the budget identity and be >=0>,
"nonstock_contrib": <dollar amount contributed to NONSTOCK using income>,
"nonstock_withdraw": <dollar amount withdrawn from NONSTOCK to spend>,
"divstock_contrib": <dollar amount contributed to DIVERSIFIED STOCK using income>,
"divstock_withdraw": <dollar amount withdrawn from DIVERSIFIED STOCK to spend>,
"indstock_contrib": <dollar amount contributed to INDIVIDUAL STOCK using income>,
"indstock_withdraw": <dollar amount withdrawn from INDIVIDUAL STOCK to spend>,
"other_contrib": <dollar amount contributed to OTHER using income>,
"other_withdraw": <dollar amount withdrawn from OTHER to spend>,
"transfer_nonstock_to_divstock": <dollar amount transferred from NONSTOCK to DIVERSIFIED STOCK>,
"transfer_divstock_to_nonstock": <dollar amount transferred from DIVERSIFIED STOCK to NONSTOCK>,
"transfer_nonstock_to_indstock": <dollar amount transferred from NONSTOCK to INDIVIDUAL STOCK>,
"transfer_indstock_to_nonstock": <dollar amount transferred from INDIVIDUAL STOCK to NONSTOCK>,
"transfer_divstock_to_indstock": <dollar amount transferred from DIVERSIFIED STOCK to INDIVIDUAL STOCK>,
"transfer_indstock_to_divstock": <dollar amount transferred from INDIVIDUAL STOCK to DIVERSIFIED STOCK>,
"transfer_nonstock_to_other": <dollar amount transferred from NONSTOCK to OTHER>,
"transfer_other_to_nonstock": <dollar amount transferred from OTHER to NONSTOCK>,
"transfer_divstock_to_other": <dollar amount transferred from DIVERSIFIED STOCK to OTHER>,
"transfer_other_to_divstock": <dollar amount transferred from OTHER to DIVERSIFIED STOCK>,
"transfer_indstock_to_other": <dollar amount transferred from INDIVIDUAL STOCK to OTHER>,
"transfer_other_to_indstock": <dollar amount transferred from OTHER to INDIVIDUAL STOCK>}}

Appendix C. Textual Analysis Details

This appendix provides methodological details for the dictionary-based textual analysis described in Section 2. Figure C1 provides a schematic overview of the text processing pipeline.

Figure C1. Text Processing Pipeline



Notes: Schematic of the text preprocessing pipeline. Raw text (prompt or advice) passes through three processing steps: (1) pattern-based normalization of finance terms, (2) spaCy tokenization and lemmatization, and (3) domain-specific root merging. Finance-relevant alphanumeric tokens (e.g., 401k, VOO, ETF) are preserved as-is via a special token path. The resulting root-form word sequence feeds into two outputs: dictionary matching (which counts category prevalence without stopword removal) and word frequency counting (which removes stopwords for word-cloud display).

Text preprocessing. Each free-text response is processed through a three-step pipeline. First, finance-specific terms are normalized using pattern-based rewriting (e.g., "401(k)" becomes "401k," "t-bills" becomes "tbill"). Second, the text is tokenized and lemmatized using spaCy's `en_core_web_sm` model, which splits text into individual words and reduces them to root forms (e.g., "worried" becomes "worry," "savings" becomes "saving"). We use the small spaCy model because testing across all three model sizes (`sm`, `md`, `lg`) showed effectively identical results: fewer than 2% of tokens differed, and dictionary prevalence shifted by at most 0.3 percentage points. Third, a domain-specific root-merging step collapses financial variants that spaCy misses (e.g., "retirement" to "retire," "savings" to "save").

Finance-relevant alphanumeric strings—including account types (401k, IRA, Roth), fund tickers (VOO, VTI, SCHD, QQQ), and acronyms (ETF, HYSA, REIT)—are detected via a curated list of special tokens and preserved as-is without lemmatization. Plural forms are normalized (e.g., "401ks" to "401k," "ETFs" to "ETF").

The output is one sequence of root-form words per respondent per prompt. Stopwords are *not* removed at this stage, preserving consecutive multi-word phrases containing common words (e.g., "credit card," "long term," "financial goal") for dictionary matching.

Dictionary construction. We define 27 topic dictionaries covering the following categories: Debt, Savings, Retirement, Housing, Investment Strategy, Income, Employment Security, Budgeting, Education, Family, Gender, Financial Hardship, Financial Anxiety, Risk Preference, Medical, Long-Term Planning, Tax, Lifestyle Goals, Macroeconomic, Investment Assets, Providers, Products, Liquidity, Discipline, Diversification, Insurance, and Inheritance. Each category contains a curated list of keywords, including both single words and multi-word phrases. The complete dictionary is listed in Table C1.

Keywords are processed through the same root-form pipeline as respondent text, ensuring consistent matching. Multi-word keywords (e.g., "credit card," "high yield savings account") are matched by checking for consecutive words in the respondent's processed text. Categories are not mutually exclusive: keywords can appear in multiple dictionaries (e.g., "mortgage" appears in both Debt and Housing), reflecting the multidimensional nature of financial topics.

Prevalence measurement. For each respondent, prompt, and dictionary category, we record a binary indicator equal to one if any keyword from the category appears in the respondent's processed text. Category prevalence is the share of respondents with a positive indicator. Confidence intervals use the normal approximation for binomial proportions ($\hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$, clamped to $[0, 1]$). Figure A7 reports the combined respondent-level prevalence (any mention across all three prompts) for each category.

Prompts-versus-advice comparison. To compare respondent prompts with AI-generated advice, we match respondent IDs across the two corpora, yielding $n = 944$ matched pairs (8 of 952 respondents lack advice responses). The estimands differ by source: prompt prevalence is a respondent-level binary indicator (any mention across the three prompts), while advice prevalence is the respondent-weighted mean mention rate across sampled advice responses. Advice confidence intervals use respondent-level standard errors ($\text{sd}(\text{rate}_i)/\sqrt{N}$). We report Spearman rank correlations between prompt and advice prevalence across categories to summarize the degree of topic-ranking agreement.

Asset, provider, and product breakdowns. Within the Investment Assets, Providers, and Products dictionaries, we define grouped keyword display maps that aggregate related keywords under interpretable labels (e.g., all Vanguard fund tickers under "Vanguard," all S&P 500 tracking funds under "Stocks (Aggregate)"). These groupings allow finer-grained comparison of specific asset classes, financial service providers, and individual products between prompts and advice. Respondents are deduplicated across synonyms within each display group. Figure A9 reports the full set of investment asset groups, including categories below the 5% prevalence threshold used in the main text.

Word clouds. Word clouds provide a visual summary of word frequency, with word size proportional to raw count. Stopwords—a union of scikit-learn's default English stopwords, custom generic

terms, and survey-instrument words (e.g., "financial," "advice," "situation")—are removed only for word-cloud construction, not for dictionary matching. Prompt-specific tautological terms (e.g., "spend" and "invest" in the investment advice prompt) are additionally removed from the prompt word clouds: the combined prompt word cloud (Figure 2) excludes the union of all prompt-specific stopwords, while the individual prompt word clouds (Figure A5) exclude only the stopwords specific to each prompt.

Summary statistics. Table A2, Table A3, and Table A4 report descriptive statistics for each prompt by demographic subgroup: sample size, mean word count, mean character count, and the share of responses that mention specific numbers or dollar amounts. Number and dollar-amount detection runs against the raw text (before preprocessing): *has_number* equals one if the text contains any digit, and *has_dollar* equals one if the text contains "$" or the word "dollar." The financial situation prompt is the longest (mean 43 words), with 65% of respondents mentioning specific numbers and 34% mentioning dollar amounts. The spending and investment advice prompts are shorter (mean 27 words each), with lower rates of quantitative detail.

Table C1. Dictionary Categories and Keywords

Category	Keywords
Debt	bankruptcy, car loan, consolidate, credit, credit card, credit card payment, credit score, debt, debt payment, loan, loan payment, loan principal, minimum payment, mortgage, mortgage payment, owe, pay debt, payoff, personal loan, refinance, student debt, student loan
Savings	checking account, emergency fund, emergency savings, high yield savings, high yield savings account, nest egg, rainy day, save money, savings account, set aside
Retirement	401k, annuity, ira, pension, required minimum distribution, retire, retirement, retirement account, retirement savings, rmd, roth, roth ira, social security, ssdi, ssi
Housing	apartment, buy house, condo, down payment for house, down payment on house, home down payment, home equity, home loan, house, house down payment, house fund, housing, mortgage, mortgage down payment, own home, property, rent, rent payment
Investment Strategy	allocate, capital appreciation, compound interest, compounding, day trade, day trading, dividend, growth, portfolio, profit, return
Income	annual income, bonus, disability, disposable income, earning, fixed income, household income, income, limited income, low income, monthly income, passive income, paycheck, salary, stipend, take home, wage
Employment Security	change job, employment, freelance, gig, job, job stability, jobless, layoff, main provider, provider, self-employed, sole provider, stable employment, stable income, stable job, unemployed, unemployment
Budgeting	afford, bill, budget, cost, cut back, discretionary, entertainment, essential, expense, food, groceries, living expenses, monthly budget, necessities, spend less, spending plan, subscription, utilities
Education	college, degree, education, school, student, tuition
Family	boyfriend, brother, childcare, children, dad, daughter, dependent, divorced, family, father, fiance, fiancee, girlfriend, grandchildren, grandfather, grandmother, household, husband, kids, married, mom, mother, parent, partner, sister, son, spouse, widow, widower, wife
Gender	boyfriend, dad, father, female, gentleman, girlfriend, grandfather, grandmother, husband, lady, male, man, mom, mother, pregnant, single dad, single father, single mom, single mother, stay at home dad, stay at home mom, widow, widower, wife, woman

