Generative AI and Retail Investment*†

Valerie Baldinger‡

November 8, 2026

Updated regularly. Click here for latest version.

Abstract

Generative AI is a general purpose technology with rapidly growing adoption among retail investors. How does the adoption of this technology change household investment decisions? I study the adoption of a generative AI–based conversational assistant by retail investors, using data that link investors' prompts and AI answers to transaction-level trading records and portfolio holdings. Using a staggered difference-in-differences design, I find that generative AI adoption increases investment in index funds relative to individual stocks by 7.6 percentage points. As a result, AI adoption reduces investors' return loss, which measures the expected return forgone due to underdiversification at a given level of risk, by 43 basis points on average, with substantial heterogeneity across investors. Less diversified portfolios improve the most, with return loss declining by 101 basis points.

*I am deeply grateful to Arpit Gupta (committee chair), Toomas Laarits, Thomas Philippon, Indira Puri, and Olivier Wang for their invaluable guidance, encouragement, and support. I also thank Viral Acharya, Yakov Amihud, John Bonney, Amanda Dos Santos, Ian D'Souza, Shan Ge, Daniel Greenwald, Peter Koudijs, Theresa Kuchler, Daniel Milo, Alan Moreira, Holger Mueller, Sungho Park, Dean Parker, Alexi Savov, Philipp Schnabl, Johannes Stroebel, Sean (Lingxuan) Wu, and Xuran Zeng, as well as seminar participants at NYU Stern, for helpful comments and suggestions.

†I thank Mezzi.com for providing access to the de-identified data used in this paper. I served as an unpaid contractor for Mezzi and have no financial interest in the company.

‡New York University, Stern School of Business. Email: vlb7599@stern.nyu.edu.

1 Introduction

Generative AI has become one of the most rapidly adopted technologies, outpacing the adoption speed of the personal computer and the internet (Bick et al., 2026; Lynch, 2026). This adoption extends to retail investors: recent surveys report that 54–70 percent of investors have used AI to obtain investment information or support financial decisions (Investing.com, 2026; Amundi, 2026; HSBC and Ipsos, 2026; Opinium, 2026).

How do retail investors use generative AI, and how does the adoption of this technology change their investment behavior? On the one hand, generative AI may improve investment choices by providing sound information and advice consistent with portfolio theory. Such guidance could address a central issue in household finance: under-diversification and its costs for investors (Calvet, Campbell, and Sodini, 2007). On the other hand, generative AI may reduce the cost of researching individual companies or increase investors' confidence in their ability to select stocks, leading to more individual stock trading and possibly decreased diversification.[1] A third possibility is that AI adoption leaves investment behavior unchanged: sycophantic responses may simply affirm investors' existing beliefs and investment decisions.

To study how retail investors use generative AI and how adoption affects their investment behavior, I use data from a wealth management platform serving U.S. retail investors that introduced an assistant based on ChatGPT and Gemini in November 2024. These data link investors' prompts and generative AI answers to their subsequent transactions and portfolio holdings. To my knowledge, this is the first study to examine transaction- and portfolio-level evidence on the effects of generative AI on retail investment. Using an LLM to encode investors' prompts and AI answers, I find that investors use generative AI to seek information and advice about securities and investment products, portfolio allocation and risk, existing holdings and performance, retirement and financial planning, and taxes. Generative AI answers to investor questions strongly favor passive index fund investing over individual stocks, a pattern that persists over time and across ChatGPT and Gemini models. The linked data allow me to match AI recommendations to investors' subsequent transactions, and I find high rates of implementation of AI recommendations.

Identifying the causal impact of generative AI adoption on investment behavior is challenging because investors choose whether and when to use generative AI, as well as what to ask about. I address this identification challenge using two complementary identification strategies. First, I study the staggered adoption of generative AI by retail investors and its effects on their investment decisions and portfolio allocations, following the introduction of the generative AI financial assistant on the platform in November 2024. I find that generative AI adoption increases investment in index funds relative to individual stocks by 7.6 percentage points. As a result, AI adoption reduces investors' return loss, which measures the expected return

[1] This view is consistent with Van Nieuwerburgh and Veldkamp (2010), who show in a model that information acquisition can make a portfolio consisting of individual stocks and a diversified fund optimal.

forgone due to underdiversification at a given level of risk, by 43 basis points on average, with substantial heterogeneity across investors. Less diversified portfolios improve the most, with return loss declining by 101 basis points. Second, I exploit changes in the instructions given to the AI assistant that altered how often it provided explicit buy recommendations. Since these recommendations predominantly concern funds, the changes provide a source of exogenous variation in fund advice received by existing AI users. I use a difference-in-differences design and find that investors allocate a larger share of their purchases to funds when AI recommendations to buy funds become more frequent, consistent with the main staggered DiD results.

I have four main findings.

First, I find that generative AI answers guide investors toward passive index investing. I use an LLM to encode investors' prompts and AI answers and examine how the advice varies with the topic and nature of investors' questions. A simple descriptive comparison that illustrates this finding is the difference between investors' prompts and AI answers: investors express interest in buying ETFs or mutual funds in 3.5 percent of prompts, while 18 percent of AI answers recommend buying funds. Moreover, passive funds are recommended more frequently than individual stocks across financial topics and in responses to both information requests and explicit requests for advice. This result remains after accounting for investor fixed effects, indicating that it is not explained by persistent differences in investor characteristics. In addition, the guidance toward passive investing persists over time and across ChatGPT and Gemini models.

Second, investors shift toward passive funds following AI adoption, in line with the advice they receive. Using a staggered difference-in-differences design, I find that adoption increases the passive fund share of new investment by 7.60 percentage points and the passive fund share of month-end equity holdings by 2.73 percentage points. I find no evidence that this shift is offset by riskier individual stock purchases; the idiosyncratic risk of new purchases is lower following adoption. Purchased funds also have lower expense ratios.

A complementary design provides further evidence that AI advice shifts investment toward funds. I exploit changes in the assistant's instructions that increased the frequency of ticker-specific fund buy recommendations. Using a difference-in-differences design, I find that the fund share of new investments increases among exposed investors relative to nonusers when ticker-specific buy recommendations become more frequent.