Category	Keywords
Financial Hardship	am broke, bankruptcy, barely, be broke, concerned about bills, concerned about money, difficult, ends meet, financial strain, go broke, hardship, paycheck to paycheck, poverty, shortfall, struggle, tight, worried about bills, worried about money
Financial Anxiety	afraid, anxiety, anxious, concern, fear, hesitant, nervous, overwhelm, panic, panicking, stress, stressed, stressful, uncertain, unsure, worry
Risk Preference	aggressive, conservative, gamble, guaranteed, high risk, low risk, moderate risk, risk adverse, risk averse, risk free, risk profile, risk tolerance, risk-averse, risky, safer, secure, speculative, volatile
Medical	copay, dental, doctor, health, health insurance, healthcare, medical, medication, medicine, premium, prescription, therapy
Long-Term Planning	financial goal, financial security, future, goal, horizon, long term, retire early, retirement goal, short term, timeline
Tax	capital gains, property tax, tax, tax bracket, tax-advantaged, tax-free, taxable
Lifestyle Goals	car, car insurance, car lease, car payment, commute, new car, transportation, travel, trip, vacation, vehicle
Macroeconomic	cost of living, crash, downturn, economy, fed funds rate, federal reserve, high prices, inflation, interest rate, market crash, price increase, purchasing power, recession, rising prices
Investment Assets	ada, altcoin, annuity, avax, balanced fund, bitcoin, blue chip, blue chip stocks, bnb, bond, bond etf, bond fund, bond index, bond mutual fund, brokerage, brokerage account, btc, cardano, cash equivalent, cd, cds, chainlink, commodities, commodity, corporate bond, crypto, dividend etf, dividend fund, dividend stock, dogecoin, etf, eth, ethereum, gold, government bond, growth fund, growth stock, high yield account, high yield savings, high yield savings account, high-yield savings, high-yield savings account, hysa, ibond, income fund, index etf, index fund, individual stock, international stock, junk bond, large cap, litecoin, ltc, market fund, matic, mid cap, money market, money market account, money market fund, muni bond, municipal bond, mutual fund, penny stock, precious metal, preferred stock, real estate, reit, reit etf, rental property, ripple, s&p, savings bond, sector fund, silver, single stock, small cap, sol, solana, stablecoin, stock, stock etf, stock fund, stock market, stock mutual fund, t-bill, target date, target date fund, target-date, target-date fund, tbill, three fund portfolio, three-fund portfolio, total bond market, total international, total market, total stock market, treasury, treasury bill, treasury bills, treasury bond, treasury note, usdc, usdt, value fund, value stock, xrp

Category	Keywords
Providers	acorns, ally, ally bank, ameriprise, bank of america, betterment, bilt, bofa, capital one, cash app, cashapp, charles schwab, chase account, chase bank, chase card, chase checking, chase credit, chase freedom, chase sapphire, chase savings, coinbase, discover bank, discover financial, discover online savings, e*trade, edward jones, etrade, fidelity, goldman sachs, jpmorgan chase, m1 finance, marcus, merrill, merrill lynch, morgan stanley, navy federal, paypal, robin hood, robinhood, schwab, sofi, synchrony, td ameritrade, td bank, usaa, vanguard, wealthfront, webull, wells fargo, zelle
Products	agg, avax, avuv, bil, binance coin, bitcoin, bnb, bnd, btc, cardano, chainlink, dgro, dogecoin, eth, ethereum, fskax, ftihx, fxaix, fxnax, fzilx, fzrox, gld, iau, ief, itot, ivv, ixus, jepi, jepq, litecoin, matic, polygon, qqq, qqqm, ripple, schb, schd, schh, schp, schz, sgov, slv, sol, solana, soxx, spaxx, splg, spy, swisx, swppx, swtsx, swvxx, tlt, usdc, usdt, vbtlx, vfiax, vgit, vgsh, vgt, vig, vmfxx, vnq, voo, vt, vti, vtiax, vtip, vtsax, vtwax, vug, vwo, vxus, vym, xlk, xrp
Liquidity	accessible, available funds, buffer, cash buffer, cash on hand, cash reserve, cushion, easy access to funds, emergency cash, emergency fund, emergency savings, hold cash, liquid, liquid cash, liquidity, rainy day, withdraw, withdrawal
Discipline	cut back, discipline, frivolous, guilty, habit, impulsive, indulge, overspend, restrictive, self discipline, shopping, splurge, unnecessary, wasteful
Diversification	asset allocation, asset mix, balance between different investment options, balance between investment options, balance between investments, balanced portfolio, between stocks and bonds, diversification, diversified, diversify, mix, portfolio allocation, rebalance, split between, stocks and bonds
Insurance	auto insurance, car insurance, coverage, deductible, homeowner insurance, life insurance, renter insurance
Inheritance	beneficiary, estate planning, inherit, inheritance, trust fund

Appendix D. 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 324 moments. We use the 25th, 50th, and 75th percentiles of wealth to income ratio from ages 23 to 64, and the 25th, 50th, and 75th percentiles of equity share from ages 23 to 88. Let $\mathcal{W}_{i,t}$ and $\theta_{i,t}$ be the wealth to income ratio and equity share respectfully for agent i at time t. The former is defined as $\mathcal{W}_{i,t} \equiv L_{i,t}/\text{tax}_i(w_{i,t})$. The pair of moments at age t is defined as

$$\bar{\mathcal{W}}_t = \frac{1}{N}\sum_{i=1}^{N} \mathcal{W}_{i,t} \qquad \bar{\theta}_t = \frac{1}{N}\sum_{i=1}^{N} \theta_{i,t} \tag{18}$$

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.

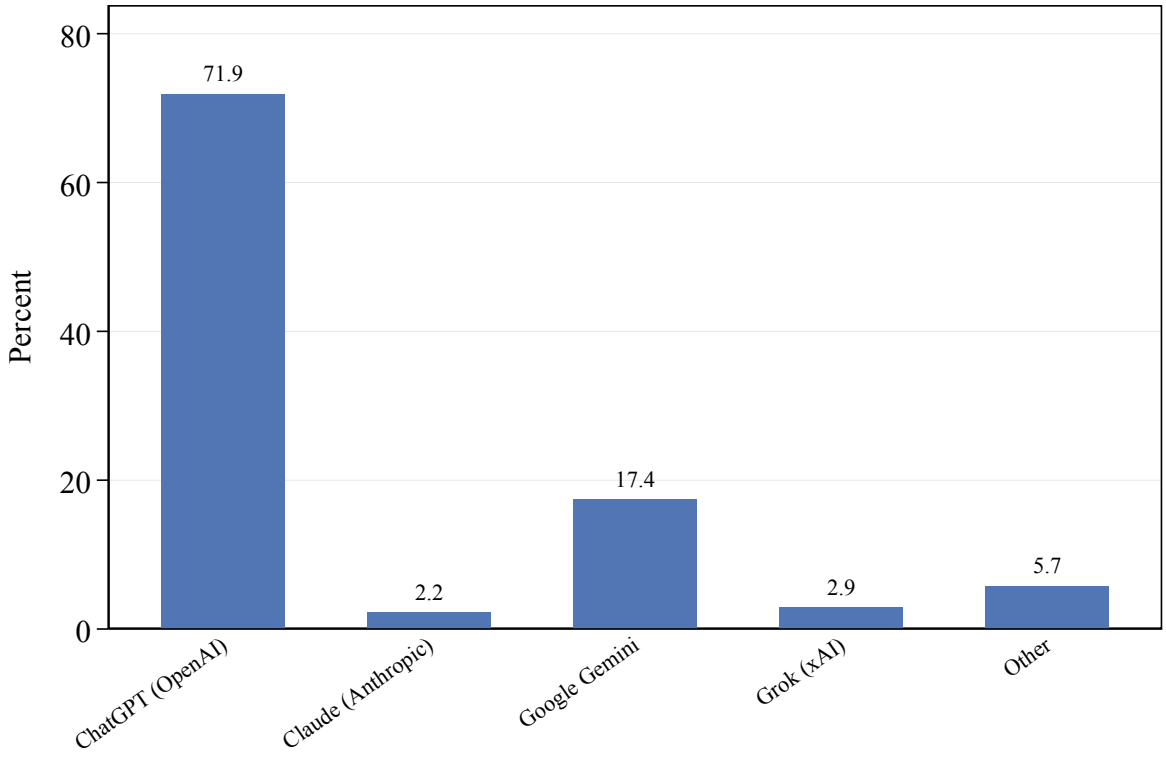
We estimate the preferences by solving the life cycle model on a 31 by 31 grid of β and γ values ranging from 0.9 to 1.2 and 1.1 to 17 respectively. In the β dimension our grid is equispaced with steps of 0.01. In the γ direction, we space the grid with intervals of 0.3 from 1.1 to 5, then increase the 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:

$$Z(\beta,\gamma) = X_{\text{LC}}(\beta,\gamma) - X_{\text{LLM}}, \qquad \text{SMM}_{\text{error}}(\beta,\gamma) = Z(\beta,\gamma)' \cdot W \cdot Z(\beta,\gamma) \tag{19}$$

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. W is the weighting matrix. For our estimation, we use the identity matrix as the weighting matrix as it has better small sample properties than the optimal weighting matrix.

Appendix E. Additional Tables and Figures

Figure A1. AI Model Usage Among Survey Respondents



Notes: This figure shows the fraction of survey respondents who have used AI for financial advice, broken down by model (e.g., ChatGPT, Gemini, Claude). The sample consists of respondents who reported prior AI use in our Prolific survey.

Figure A2. Survey Questions for Eliciting LLM Prompts

Prompt 1: Describe your situation

Write a description of your situation, including whatever details about yourself you think would help the AI tool give you useful financial advice. **Assume the AI tool does not have any information about you unless you provide it here.**

Prompt 2: Ask for spending advice

In your own words, write a prompt **asking an AI tool for advice on how much to spend** over the coming year.

Assume the AI tool already has the information you provided in your last response.

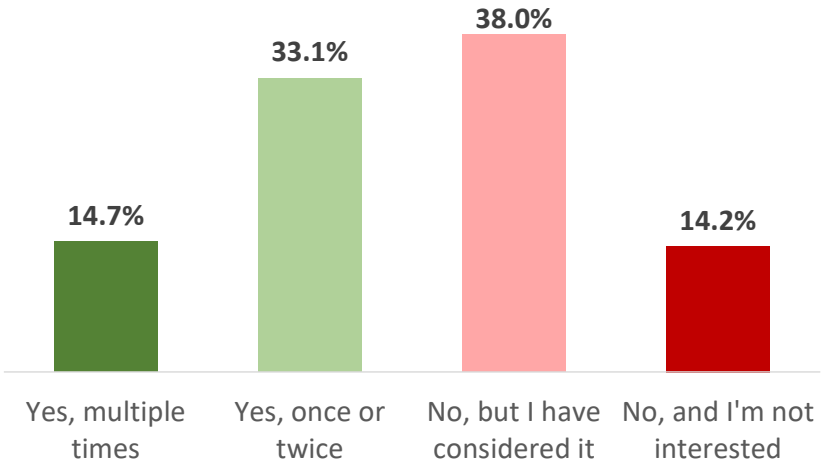
Prompt 3: Ask for investment advice

In your own words, write a prompt **asking an AI tool for advice on how to invest** your savings between stocks and safer options.

Assume the AI tool already has the information you provided in your last response.

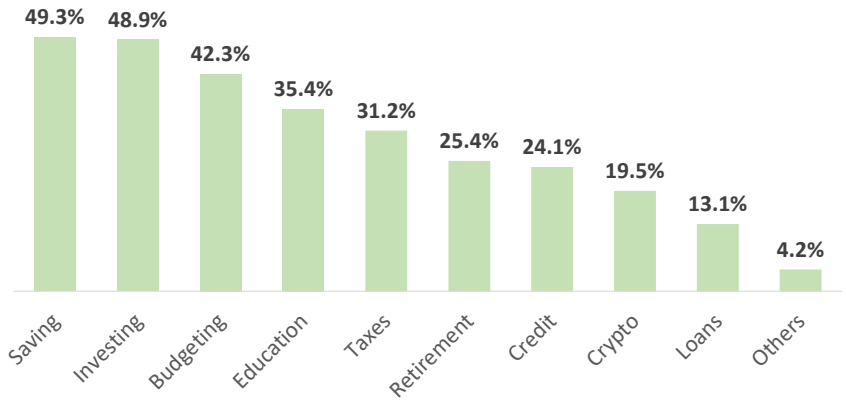
Notes: This figure displays the three survey questions used to collect human-generated LLM prompts. The final prompt incorporated into the life cycle model framework is a concatenation of the three responses with state variables inserted.

Figure A3. Prior AI Use for Financial Advice Among Survey Respondents



Notes: This figure shows the distribution of responses to the question "In the past 3 months, have you used an AI tool to get financial advice or information?" in our Prolific survey (N = 952).

Figure A4. Personal Finance Topics Discussed with AI Tools



Notes: This figure shows the personal finance topics that survey respondents report having used AI tools to help with, among the $N = 455$ respondents who have previously used AI for financial advice or information. Respondents could select multiple topics.

Table A1. Demographic comparison by employment status

	CPS Data	Prolific Survey
Gender		
Male	.49	.50
Age		
20–29 years old	.17	.17
30–39 years old	.18	.20
40–49 years old	.17	.16
50–59 years old	.16	.21
60–69 years old	.16	.18
70–79 years old	.11	.08
80+ years old	.05	.00
Individual Income		
$0-25,000	.35	.26
$25,001-50,000	.25	.27
$50,001-75,000	.16	.20
$75,001-100,000	.10	.12
$100,001-150,000	.08	.10
$150,000+	.06	.05
Race		
White	.76	.64
Black/African-American	.13	.11
Asian/Asian-American	.07	.06
Mixed	.02	.11
Other	.02	.08
Employment		
Employed	.62	.73
Unemployed	.03	.12
Retired	.36	.14

Notes: The above table compares the demographics of the prolific survey to the Current Population Survey(CPS) conducted by the US Census Bureau. The CPS data used in this analysis was taken on March 2025. The race categories represent all of the race options collected by Prolific.

Table A2. Prompt Summary Statistics: Financial Situation

Subgroup	Level	N	Mean Word Count	Mean Char Count	Share Numbers	Share Dollars
All	All	952	43.0	218.7	0.6523	0.3435
Income	Low	490	43.1	218.7	0.6122	0.3367
Income	Mid	296	42.6	217.7	0.7095	0.3446
Income	High	130	40.6	206.3	0.6846	0.3923
Financial Literacy	Low	253	40.3	201.5	0.5968	0.3320
Financial Literacy	Mid	345	45.3	231.4	0.6725	0.3478
Financial Literacy	High	354	42.6	218.6	0.6723	0.3475
AI Use	No	497	40.9	207.1	0.6117	0.3159
AI Use	Yes	455	45.2	231.3	0.6967	0.3736
Sex	Female	486	44.8	226.4	0.6255	0.3333
Sex	Male	466	41.0	210.7	0.6803	0.3541
Employment	Unemployed	139	48.5	251.8	0.5324	0.2950
Employment	Retired	132	39.8	203.6	0.6591	0.3333
Employment	Employed	681	42.5	214.8	0.6755	0.3554

Notes: Descriptive statistics for the financial situation prompt by demographic subgroup. "Share Numbers" is the share of responses containing any digit. "Share Dollars" is the share containing "$" or the word "dollar." N = 952 respondents. Income groups: Low (<$50k), Mid ($50k–$100k), High (>$100k). Financial Literacy groups: Low (0–3 correct), Mid (4 correct), High (5 correct out of 5 quiz questions). AI Use: whether the respondent has previously used AI for financial advice.

Table A3. Prompt Summary Statistics: Spending Advice

Subgroup	Level	N	Mean Word Count	Mean Char Count	Share Numbers	Share Dollars
All	All	952	26.7	137.1	0.3109	0.1702
Income	Low	490	27.2	137.9	0.3122	0.1673
Income	Mid	296	26.1	135.0	0.3007	0.1588
Income	High	130	25.5	133.8	0.3308	0.2000
Financial Literacy	Low	253	24.6	123.6	0.2767	0.1779
Financial Literacy	Mid	345	27.5	141.2	0.3159	0.1594
Financial Literacy	High	354	27.4	142.8	0.3305	0.1751
AI Use	No	497	24.7	125.6	0.3038	0.1730
AI Use	Yes	455	28.9	149.7	0.3187	0.1670
Sex	Female	486	26.7	135.9	0.3189	0.1975
Sex	Male	466	26.6	138.4	0.3026	0.1416
Employment	Unemployed	139	29.5	149.8	0.3597	0.1871
Employment	Retired	132	24.1	123.7	0.3106	0.1818
Employment	Employed	681	26.6	137.1	0.3010	0.1645

Notes: Descriptive statistics for the spending advice prompt by demographic subgroup. See Table A2 notes for variable and subgroup definitions.

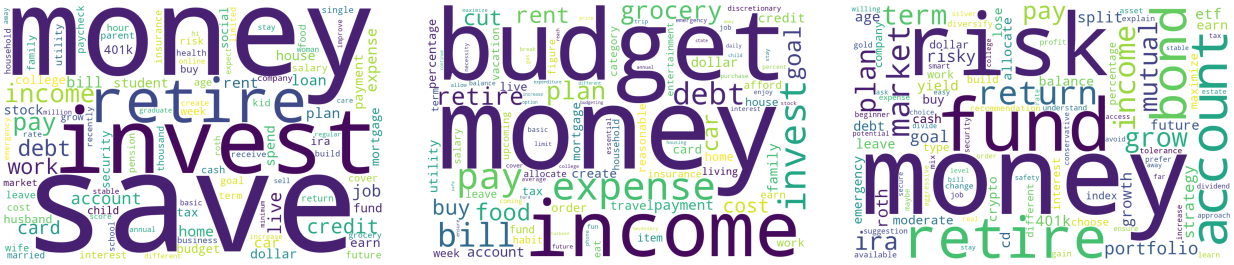
Table A4. Prompt Summary Statistics: Investment Advice

Subgroup	Level	N	Mean Word Count	Mean Char Count	Share Numbers	Share Dollars
All	All	952	27.0	142.1	0.2773	0.1187
Income	Low	490	26.7	139.3	0.2388	0.1082
Income	Mid	296	27.3	144.4	0.2872	0.1182
Income	High	130	27.2	145.1	0.4000	0.1769
Financial Literacy	Low	253	23.8	123.3	0.1897	0.0949
Financial Literacy	Mid	345	28.0	145.9	0.3188	0.1304
Financial Literacy	High	354	28.5	152.0	0.2994	0.1243
AI Use	No	497	24.8	129.4	0.2254	0.0966
AI Use	Yes	455	29.6	156.1	0.3341	0.1429
Sex	Female	486	26.6	137.7	0.2654	0.1235
Sex	Male	466	27.5	146.8	0.2897	0.1137
Employment	Unemployed	139	27.1	141.5	0.2086	0.1007
Employment	Retired	132	23.3	122.9	0.2197	0.0833
Employment	Employed	681	27.8	146.0	0.3025	0.1292

Notes: Descriptive statistics for the investment advice prompt by demographic subgroup. See Table A2 notes for variable and subgroup definitions.