Taken together, the two designs address different identification concerns while implying similar investment responses to fund recommendations, providing causal evidence that AI advice shifts retail investment toward funds. Changes in the assistant's instructions provide exogenous variation in the advice received by existing AI users, addressing the concern that investors adopt AI when they already intend to change their investments. The staggered adoption design estimates effects across different calendar months and market conditions, reducing concerns that the results are driven by contemporaneous shocks around the AI assistant's instruction

changes. The magnitudes are also consistent: after scaling by the corresponding increase in exposure to fund recommendations, both designs imply a similar increase in the share of new investment allocated to funds in response to AI advice.

I then examine the effects of AI adoption on portfolio diversification. I find lower portfolio volatility and a smaller idiosyncratic share of portfolio risk. Return loss, defined by Calvet, Campbell, and Sodini (2007) as the expected return investors forgo by holding their portfolio instead of an efficient portfolio with the same level of risk, falls by 43 basis points. This improvement is economically meaningful: the average reduction in return loss is comparable to the 44 basis point difference in return loss between investors with low and high financial numeracy documented by von Gaudecker (2015). There is substantial heterogeneity underlying these results, with investors who were less diversified before adoption experiencing greater diversification benefits. For the bottom half of investors in terms of diversification before adoption, annual return loss falls by 101 basis points, or 25 percent of their pre-AI mean. Overall, these findings suggest that generative AI may bring investment choices more in line with the advice from portfolio theory by improving diversification and thereby reducing investors' exposure to uncompensated risk. In Section 7, I then compare these effects with those of other technological innovations studied in the literature and find that the improvements in fund allocation and portfolio efficiency lie between those reported for broadband internet and robo advice (Hvide et al., 2024; Rossi and Utkus, 2024).

Third, I provide evidence on personalization as a mechanism through which generative AI affects investment behavior. Personalization is a key attraction of generative AI chatbots and has been explored in the nudge literature as a way of altering behavior (DellaVigna and Linos, 2022). I find that investors are more likely to implement recommendations when answers use more personal details. To address the concern that investors' questions and investment intentions may jointly affect personalization and implementation, I construct an instrument based on the model generating each answer. This leverages two key features of my setting. First, most personalization concerns portfolio information supplied directly by the platform, giving models access to investors' circumstances without requiring investors to disclose this information in their questions. Second, improvements across model generations in processing long contexts allow models to make greater use of this information. I therefore use model indicators as instruments to exploit differences in how models use platform-supplied portfolio information and estimate the effect of portfolio personalization on implementation. I find that a one-standard-deviation increase in portfolio personalization raises implementation by 11.9 percentage points. Personalization relies on personal data that allow chatbots and AI agents to tailor their responses to consumers' circumstances, so this finding also has important implications for the value of data in the age of AI (see Veldkamp, 2023, for a discussion of data as an asset).

Finally, I find that ticker-level recommendations by Gemini and ChatGPT models exhibit mode collapse,

meaning that the distribution of recommended tickers is concentrated on a small set of securities. Across models, the ten most frequently recommended tickers account for 50 percent of ticker recommendations. This concentration in recommendations translates into the investments of AI adopters. Following AI adoption, the share of their fund investment going to the twenty most widely recommended funds increases by 7.19 percentage points. Given widespread adoption of AI, this concentration of retail investment may have implications for financial markets.

Overall, my findings thus provide evidence that generative AI adoption shifts investment toward passive index funds and can thereby improve diversification and portfolio efficiency for individual investors. The widespread adoption of generative AI by consumers may therefore further increase passive index investing, with potential implications for asset prices (Haddad and Muir, 2025). In addition to this shift towards index investing, AI adoption may make investment choices more similar across investors by directing retail demand towards a smaller set of tickers.

Related literature. This paper contributes to four strands of research. First, I contribute to the growing literature on generative AI and its applications. Mo and Ouyang (2025) provide a comprehensive review of this literature. Ecker et al. (2026) analyze investment related questions submitted to a Chinese generative AI platform. Studies by OpenAI and Anthropic document what people ask and the tasks for which they use generative AI (Chatterji et al., 2025; Handa et al., 2025). Choukhmane et al. (2026) ask a representative sample to write prompts seeking financial advice from LLMs and simulate the effects of following the advice. They find that following this advice would lead to choices more consistent with life-cycle theory. Ngangoue et al. (2026) find that ChatGPT's advice on major life decisions varies with socioeconomic background. Other work examines the preferences and biases of LLM models (Ross et al., 2024; Bini et al., 2026; Ouyang et al., 2025; Fedyk et al., 2026) and studies how financial advice responds to investor characteristics (Rumpf et al., 2026; Fieberg et al., 2025; Takayanagi et al., 2025; Foltyn and Olsson, 2026). Existing research on consumer use of generative AI thus studies what individuals ask and the answers they receive. My main contribution is to provide evidence on the implementation of AI advice in a high-stakes setting: individual investment decisions. To my knowledge, I provide the first evidence on how AI adoption affects retail portfolios, using data that link investors' AI conversations to subsequent transactions and portfolio holdings.

Second, I contribute to the literature on how the adoption of general-purpose technologies affects household investment decisions. Existing evidence documents mixed effects for the major technological innovations of the last thirty years: the internet and mobile phones. Studies of online trading find increased trading activity following adoption of online brokerage accounts and web-based trading in 401(k) plans, including higher trading frequency and portfolio turnover (Barber and Odean, 2002; Choi et al., 2002). Kalda et al. (2024)

find that smartphone trading encourages purchases of riskier, less diversifying assets. By contrast, Hvide et al. (2024) find that broadband internet access increases investment in equity funds and improves portfolio efficiency, with no adverse impact on trading. I extend this literature to generative AI, a general-purpose technology that can provide and explain investment recommendations, and find that AI adoption increases investment in index funds and mitigates under-diversification, with no adverse impact on trading frequency.