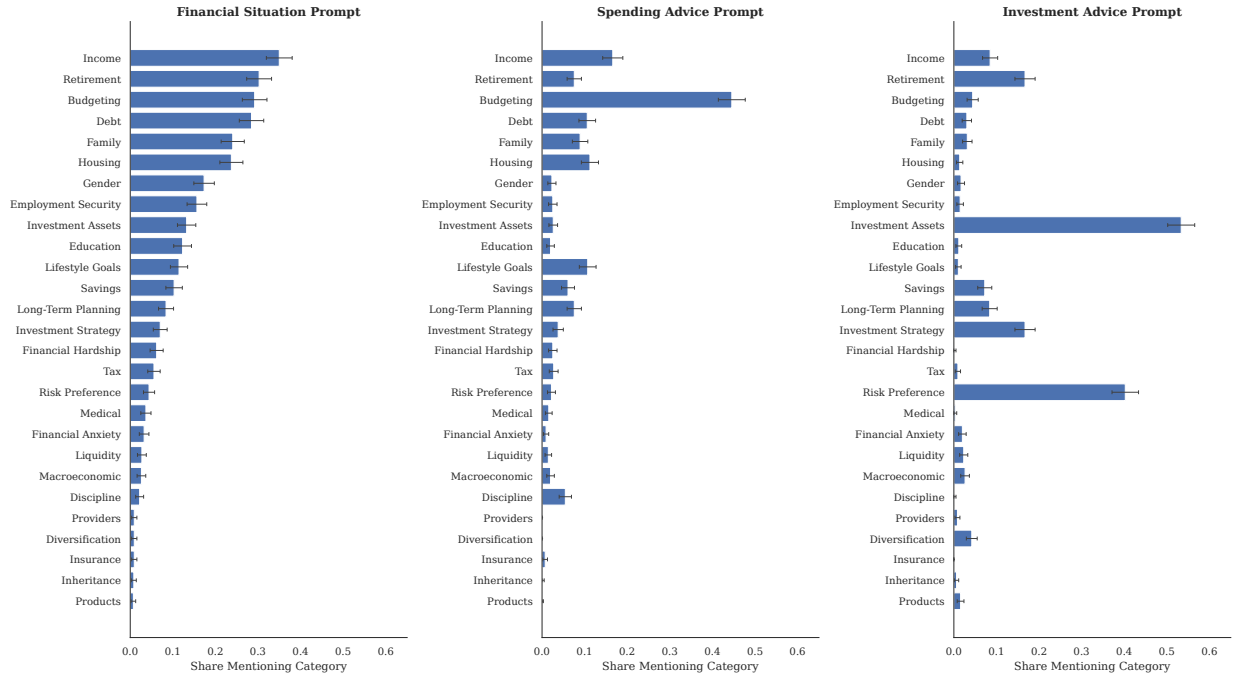
Figure A5. Word Clouds by Individual Prompt

Financial Situation | **Spending Advice** | **Investment Advice**



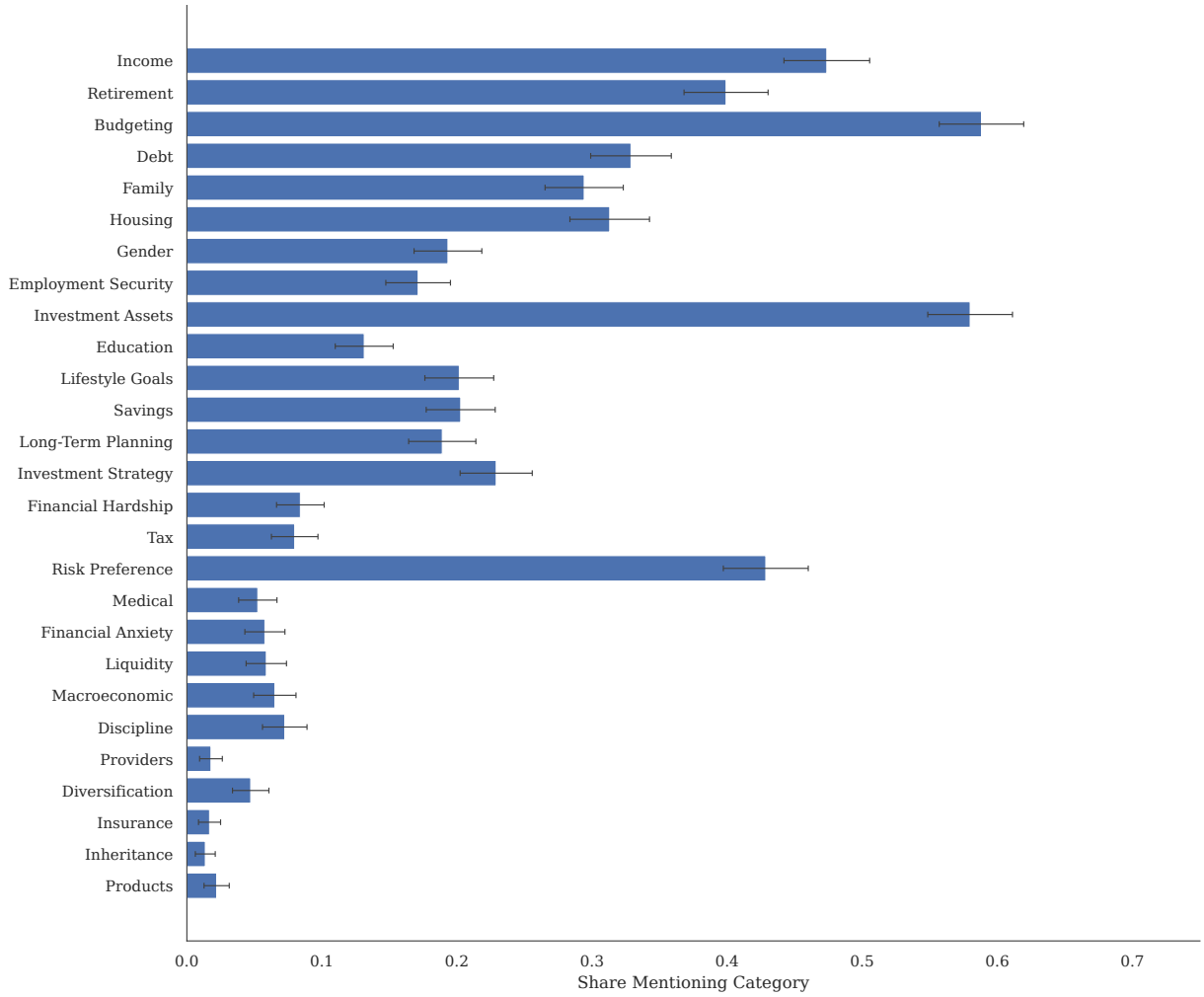
Notes: Each panel shows a word cloud for one of the three respondent prompts. Word size is proportional to raw frequency. Stopwords—including scikit-learn's default English stopwords, custom generic terms, and survey-instrument words—and prompt-specific tautological terms (e.g., "spend" and "invest" in the investment advice prompt) are excluded from each panel. Stopwords are removed only for word-cloud display, not for dictionary matching. Text is preprocessed using spaCy tokenization, lemmatization, and domain-specific root merging. Figure 2 shows the combined prompt word cloud alongside the advice word cloud.

Figure A6. All Dictionary Categories by Prompt



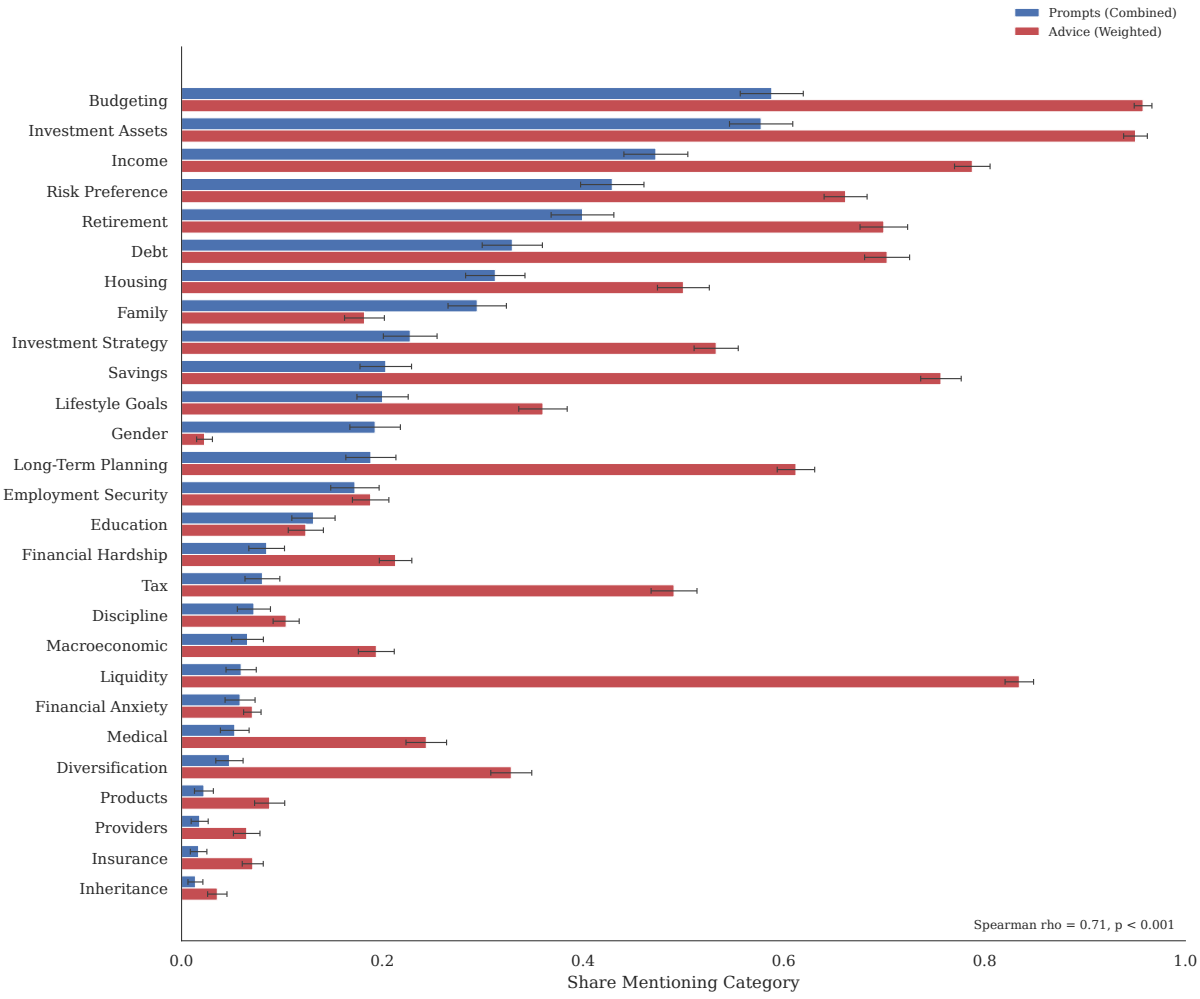
Notes: Each panel shows the prevalence of all 27 dictionary categories for one of the three prompts (952 respondents). Bars report the share of respondents whose prompt text contains at least one keyword from the category. Error bars show 95% confidence intervals based on the normal approximation for binomial proportions. Categories are ordered by descending prevalence in the financial situation prompt. All panels share a common x-axis upper bound.