Third, I contribute to the literature on financial advice by examining a new form of financial advice. Existing literature focuses on human financial advisors and robo advice. The motivation behind this literature is that, in theory, financial advice might mitigate under-diversification and other mistakes observed in retail investors. However, in practice, the literature has identified several shortcomings of human advice. First, incentives may distort advice (Mullainathan et al., 2012; Chalmers and Reuter, 2020). Further, Bhattacharya et al. (2012) provide evidence that even unbiased and beneficial advice is rarely followed. They offer advice to a sample of German households and provide evidence that few households follow the provided advice. Other work examines advisor beliefs and limited tailoring (Linnainmaa et al., 2021; Reuter and Schoar, 2024), including Canadian evidence that clients' portfolios closely reflect their advisers' own asset allocations (Foerster et al., 2017). The literature on robo advice finds more promising evidence of its benefits (D'Acunto et al., 2019; Rossi and Utkus, 2024; see D'Acunto and Rossi, 2023, for a review). Generative AI advice differs from automated robo advice in implementation: individuals must choose which recommendations to follow and execute the trades themselves.

Finally, I contribute to the literature on household investment mistakes more broadly, and on what forms of treatment can improve financial decisions (see Campbell, 2006 for a review). This literature documents under-diversification (Calvet et al., 2007; von Gaudecker, 2015), excessive trading (Barber and Odean, 2000), and costly fund selection (Choi et al., 2010; Bailey et al., 2011; Bhattacharya et al., 2017). Explanations for under-diversification, and in particular the combination of single stocks with some weight in diversified funds, include probability weighting (Polkovnichenko, 2005; Dimmock et al., 2021) and incentives to specialize in learning about particular assets (Van Nieuwerburgh and Veldkamp, 2010). Other studies examine financial inertia (Madrian and Shea, 2001; Andersen et al., 2020) and the effects of financial education (Fernandes et al., 2014). I provide evidence that AI advice has the potential to mitigate under-diversification.

The remainder of the paper proceeds as follows. Section 2 describes the platform and linked data. Section 3 documents what investors ask and what the AI recommends. Section 4 studies recommendation implementation and compares recommended investments with other purchases. Section 5 presents the empirical strategy. Section 6 reports the main results and robustness checks. Section 7 assesses the welfare implications, heterogeneity in diversification gains, and comparisons with other technological innovations. Section 8 examines personalization as a mechanism through which generative AI affects investment behavior

and explores investor learning. Section 9 examines common ticker recommendations across models and the concentration of investment demand. Section 10 concludes.

2 Setting and Data

This section describes the data I analyze. I first describe the Mezzi platform and linked financial account records, and then explain how I construct portfolio and conversation measures.

2.1 Firm Background

I use data on U.S. retail investors from Mezzi, a wealth management platform that launched as a portfolio tracking application in November 2023. In November 2024, Mezzi introduced a generative AI–based conversational assistant that provides financial advice to its customers. Investors interact with generative AI through a chat interface, submitting prompts about their investments. These prompts are routed to ChatGPT and Gemini models, which generate the responses. Generative AI does not execute trades; investors decide whether and how to implement the advice they receive. I define adoption as the date of an investor's first observed prompt.

The app integrates generative AI with investors' financial account data through third-party aggregators, covering checking accounts, brokerage accounts, and tax-advantaged accounts such as 401(k)s, Roth/IRA accounts, and 529 accounts. To ensure a complete financial picture of its investors, Mezzi provides onboarding calls and assists investors with connecting brokerage, retirement, and other accounts.

2.2 Financial and Conversation Data

Through third-party aggregators, Mezzi obtains the transaction histories of investors' linked accounts.[2] Transactions and dividend receipts are recorded on a daily basis. For each transaction, I observe the date, security identifier, transaction type, quantity, price, dollar amount, and fee. The records include purchases, sales, transfers, and reinvestments and extend back up to two years before investors first connect their accounts. The classified conversational sample contains 69,598 initial and follow-up prompts and their corresponding AI answers, with a median of 12 prompts per investor.

Information such as prices and returns comes from CRSP and CRSP Mutual Funds.[3] I match the securities in the account records to the CRSP stock and mutual fund databases. The account data, in conjunction with the

[2] Before being made available for this research, the records are stripped of personally identifiable data. All analyses are conducted on these deidentified records, and only aggregate results are reported.

[3] Security identifiers are supplemented using the CRSP–Compustat linking table; option prices are from HistoricalOptionData.com.

CRSP market data, allow me to compute daily portfolio holdings, values, and returns. I first infer daily holdings from transactions, transfers, reinvestments, and dividend receipts. When a position predates the available transaction history, I recover the number of units held by dividing the cash dividend or fund distribution recorded in the account by the corresponding CRSP distribution per share. Purchases, sales, reinvestments, and split adjustments are then used to update these holdings. I value each position at the daily closing price and aggregate holdings across linked accounts. In addition to transaction and dividend data, Mezzi periodically obtains snapshots of holdings data, which are used to validate the thus constructed portfolios. The constructed daily portfolio snapshots contain market values, asset class composition, concentration measures, and returns.

I measure purchase shares relative to total purchase dollars, and the passive fund holdings share relative to month-end equity holdings. Appendix B provides the detailed variable definitions.

2.3 Summary Statistics

Table A.1 provides the key summary statistics for the main sample. Panel A reports sample size and demographic characteristics, and Panel B reports portfolio characteristics. For comparison, the table also reports the corresponding statistics for SCF respondents with equity holdings in the 2022 Survey of Consumer Finances. Table A.2 compares monthly turnover of investors in my sample with that in Giglio et al. (2021) and Barber and Odean (2000). An important distinction of my data is that they track investments across accounts at different institutions, while Giglio et al. (2021) and Barber and Odean (2000) use accounts at Vanguard and one discount brokerage, respectively. When I restrict my data to each investor's largest account to make the estimates more comparable, investors in my sample have higher mean monthly turnover than Vanguard investors (7.50 versus 2.30 percent), but turnover comparable in magnitude to the purchase turnover reported by Barber and Odean (2000) (7.50 versus 6.49 percent).