Figure A7. Combined Dictionary Category Prevalence Across All Prompts



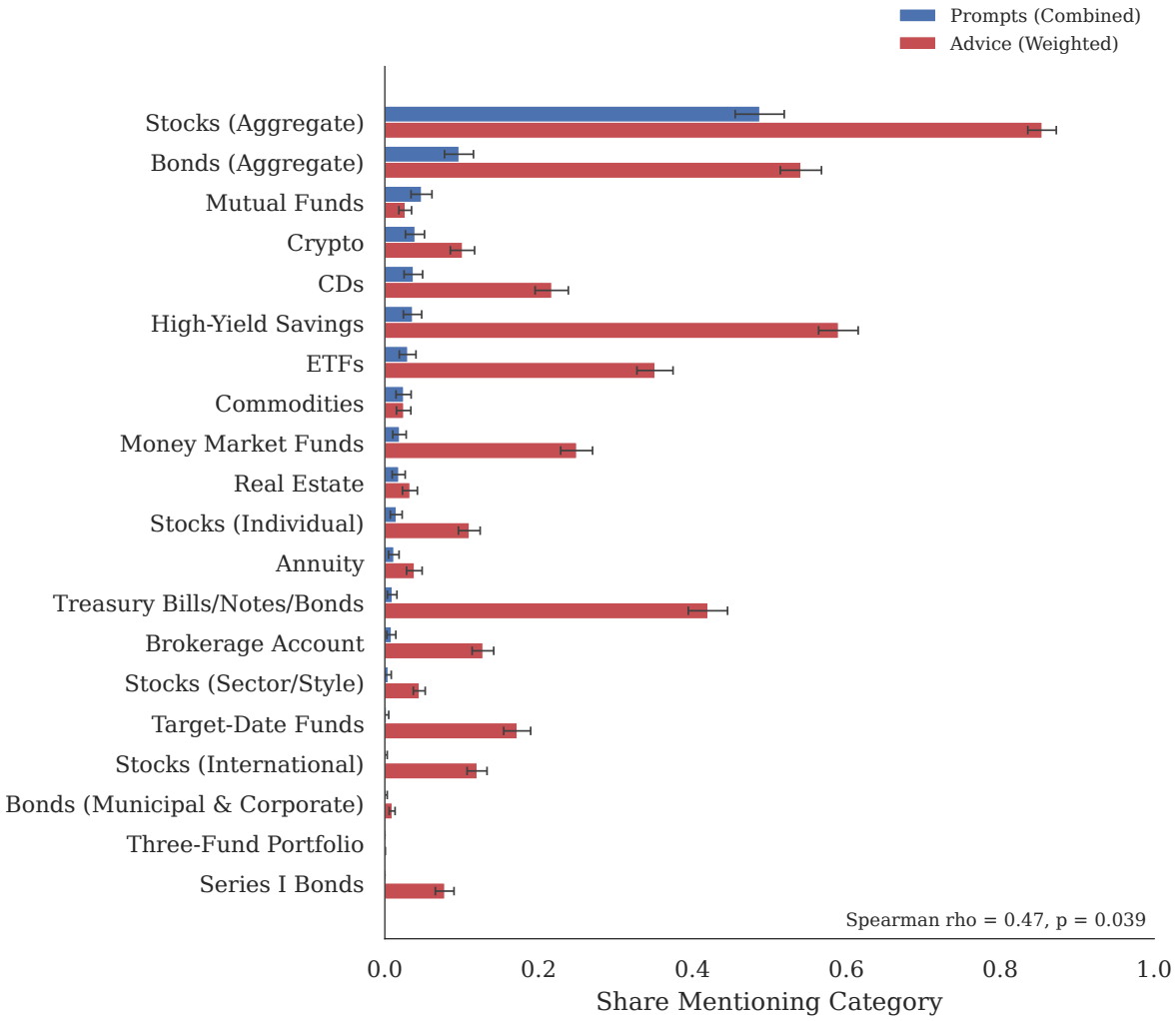
Notes: Combined respondent-level prevalence for all 27 dictionary categories (952 respondents). Each bar reports the share of respondents whose text contains at least one keyword from the category in any of the three prompts. Error bars show 95% confidence intervals based on the normal approximation for binomial proportions. Categories are ordered by descending prevalence.

Figure A8. All Dictionary Categories: Prompts vs. Advice



Notes: Grouped horizontal bar chart comparing prompt and advice prevalence across all 27 dictionary categories (n = 944 matched respondent-advice pairs). Blue bars show the share of respondents mentioning the category in any of the three prompts. Red bars show the respondent-weighted mean mention rate in AI-generated advice. Categories are ordered by descending combined-prompt prevalence. Error bars show 95% confidence intervals. Spearman rank correlation: $\rho = 0.71$, $p < 0.001$.

Figure A9. All Investment Asset Groups: Prompts vs. Advice



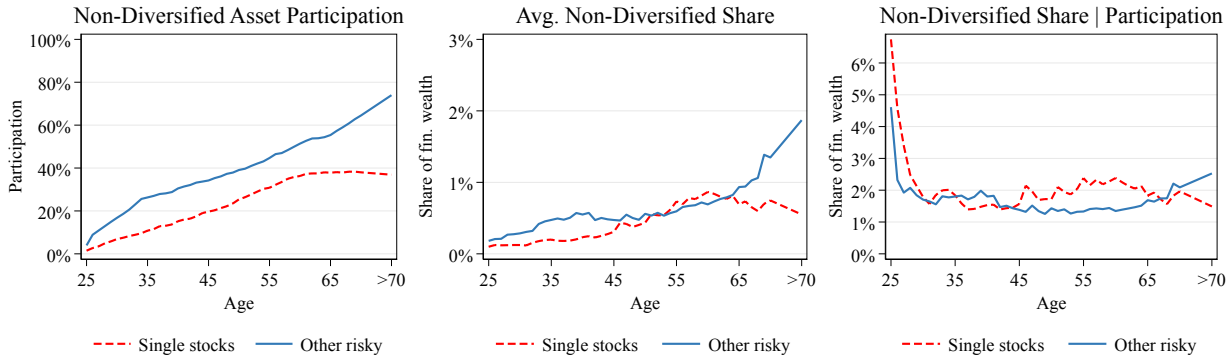
Notes: Grouped bar chart comparing prompt and advice prevalence for all investment asset keyword groups (n = 944 matched pairs). Blue bars show the share of respondents mentioning the asset type in any of the three prompts. Red bars show the respondent-weighted mean mention rate in AI-generated advice. Categories are ordered by descending advice prevalence. Error bars show 95% confidence intervals. The main text (Figure 4, Panel A) shows the subset of categories with at least 5% prevalence in either series.

Table A5. 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 ($)	Employed	46,755	245,943	636,464	1,273,949	1,879,185	674,186
	Unemployed	29,883	178,785	526,523	1,118,812	1,654,561	921,878
Stock Participation (%)	Employed	100%	100%	100%	100%	100%	100%
	Unemployed	93%	100%	100%	100%	100%	99%
Equity Share (%)	Employed	79%	80%	71%	62%	55%	71%
	Unemployed	75%	84%	78%	68%	61%	70%
Net Saving Rate	Employed	27%	21%	12%	-8%	-52%	7%
(% of earnings)	Unemployed	-4%	-41%	-90%	-193%	-307%	-167%
Panel C: Life Cycle Model (LCM) Endogenous Choices							
Consumption ($)	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	Employed	45%	27%	-3%	-56%	-118%	-8%
(% of earnings)	Unemployed	-16%	-55%	-133%	-279%	-451%	-245%

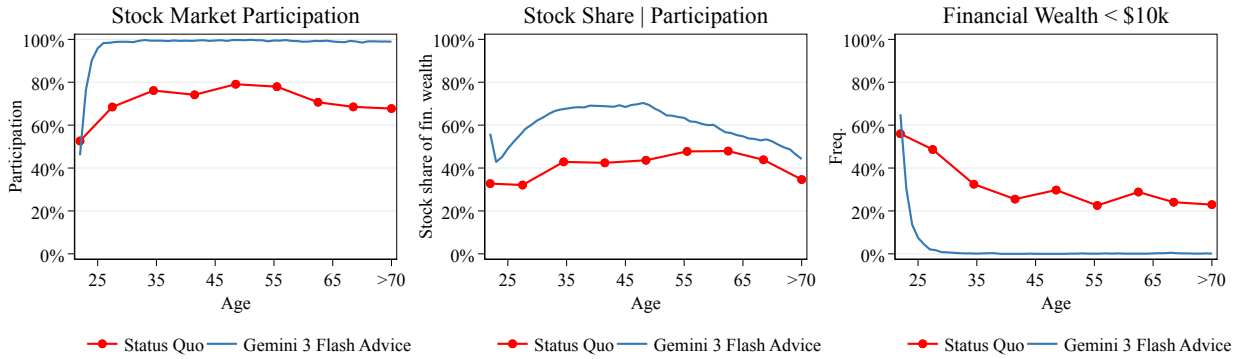
Notes: This table reports mean values for key variables from both the LLM and life cycle model simulations, broken down by age group (22-29, 30-39, 40-49, 50-59, 60-64). The Total column shows averages over the entire working life (ages 22-64). Panel A presents exogenous inputs that are identical across 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.

Figure A10. Non-Diversified Asset Participation and Portfolio Shares



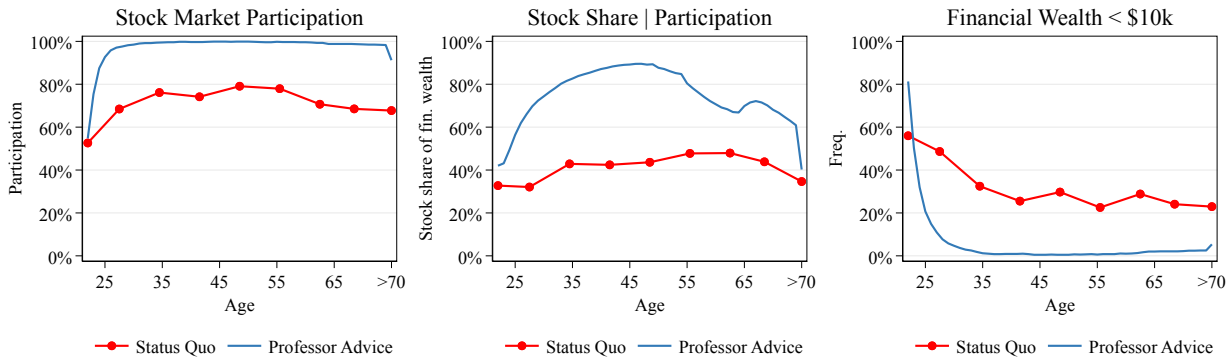
Notes: This figure plots the life cycle profiles of non-diversified asset holdings—individual stocks and other risky assets (e.g., crypto, gold, commodities, collectibles)—for individuals following the LLM's advice using the survey prompts, as described in Section 1.5. The left panel shows unconditional participation rates by age; the middle panel shows the average portfolio share of financial wealth allocated to each asset class; and the right panel shows the average portfolio share conditional on participation. Ages (22-89) are grouped into 20 equally sized bins.

Figure A11. Current Behavior vs. LLM-Recommended Behavior: `Gemini 3 Flash`



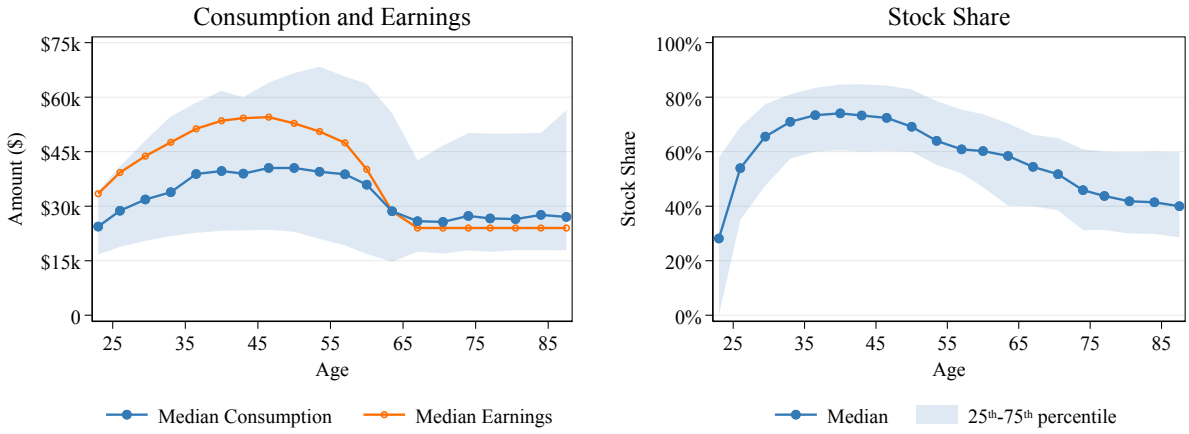
Notes: This figure compares survey respondents' current financial behavior with the LLM's recommendations using `Gemini 3 Flash`. Current behavior is reported by individuals in our Prolific survey. LLM recommendations are generated using survey prompts as described in Section 1.5. Ages (22-89) are grouped into 20 equally sized bins. All values are in 2025 dollars.

Figure A12. Current Behavior vs. LLM-Recommended Behavior: Academic Prompt



Notes: This figure compares survey respondents' current financial behavior with the LLM's recommendations using the academic prompt described in Section 1.5. Current behavior is reported by individuals in our Prolific survey. Ages (22-89) are grouped into 20 equally sized bins. All values are in 2025 dollars.

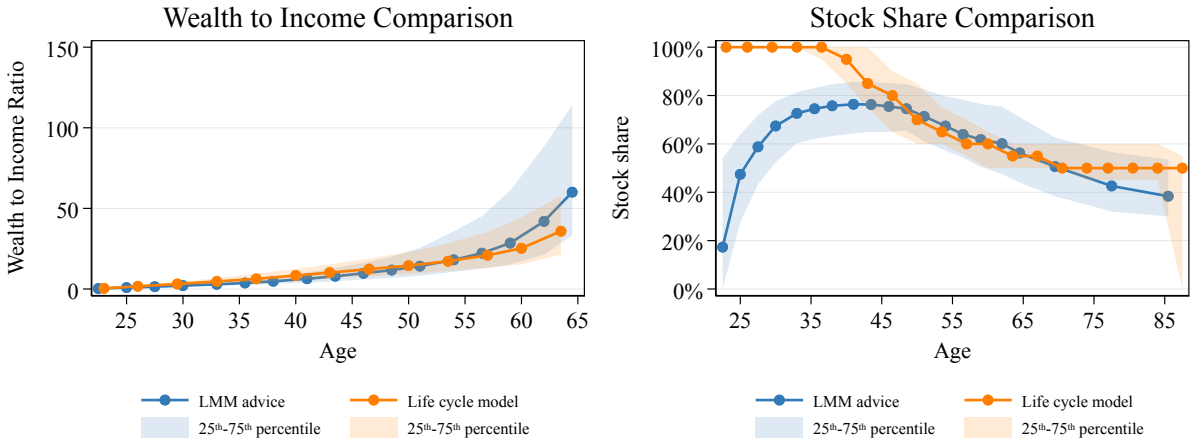
Figure A13. Life Cycle Profiles of LLM-Recommended Consumption, Wealth, and Stock Shares: `Gemini 3 Flash`



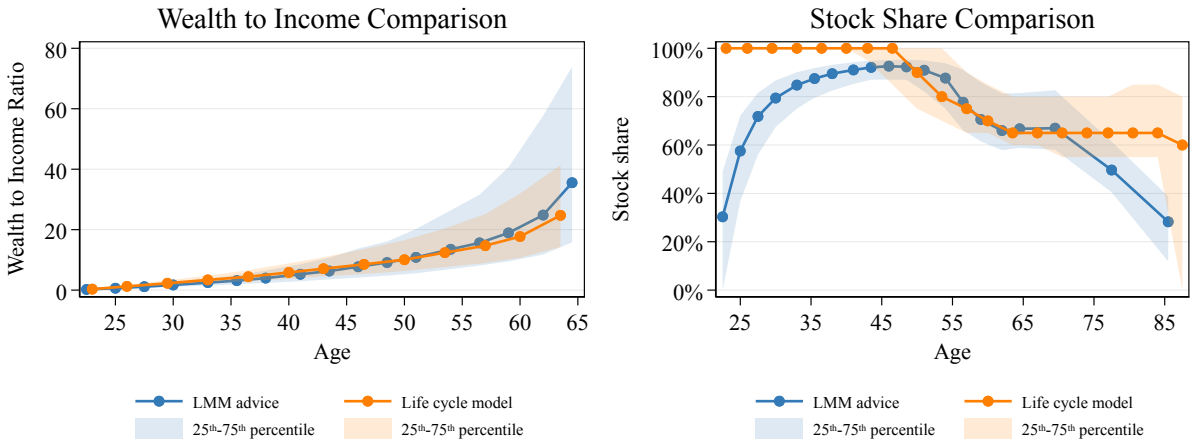
Notes: This figure plots the life cycle profiles for individuals following the LLM's recommended choices using `Gemini 3 Flash` with the survey prompts, as described in Section 1.5. Ages (22-89) are grouped into 20 equally sized bins. The left panel shows consumption and post-tax salary; the middle panel 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). The estimated preference parameters are $\hat{\beta} = 1.054$ and $\hat{\gamma} = 5.1$. All values are in 2025 dollars.

Figure A14. Life Cycle Model Fit to LLM Advice Moments

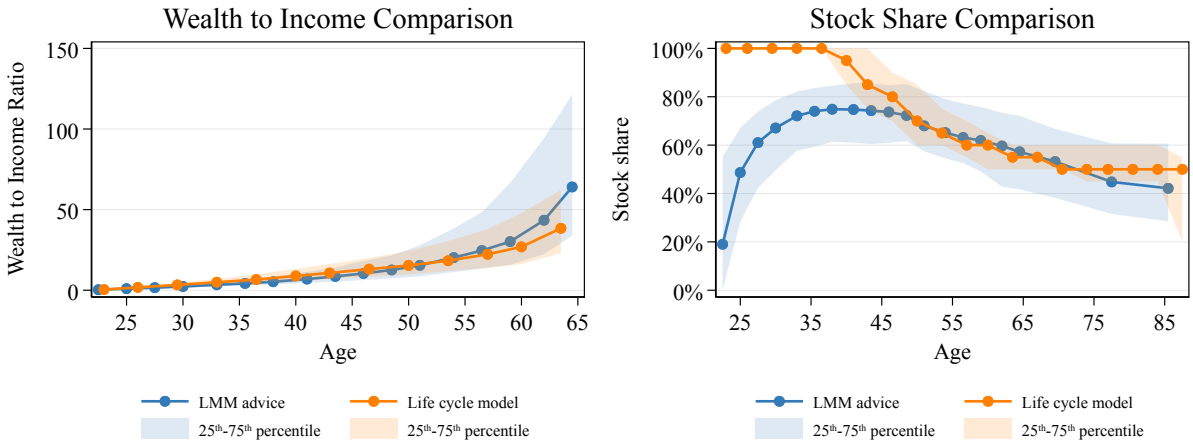
Panel A: Survey Prompts (`GPT-5.2`)



Panel B: Academic Prompts (`GPT-5.2`)