Most equity owners in my sample combine indirectly held equity with substantial investments in individual stocks. Among direct stock owners, the median investor holds 13 stocks and allocates 41.0 percent of total equity to direct stocks. The median total equity, indirect equity, and direct equity shares of the observed portfolio are 75.3 percent, 43.4 percent, and 25.3 percent, respectively. Applying the same definitions to the 2022 Survey of Consumer Finances yields corresponding medians of 35.9 percent for the direct fraction of equity, four directly held stocks, and portfolio shares of 44.0 percent, 37.8 percent, and 10.9 percent.

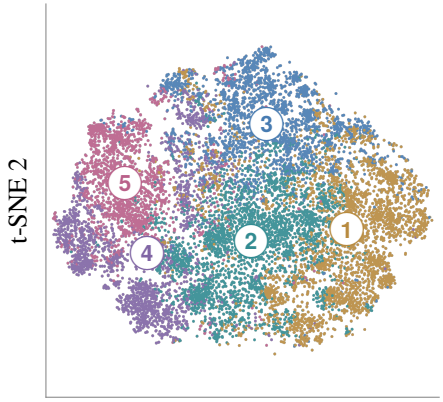
Investors in my sample lean wealthier than equity-owning SCF households (median observed portfolio value of approximately $556,000 versus $155,000) and are more active investors (75.0 percent versus 28.2 percent hold more than one individual stock). However, these characteristics are consistent with survey evidence on who uses GenAI for financial information. A survey by HSBC and Ipsos (2026) finds that 82 percent of high-net-worth investors with at least $2 million in investable assets use AI for finance and investment tasks, compared with 73 percent across its full sample of affluent and high-net-worth investors. Only 9 percent of high-net-worth respondents report that AI has not meaningfully changed how they manage their wealth, compared with 26 percent of investors with $100,000–$250,000 in investable assets (HSBC and Ipsos, 2026). Blankespoor et al. (2026) similarly find that more sophisticated retail investors lead adoption. Thus, while my sample is not representative of the overall SCF population, its characteristics are consistent with those of the retail investors most likely to use GenAI for investment decisions.

3 Investor Use of Generative AI

This section first describes the content of investors' questions and AI's corresponding responses. I then examine when investors consult AI and how consultation relates to subsequent trading.

3.1 The Content of Investor Prompts

Generative AI is a multipurpose technology that can be used for a wide array of tasks (Chatterji et al., 2025). As such, it is useful to start by classifying how AI is used by investors in this sample. I first obtain text embeddings of prompt summaries using OpenAI's large embedding model. These embeddings provide numerical representations of the prompts' meaning. I apply principal component analysis to these embeddings and use k-means clustering to group similar prompts. To visualize these embeddings in a two-dimensional space, I use the t-distributed Stochastic Neighbour Embedding (t-SNE) algorithm, which reduces the dimensionality of document embeddings (van der Maaten and Hinton, 2008). The proximity of two points in the figure can be interpreted as similarity in the meaning of the corresponding prompts. Figure 1 displays five broad groups: securities and investment products, portfolio allocation and risk, existing holdings and performance, retirement and financial planning, and taxes. These labels were generated using privacy-preserving LLM-assisted labeling and then manually reviewed.



	Financial subject	Share (%)
1	Securities and investment products	25.6
2	Portfolio allocation and risk	25.0
3	Existing holdings and performance	18.8
4	Retirement and financial planning	16.4
5	Taxes	14.2

Figure 1. Topics in Investor Prompts

Notes: This figure plots a t-SNE map of investor prompts using embeddings of prompts. Each point represents one investor-written prompt, and colors identify five k-means groups.

Semantic similarity can group prompts about similar topics but with different intent. For example, "How does an ETF work?" and "Should I buy an ETF for my retirement portfolio?" concern the same financial subject, but the first seeks information while the second seeks advice. To account for this, I next group investor prompts according to their intent into information-seeking and advice-seeking requests, again using LLM classification. Combining the prompt intent with the financial subjects identified above thus sorts investor prompts along two dimensions. The resulting distribution across these categories is displayed in Table 1. Securities and investment products account for 25.6 percent of prompts, followed by portfolio allocation and risk at 25.0 percent. Information seeking and advice seeking account for 34.3 and 59.9 percent of prompts, respectively. Panel B reports the share of investors with at least one prompt in each subject–request cell. Overall, investors use generative AI for a broad range of financial topics, seeking both advice and purely informational responses.

3.2 Investor Prompts and AI Answers

The previous section provides evidence that investors use Gen AI for a broad range of topics. In this section, I explore the investment advice contained in AI answers and how it varies with the topic and nature of investors' prompts. AI answers recommend index investing, and this pattern is present across topics and both information- and advice-seeking prompts.

Table 1. Financial Subjects and Nature of Investor Requests

Financial subject	Information seeking	Advice seeking
Panel A. Share of prompts (%)		
Securities and investment products	11.3	13.6
Portfolio allocation and risk	2.1	20.9
Existing holdings and performance	9.4	8.5
Retirement and financial planning	6.4	8.6
Taxes	5.1	8.3
Panel B. Share of investors with at least one prompt in the cell (%)		
Securities and investment products	39.9	49.1
Portfolio allocation and risk	16.9	62.1
Existing holdings and performance	42.4	40.1
Retirement and financial planning	30.2	36.1
Taxes	28.6	40.7

Notes: This table reports topic shares of plotted prompts (Panel A) and investors with at least one prompt in each cell (Panel B). The difference between the topic totals and the sum of the two columns reflects other requests not assigned to either category.

I begin by comparing buy interest in investors' prompts with buy recommendations in AI answers, and find that AI answers are highly tilted towards recommending funds and ETFs. Investors consider buying stocks in 1.9 percent of prompts, and stocks are recommended in 4.6 percent of AI answers. By comparison, investors consider buying ETFs and funds in 3.5 percent of prompts, and these are recommended in 18.0 percent of answers. AI answers are therefore about 5.14 times as likely to recommend buying a fund as investors are to express interest in buying one in their prompts. Panel C of Table 2 instead reports the share of investors receiving at least one buy recommendation for each asset type. Among investors receiving buy advice, 96.9 percent receive at least one answer containing advice to buy funds, compared with only 48.0 percent receiving at least one individual stock recommendation.