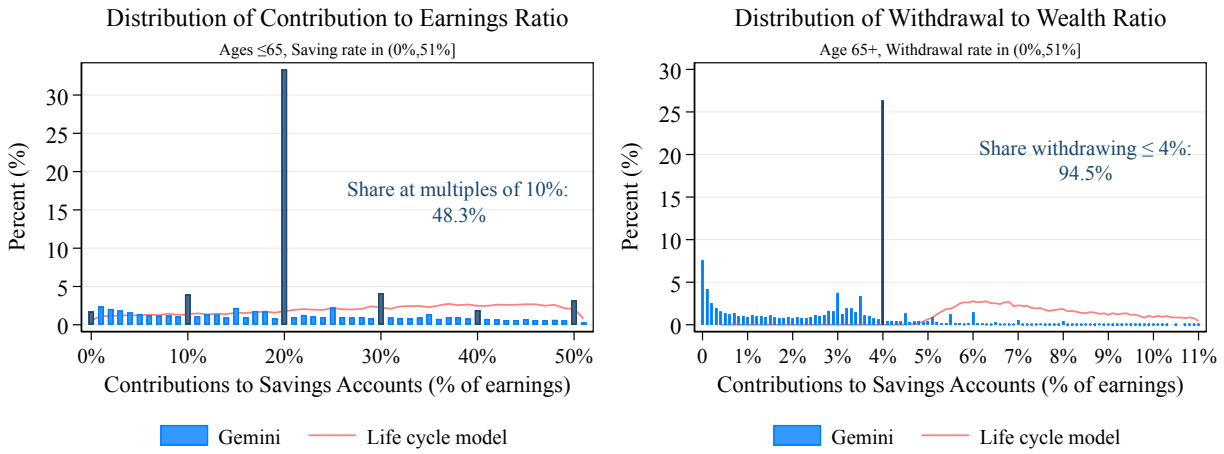


Panel C: Survey Prompts (`Gemini 3 Flash`)



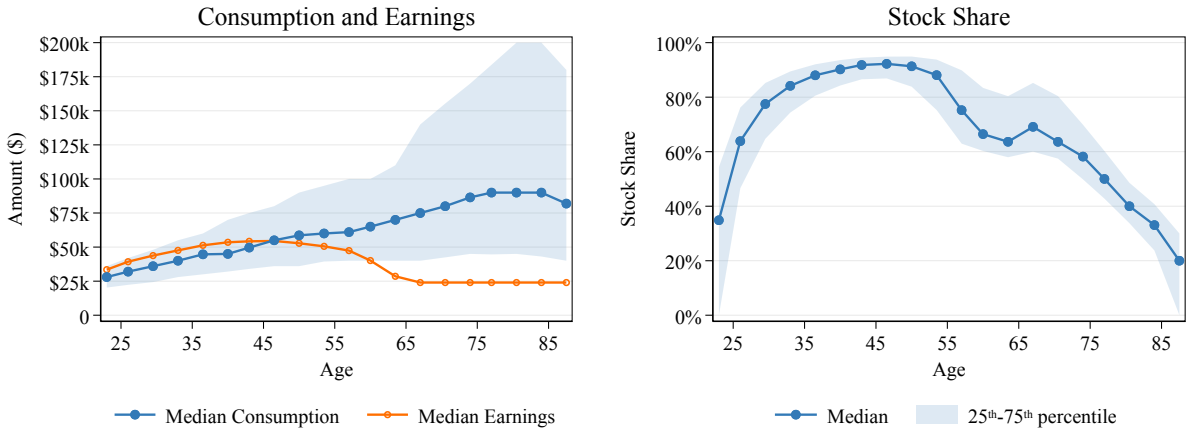
Notes: This figure shows the fit of the estimated life cycle model (with the SMM-estimated $\hat{\beta}$ and $\hat{\gamma}$ from Table 2) to the LLM advice moments. Panel A corresponds to survey prompts with `GPT-5.2` ($\hat{\beta}$ = 1.044, $\hat{\gamma}$ = 5.2), Panel B to academic prompts with `GPT-5.2` ($\hat{\beta}$ = 0.99, $\hat{\gamma}$ = 4.7), and Panel C to survey prompts with `Gemini 3 Flash` ($\hat{\beta}$ = 1.054, $\hat{\gamma}$ = 5.1). Each panel plots the targeted moments (wealth-to-income ratio in working life and equity share in the full life-cycle by age) for both the LLM advice and the estimated model. $D_{\varepsilon} w^{*}(\varepsilon)$

Figure A15. Saving and Withdrawal Heuristics in LLM Advice: `Gemini 3 Flash`



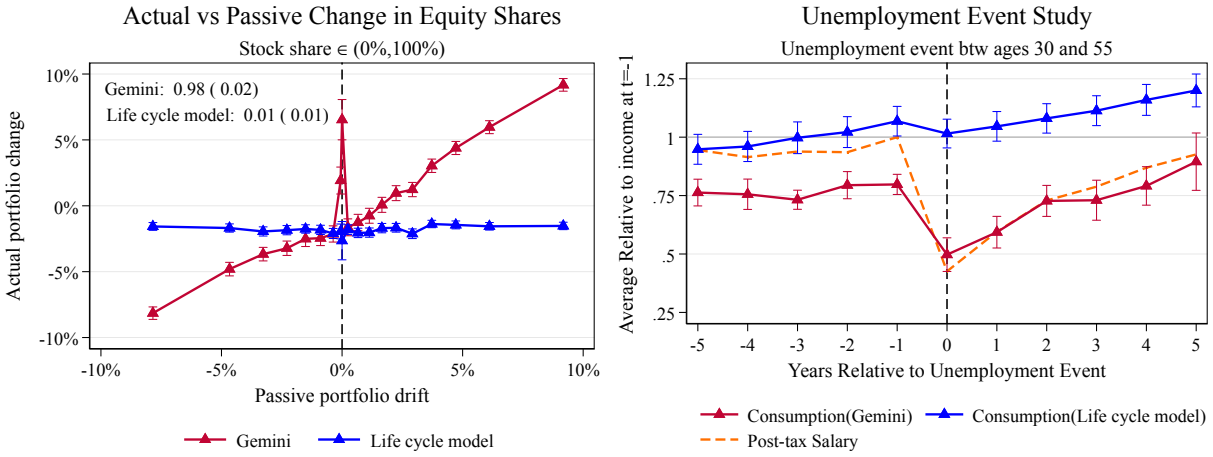
Notes: This figure plots the distributions of savings rates, savings amounts, and withdrawal heuristics in both the LLM advice and the life cycle model. The LLM advice is generated using `Gemini 3 Flash` with the survey prompts as described in Section 1.5. Blue histograms represent the LLM advice, and dark blue bars highlight heuristic values. The red lines represent the corresponding distributions from the life cycle model described in Section 1.1. All values are in 2025 dollars.

Figure A16. Life Cycle Profiles of LLM-Recommended Consumption, Wealth, and Stock Shares: Academic Prompt



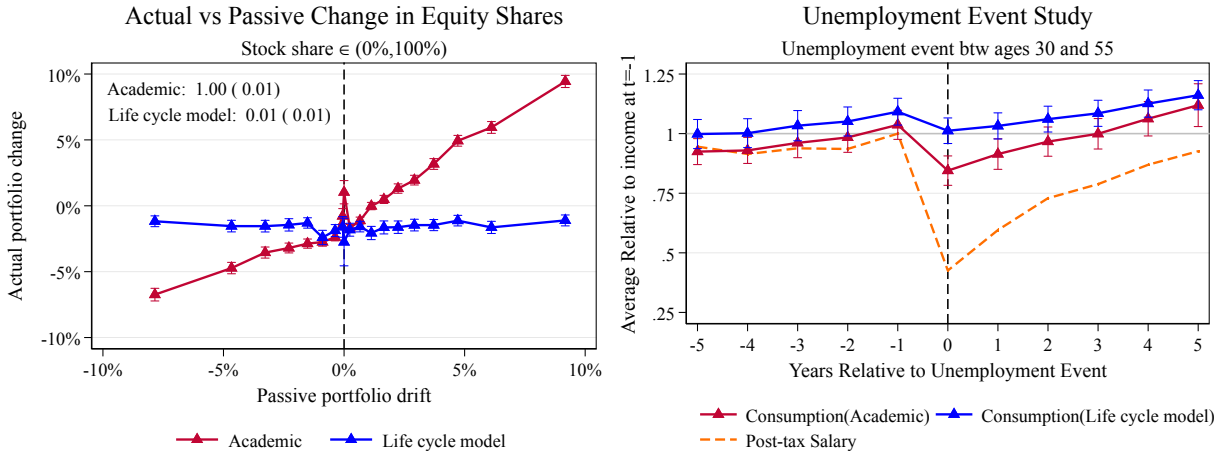
Notes: This figure plots the life cycle profiles for individuals following the LLM's recommended choices using the academic prompt, as described in Section 1.5. Ages (22-89) are grouped into 20 equally sized bins. The left panel shows consumption and post-tax salary; the middle panel shows wealth accumulation; 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). The estimated preference parameters are $\hat{\beta} = 0.99$ and $\hat{\gamma} = 4.7$. All values are in 2025 dollars.

Figure A17. Responsiveness of LLM Advice to Shocks: `Gemini 3 Flash`



Notes: This figure compares the responsiveness to shocks between the LLM advice and the life cycle model using `Gemini 3 Flash`. 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 1.1. The left panel presents a binscatter plot of average actual portfolio change against passive portfolio drift. Passive portfolio drifts are grouped into 20 equally sized bins. The right panel plots the unemployment event study, constructed in the same way as the right panel of Figure 9.

Figure A18. Responsiveness of LLM Advice to Shocks: Academic Prompt



Notes: This figure compares the responsiveness to shocks between the LLM advice and the life cycle model using the academic prompt described in Section 1.5. 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 1.1. The left panel presents a binscatter plot of average actual portfolio change against passive portfolio drift. Passive portfolio drifts are grouped into 20 equally sized bins. The right panel plots the unemployment event study, constructed in the same way as the right panel of Figure 9.