Using the classification introduced in Section 3.1, I next examine how AI's buy recommendations vary across financial topics and request types. Panel A of Appendix Table F.1 provides evidence that passive funds are recommended more frequently than individual stocks across topics for both information requests and explicit requests for advice, with likelihood ratios ranging from approximately 2.6 to 14.4. To examine whether the greater frequency of passive fund recommendations reflects persistent differences across investors, Panel B of Appendix Table F.1 additionally includes investor, calendar month, and weekday fixed effects. Passive funds remain more frequently recommended than individual stocks across topics and request types.

AI also favors funds when investors leave the choice of securities open. Panel B of Table 2 reports the composition of ticker-level buy and sell recommendations, separately for all prompts and for prompts containing neither a recognized ticker nor the words "stock" or "stocks." Among suggested purchases in AI answers to the latter prompts, 89.7 percent concern an ETF or mutual fund and 10.3 percent concern an

individual stock.

Table 2. Investor Prompts and AI Answers

	Individual stocks	ETFs and mutual funds
Panel A. Buy interest and recommendations		
Prompts expressing interest in buying	1.9	3.5
Including a specific ticker	1.9	1.7
AI answers containing buy recommendations	4.6	18.0
Including a specific ticker	4.4	15.7
Panel B. Composition of AI buy and sell recommendations		
All prompts		
Buy recommendations	15.0	85.0
Sell recommendations	50.4	49.6
Prompts without tickers or stock mentions		
Buy recommendations	10.3	89.7
Sell recommendations	47.4	52.6
Panel C. Investors receiving recommendations		
Among investors receiving buy advice	48.0	96.9
Among investors receiving sell advice	80.7	70.0

Notes: This table reports buy interest in investor prompts and buy and sell recommendations in AI answers. Panel A reports the share of prompts with buy interest and AI answers with buy recommendations for each asset type. Panel B reports, conditional on buy or sell advice, the share of ticker-level recommendations for each asset type. Panel C reports, conditional on receiving buy or sell advice, the share of investors receiving at least one recommendation for each asset type.

4 AI Recommendations and Investor Behavior: Descriptive Evidence

The data link AI recommendations to subsequent trades, which allows me to examine high-frequency, transaction-level evidence of how investors implement AI advice and how recommended investments differ from their other purchases before turning to the main estimates in Section 6. Section 4.1 first studies whether investors buy tickers (stocks and funds) that are recommended. Second, it restricts the analysis to fund ticker advice, which represents 89.7 percent of ticker-level advice, and examines whether investors buy the recommended Lipper categories and the specific recommended funds within those categories. Lipper categories group funds with similar investment objectives and strategies. Section 4.2 compares the composition, fees, and historical risk of recommended investments with those of other purchases made following the same advice. These analyses provide a first step toward understanding how AI advice may change investor behavior.

4.1 Implementation of AI Recommendations

I first present high-frequency, ticker-level evidence on the implementation of particular AI recommendations. Using an open-source LLM, I identify ticker-specific buy and sell recommendations in each response; the Online Appendix reports the classification prompt. These ticker-level recommendations are then linked to investments made within 10 trading days.

Table E.1 reports implementation rates within 1, 2, 5, and 10 trading days. 36.2 percent of answers with buy advice are implemented within 10 trading days. At the ticker-level, 18.5 percent of recommended securities are purchased. 72.6 percent of answers containing a ticker-level buy recommendation recommend more than one ticker, and thus the implementation rate at the answer level is higher.

I next examine the overall share of purchase dollars allocated to recommended securities. It combines the extensive margin (whether or not investors buy any recommended security) with the intensive margin (conditional on buying recommended securities, what share of purchase dollars is allocated to them). Purchases are pooled within 10 or 30 trading days of advice, counting each transaction once and matching it to any recommendation active on its purchase date. As can be seen in Panel A of Appendix Table E.2, the overall share is 24.65 percent within 10 trading days and 22.29 percent within 30 trading days.

As discussed in Section 3 (Table 2), 89.7 percent of ticker-level buy recommendations in response to open-ended prompts concern ETFs or mutual funds. Motivated by this prevalence of ticker-level fund recommendations, I next document how investors' fund purchases are associated with the categories and funds recommended in AI answers. I distinguish between purchasing a fund in a recommended category and selecting the particular recommended fund within that category. Purchases are measured within 10 trading days. Standard errors are double-clustered by individual and time.

Lipper category choice. I begin by considering the decision to purchase a fund in a recommended category. Specifically, I link category recommendations to subsequent purchases using a linear probability model of the form:

$$Y_{iac}^{C} = \alpha_i + \lambda_{w(a)} + \eta_{d(a)} + \beta_C R_{iac}^{C} + \gamma_C \log\left(\mathrm{AUM}_{c,m(a)-1}\right) + \varepsilon_{iac}. \tag{1}$$

The unit of observation is an advice event–category. The dependent variable is an indicator for purchasing any fund in the category. Here i denotes the investor, a the advice event (one AI answer to a user question), and c the Lipper category. The explanatory variable of interest is R_{iac}^{C}, an indicator equal to one if the answer recommends the category. The coefficient of interest, β_C, is the difference in the probability of purchasing a fund in a recommended rather than a nonrecommended category.

In Panel A of Table 3, I present evidence that investors are more likely to buy funds in the categories recommended by AI. Column (1) provides evidence that recommended categories are 23.96 percentage

points more likely to be purchased than nonrecommended categories. Column (2) adds investor fixed effects for persistent differences in purchase propensity, and calendar week and weekday fixed effects for common time variation and weekly purchase patterns. The recommendation gap changes little with these controls, at 24.27 percentage points. Column (3) corresponds to Equation (1) and adds log Lipper category AUM, measured as total fund assets in the category in the preceding month, to control for differences in category size. The recommendation coefficient is 20.26 percentage points. Doubling category AUM is associated with a 1.34-percentage-point higher purchase probability.

Fund choice within a recommended category. Next, Panel B of Table 3 compares recommended funds with other funds in the same recommended categories, distinguishing the choice of a particular fund from the choice of a category. ETF and mutual fund share classes are combined. I regress the fund purchase indicator on a dummy variable for whether the fund is recommended, controlling for investor, calendar week, and weekday fixed effects and fund assets under management (AUM):

$$Y^F_{iacf} = \alpha_i + \lambda_{w(a)} + \eta_{d(a)} + \beta_F R^F_{iacf} + \gamma \log\big(\text{AUM}_{f,m(a)-1}\big) + u_{iacf}. \tag{2}$$

The unit of observation is an advice event–category–fund. The dependent variable is an indicator for purchasing the fund. Here f denotes the underlying fund. The explanatory variable of interest is R^F_{iacf}, an indicator equal to one if the answer recommends the fund. The terms α_i, $\lambda_{w(a)}$, and $\eta_{d(a)}$ denote investor, calendar week, and weekday fixed effects. Log AUM, measured in the preceding calendar month, $m(a) - 1$, controls for differences in fund size.

The estimated probability of purchase remains higher for recommended funds within the same category after controlling for fund size and fixed effects. Recommended funds are 16.86 percentage points more likely to be purchased than other funds in the recommended categories, as reported in column (1) of Panel B of Table 3. Adding investor, calendar week, and weekday fixed effects has little effect on the estimate, which is 16.92 percentage points in column (2). Column (3) corresponds to Equation (2): adding the AUM control reduces the recommendation coefficient to 16.29 percentage points. The coefficient on log AUM implies that doubling AUM is associated with a 0.22-percentage-point higher purchase probability.

Table 3. Recommendations and subsequent fund purchases

Panel A. Across categories			
	Category purchase		
	(1)	(2)	(3)
Category recommended	23.96***	24.27***	20.26***
	(2.52)	(2.50)	(2.18)
Log Lipper category AUM			1.93***
			(0.23)
Constant	3.63***		
	(0.58)		
Investor FE	No	Yes	Yes
Week FE	No	Yes	Yes
R^2	0.03	0.07	0.10
Panel B. Within a recommended category			
	Fund purchase		
	(1)	(2)	(3)
Fund recommended	16.86***	16.92***	16.29***
	(2.20)	(2.22)	(2.22)
Log fund AUM			0.32***
			(0.05)
Constant	0.63***		
	(0.09)		
Investor FE	No	Yes	Yes
Week FE	No	Yes	Yes
R^2	0.10	0.12	0.12

Notes: This table reports OLS estimates (i.e., a linear probability model) of the likelihood that an investor purchases a fund following an AI answer. The dependent variable is an indicator for purchasing any fund in a category (Panel A) or a specific fund within a recommended category (Panel B) within ten trading days of the answer. The recommendation indicator equals one if the AI recommends the category or fund, respectively. Columns (2) and (3) include individual, calendar week, and day of week fixed effects. Column (3) additionally controls for log AUM in the preceding month, measured at the category level in Panel A and the fund level in Panel B. Coefficients and standard errors are in percentage points. Standard errors are double-clustered by individual and calendar month and shown in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated with ***, **, and *, respectively.

The results above indicate that ticker recommendations are associated with subsequent purchases. However, these patterns are subject to endogeneity concerns: most importantly, AI may recommend the same or similar investments to those investors would have made anyway. The next section, Section 4.2, examines how recommended investments differ from other investments made by the same investors. Section 6 presents causal evidence on the impact of AI usage on investment.

4.2 AI Recommendations and Other Investments

Section 4.1 provides evidence consistent with investors buying recommended securities. However, implementation is incomplete, and even investors who buy recommended securities may also purchase other securities.

To further understand how AI recommendations differ from investors' choices, I move beyond the textual differences between questions and answers described in Section 3. This section compares the composition, fees, and historical risk of recommended investments with those of other purchases made following the same advice.

For each advice event, I construct two portfolios. The first consists only of the recommended securities. The second consists of alternative purchases, i.e. purchases by the same investor following advice that do not match any of the recommendations. For both of these portfolios, I calculate market beta, annualized volatility, relative Sharpe ratio loss, and maximum drawdown.

Following Calvet, Campbell, and Sodini (2007), I use relative Sharpe ratio loss to measure underdiversification. Calvet, Campbell, and Sodini (2007) define diversification loss as the expected return a household loses by not choosing a position on the efficient frontier with the same level of risk as its portfolio. Assuming that assets are priced according to the CAPM, relative Sharpe ratio loss expresses the expected excess return shortfall as a fraction of that efficient portfolio's expected excess return. The resulting measure is $\text{RSL}_p = 1 - S_p/S_B$, where S_p and S_B denote the portfolio and benchmark Sharpe ratios. Table E.3 reports how these portfolios differ.

Recommended portfolios exhibit less underdiversification and lower historical risk than portfolios of alternative purchases made by the same investors. The relative Sharpe ratio loss is 10.7 percentage points lower for recommended portfolios, falling from 28.1 to 17.4 percent, a 38.1 percent reduction. The annualized volatility is 11.27 percentage points lower and the market beta is 0.23 lower; the maximum peak-to-trough decline is 10.35 percentage points smaller. The purchased-fund expense ratios are 9.09 basis points lower for advice-following purchases than for other post-advice purchases (Table E.4).

The differences between advice-following and other purchases are also visible when investors are grouped by their pre-advice portfolio characteristics. Appendix Figure A.1 compares the two types of purchases within these groups. It provides evidence that market betas tend to be lower for advice-following purchases than for other post-advice purchases. Annualized volatility is also lower. Underlying asset HHI is lower for advice-following purchases than for other purchases. The ETF/fund share of purchases is higher for advice-following purchases. Relative Sharpe ratio loss is lower for advised purchase portfolios than for other purchase portfolios. Comparisons are made within investor and response window. Appendix Figure A.1 also addresses concerns about the sycophancy of generative AI advice in financial contexts: recommendations differ from non-recommended investments along several dimensions, and similarly also differ from existing portfolio holdings.

Advice-following purchases are more heavily allocated to diversified funds and carry lower fund fees than other post-advice purchases. Table E.4 quantifies these differences. Transactions that match AI advice,

relative to other post-advice purchases, are 36.34 percentage points more likely to be diversified funds, 40.86 percentage points less likely to be individual stocks, 10.66 percentage points more likely to be broad passive equity funds, and have expense ratios that are 9.09 basis points lower.