Figure A19. Example Prompts by Financial Literacy

High Financial Literacy User:

I am a short term futures trader usually executing **3 to 4 trades per daily session**. I do not hold any positions after the market closes. I **trade the futures micro contracts of indexes, crude oil, gold, currencies and bonds**. I also place option trades on the SPX. I am looking to imptove my strategy which is based on Cumalative Delta and Orderflow. My account size is 10,000 USD

I'd like to know how much to spend over the coming year.

I'd like to know how to invest to improve my strategy and reduce my risk exposure.

Low Financial Literacy User:

Hello….Im in need of some financial advice. I need to invest some money into something that will make me money in the future. […] I have current stocks invested, however, **I know very little about stocks** and how and when to buy/sell, and **what stocks I need to buy and/or avoid**? Some advice would be very helpful! Thank you! :)

Notes: This figure shows example prompts written by individuals with high and low financial literacy, as measured by a Big 5 financial literacy test administered in our Prolific survey.

Figure A20. Differences in Prompts and Advice by Financial Literacy

Panel A: Words Overrepresented in Prompts by Financial Literacy

High Financial Literacy | **Low Financial Literacy**



Panel B: Words Overrepresented in Advice by Financial Literacy

High Financial Literacy | **Low Financial Literacy**



Notes: This figure shows word clouds of terms that are differentially used by individuals with high versus low financial literacy, as measured by a Big 5 financial literacy test. Panel A shows words overrepresented in the prompts written by individuals with high financial literacy (left) versus low financial literacy (right). Panel B shows words overrepresented in the LLM's advice to individuals with high (left) versus low (right) financial literacy. Word size is proportional to the difference in frequency between groups.

Figure A21. Example Prompts by Prior AI Experience

AI User:

Assume you are a CFA with 20+ years of experience. I am needing advice to help create a sustainable investment portfolio using E*trade or similar self-select investment platforms. I have $5K to invest currently. I have another $10k in savings and am already investing in an employer sponsored retirement account contributing the max allowing amount.

Using your CPA knowledge, and our investment strategy determined previously, how much would I bel able to continually spend to add to my investment portfolio?

What additional investment vehicles should I be looking add for my further spending? Should reinvest in high return vehicles or expand and diversify?

AI Non-User:

I would like the AI tool to create a budget for my new pay rate as I am starting a new job. I will be making $20 an hour and working 40-45 hours a week. I pay $1000 in rent & around $500 in utilities/ household need bills.

How much could I spend freely with this salary while maintaining funding for rent, household needs, & $200 savings a month?

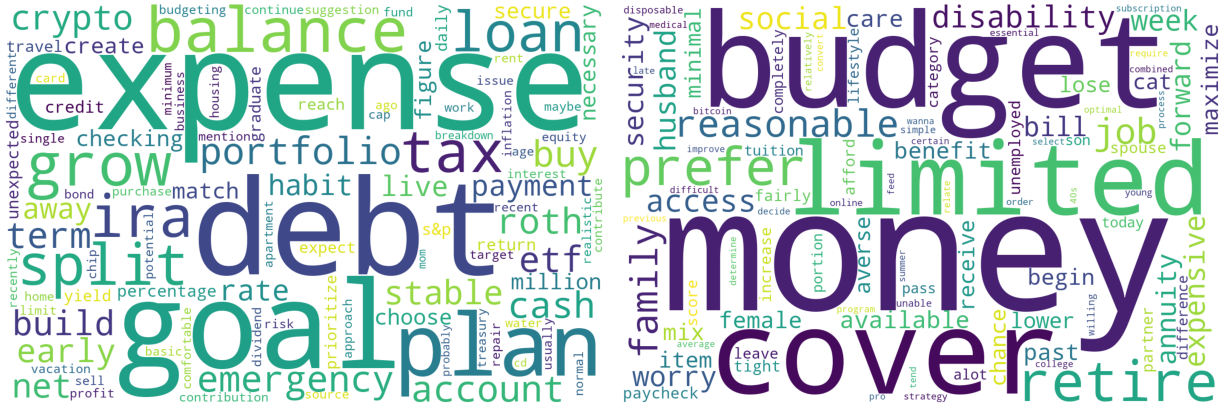
Where should I invest starting with $50 & consistently adding $25 a month after?

Notes: This figure shows example prompts written by individuals who have and have not previously used AI for financial advice, as reported in our Prolific survey.

Figure A22. Differences in Prompts and Advice by Prior AI Experience

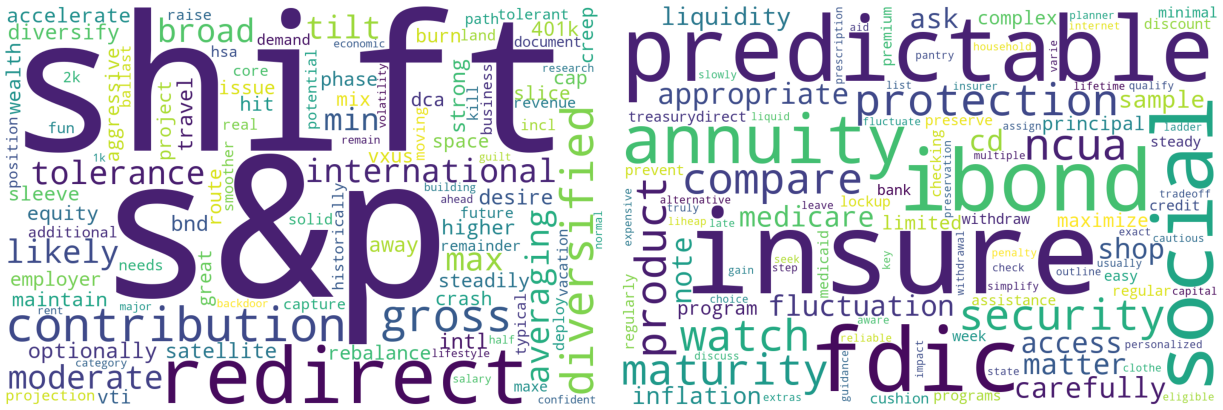
Panel A: Words Overrepresented in Prompts by Prior AI Experience

Prior AI Use | **No Prior AI Use**



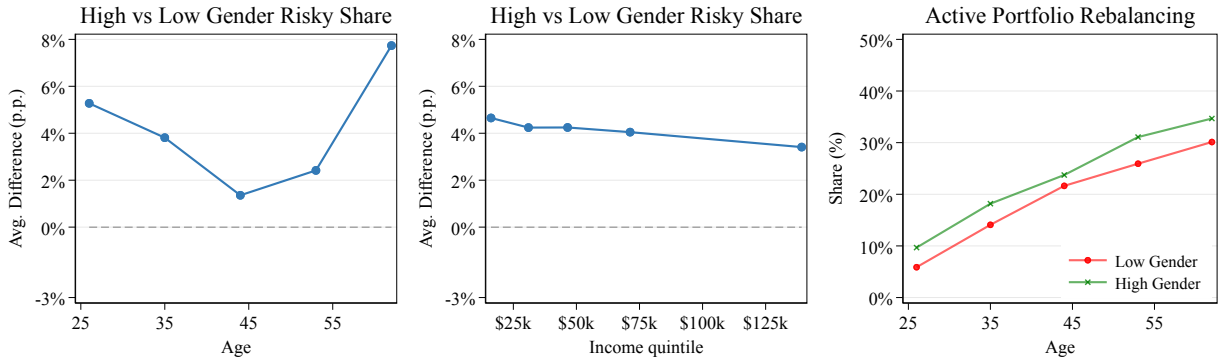
Panel B: Words Overrepresented in Advice by Prior AI Experience

Prior AI Use | **No Prior AI Use**



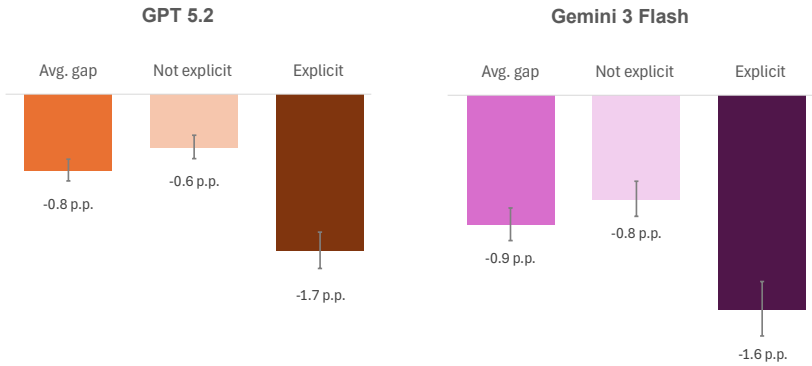
Notes: This figure shows word clouds of terms that are differentially used by individuals who have versus have not previously used AI for financial advice. Panel A shows words overrepresented in the prompts written by individuals with prior AI experience (left) versus without (right). Panel B shows words overrepresented in the LLM's advice to individuals with prior AI experience (left) versus without (right). Word size is proportional to the difference in frequency between groups.

Figure A23. Gender Gap in Recommended Equity Shares and Portfolio Rebalancing



Notes: This figure examines the gender gap in LLM-recommended equity shares by age and income. The left panel plots the average difference in recommended risky asset shares (men minus women) by age. The middle panel plots the same difference by income quintile. The right panel plots the share of individuals whose LLM-recommended portfolio change constitutes active rebalancing (rather than passive drift), separately for men and women by age.

Figure A24. Isolating the Demand Channel: Explicit vs. Implicit Gender



Notes: This figure isolates the demand channel of the gender gap in LLM-recommended equity shares by comparing the effect of explicitly stating gender in the prompt versus women-authored prompts that do not mention gender. The left panel shows results for `GPT-5.2` and the right panel shows results for `Gemini 3 Flash`. Within each panel, the first bar shows the average gender gap across all prompts, the second bar restricts to prompts that do not explicitly mention gender, and the third bar restricts to prompts that explicitly mention gender. Error bars show 95% confidence intervals.