4.3 Recommended Securities in Investor Portfolios

In addition to immediate adherence to advice, I examine whether advice accumulates in retail investors' portfolios. Even if investors implement generative AI advice, adherence could be temporary and shifts in allocations may not be economically meaningful. Securities introduced by AI appear in investors' final observed portfolios as well as in subsequent trades. Appendix Figure A.2 relates the share of final holdings in these securities to time since first advice and cumulative interactions. Panel A plots the advised share against cumulative interactions; Panel B plots it against days since first advice. At the final observed portfolio date, 30.6 percent of investors hold at least one ticker introduced to them by gen AI advice.

Heavier AI users hold a larger share of their portfolios in securities on which they have received advice. Appendix Figure A.3 reports the same accumulation statistics for any ticker recommended by AI, including securities the investor already held. Taken together, these figures provide descriptive ticker-level evidence that investors hold securities introduced by generative AI. The evidence in this section complements the main DiD estimates in Section 6 by linking transactions and holdings to recommendations received in specific conversations.

5 Empirical Strategy

The preceding sections provide descriptive evidence consistent with investors acting on AI recommendations. However, this evidence does not provide a valid counterfactual: investors may have made the same or similar investments without AI advice. I now examine the causal impact of AI advice on investment decisions using two complementary difference-in-differences designs. The first exploits staggered variation in when users begin interacting with the AI assistant. Users begin using the assistant between November 2024 and July 2026, and I observe up to two years of historical account data, enabling me to compare their investment behavior before and after adoption. The second exploits two platform-side changes in the system prompt that made the advice more explicit and actionable, in the sense that the share of advice recommending tickers first increases sharply and then decreases.

The two designs are complementary in that they have different strengths and face different threats to identification. The staggered DiD draws on adoption events distributed across users and calendar dates, but users endogenously choose whether and when to adopt. The platform change design holds users' prior

selection into AI use fixed, but is more exposed to contemporaneous developments because the system prompt changed on common calendar dates.

The results are consistent across the two estimation strategies. Both provide evidence that AI financial advice shifts investment away from individual stocks and toward diversified, passive, low-cost funds. The magnitudes also line up quantitatively across the two designs. Scaling the estimates by the corresponding increase in exposure to diversified fund recommendations, both designs imply a similar increase in fund investment per additional fund/ETF recommendation.

5.1 Staggered DiD

I examine the impact of AI adoption using a standard difference-in-differences regression, exploiting the fact that the data provide historical transactions and portfolio details up to two years prior to a user signing up for the portfolio tracking app and starting to use the AI financial assistant.

$$y_{it} = \delta_t + \alpha_i + \beta AI_{it} + \epsilon_{it}. \tag{3}$$

As documented in Section 3.2, AI answers consistently favor passive funds over individual stocks across financial topics and types of investor requests. Motivated by this finding, the main outcome variable, y_{it}, is the share of new investment and portfolio holdings allocated to index funds and individual stocks for investor i in calendar month t. Section 7 examines the implications of these allocation changes for portfolio diversification and performance. The indicator AI_{it} equals one from the month in which investor i first uses the AI financial assistant onward. Year-month fixed effects, δ_t, control for common, time-varying factors such as market-wide shocks or tax season. Investor fixed effects, α_i, control for time-invariant differences in investment behavior across investors. Observations are investor-months, and I cluster standard errors by investor and calendar month.

A rapidly growing literature has shown that two-way fixed-effects estimators can suffer from the "forbidden comparison" problem in staggered DiD designs, where already-treated units are used as controls for later-treated units (see Roth et al., 2023, for an overview). For example, improvements in AI models may lead to higher initial effect sizes for later adopters, so that their treatment effects may be larger. To account for this heterogeneity, I study the dynamic effects of AI adoption using the interaction-weighted (IW) estimator proposed by Sun and Abraham (2021). My main estimates are robust to using other estimators introduced by the literature on staggered difference-in-differences, including those of de Chaisemartin and D'Haultfœuille (2020), Callaway and Sant'Anna (2021), and Borusyak, Jaravel, and Spiess (2024), as well as using traditional two-way fixed effects OLS.

Identification requires that, absent AI adoption, allocations to index funds and individual stocks would have evolved in parallel among AI adopters and non-adopters. While this assumption cannot be directly tested, I examine its validity in the pre-period by plotting the share allocated to passive index funds in Appendix Figure C.1, comparing non-adopters with users that eventually adopt AI by the end of my data set in July 2026. The figure provides evidence that allocations are stable, with no clear pre-trend toward passive funds. In addition, there is no statistically significant difference in average index fund allocations between adopters and non-adopters before AI adoption, with a t-test of equal means yielding a p-value of 0.389.

The main identification concern with this design is selection into treatment. AI usage is not randomly assigned, and the timing of AI adoption as well as the questions asked of the AI financial assistant are endogenously chosen by users. I address this concern in two ways. First, Section 5.2 presents a complementary identification strategy that exploits exogenous variation in the advice provided to existing AI users, induced by platform-side changes in the system prompt. These changes hold users' prior selection into AI use fixed and provide variation in advice that does not depend on their individual adoption dates. Second, Section 9.2 provides additional evidence that the shift in investment toward funds is concentrated in funds that AI is most likely to recommend. Selection based on a general desire to diversify could explain a shift from individual stocks into funds, but would not by itself explain why purchases shift disproportionately toward the particular funds favored by AI.

5.2 DiD Using System Prompt Changes

Second, I provide additional evidence on the causal impact of AI advice on investment decisions, leveraging an exogenous shift in the propensity of AI answers to provide ticker-level fund recommendations. I exploit two revisions by the app provider to the system prompt supplied to ChatGPT and Gemini. A *system prompt* consists of instructions and context supplied by the app to the language model, in addition to the investor's own question. In December 2025, the app provider revised these instructions to encourage concrete investment advice that identifies specific tickers where possible. In January 2026, the app provider revised them again to instruct the assistant to avoid recommending tickers. These changes hold users' prior selection into AI use fixed and provide variation in advice that does not depend on their individual adoption dates. The changes in the system prompt alter the advice received by AI users but do not directly affect users who do not interact with the assistant.

The platform changes create a sharp increase followed by a reversal in the frequency of explicit buy recommendations. Figure 2 plots the share of AI answers containing explicit ticker buy recommendations around these changes. The share of advice containing explicit ticker buy recommendations increases from a weekly average of 18.0 percent before the change to 33.4 percent during the high fund advice period beginning

in the calendar week of December 15, 2025, and then declines to 5.6 percent during the low fund advice period beginning in the calendar week of January 12.

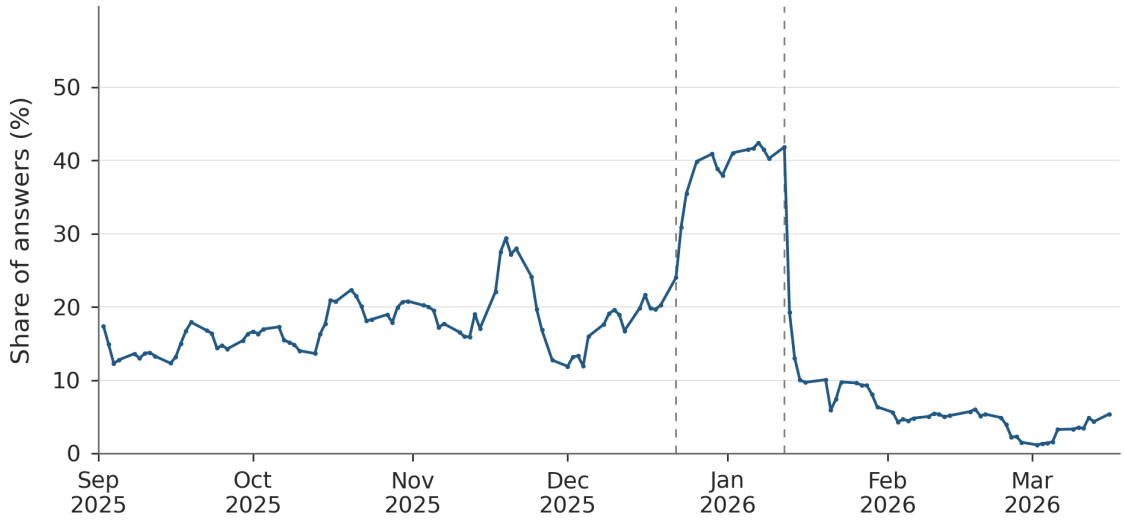


Figure 2. AI answers containing ticker buy recommendations

Notes: This figure plots the share of AI answers containing explicit ticker buy recommendations around two changes in the system prompt. The December 2025 change encouraged recommendations of specific tickers, while the January 2026 change instructed the assistant to avoid them, leading to a discrete increase and subsequent decrease in the share of answers containing such recommendations.

The treatment group consists of users who had interacted with the AI financial assistant by September 2025. I compare changes in their ETF and fund share of purchases with those of the control group, consisting of users who had not started using the assistant. The high fund advice period spans four calendar weeks, so investors must use the assistant during this window to receive advice under those instructions. To ensure a reasonable probability of receiving advice in the different treatment windows, I sort users on pre-September 2025 AI usage and compare the highest usage quartile (hereafter, AI users) with nonusers. To the extent that investors' demand for AI advice is persistent, those with higher prior usage are more likely to receive advice under the different regimes.

I estimate the following difference-in-differences regression with individual and calendar-week fixed effects:

$$y_{iw} = \alpha_i + \delta_w + \beta_H(\mathrm{AI}_i \times \mathrm{High}_w) + \beta_L(\mathrm{AI}_i \times \mathrm{Low}_w) + u_{iw}. \tag{4}$$

As in the staggered DiD, the main outcome variable, y_{iw}, is the share of new investment allocated to passive ETFs and funds by investor i in calendar week w. The coefficients β_H and β_L capture the changes

in this share for AI users relative to nonusers during the high and low fund advice periods, respectively, compared with the pre-period. Standard errors are clustered by individual and calendar week.

The identifying assumption is that, absent the shift in advice, AI users and permanent nonusers would have followed parallel trends in fund allocations. As de Chaisemartin and D'Haultfœuille (2026) discuss, this assumption is more demanding in DiD designs where groups already differ in treatment status before the change. In my setting, the effect of prior AI adoption must have stabilized over the period around the system prompt changes, so that continued AI use does not itself generate differential trends in fund allocations. Figure 3, Panel A provides supporting evidence for this condition: AI users increase the share of new investment allocated to passive funds upon adoption, and this increase persists in subsequent months. The main threat to identification in the DiD based on system prompt changes is that it relies on a narrow time period, raising concerns about contemporaneous shocks affecting AI users differently from nonusers. This concern is alleviated by the staggered DiD, since it estimates the treatment effect of AI adoption in different calendar months and thus under different market conditions.

In addition to this simple DiD, I also estimate an instrumented DiD for each prompt change (see de Chaisemartin and D'Haultfœuille, 2018, for a discussion of this methodology). An instrumented DiD is the DiD of the outcome divided by the DiD of the treatment. In this setting, it scales the DiD estimate of the change in fund investment by the change in the propensity of the AI to give fund recommendations in its answers. Intuitively, this yields an estimate of the change in fund investment per percentage point change in the fund recommendation rate. It corresponds to the usual 2SLS Wald estimator, which is simply the ratio of the reduced form (i.e. the simple DiD estimate of the change in fund investment) to the first stage, which is the change in the average rate of fund advice.

6 Main Results

6.1 Results from Staggered DiD

Figure 3 reports the Sun and Abraham (2021) event-study estimates for the monthly purchase share and the month-end portfolio share. I find an immediate 7.08-percentage-point jump in the purchase share in the adoption month. The portfolio change is more gradual as purchases accumulate. The shift in the net passive fund share is less pronounced, consistent with sale decisions also being driven by tax and rebalancing considerations.

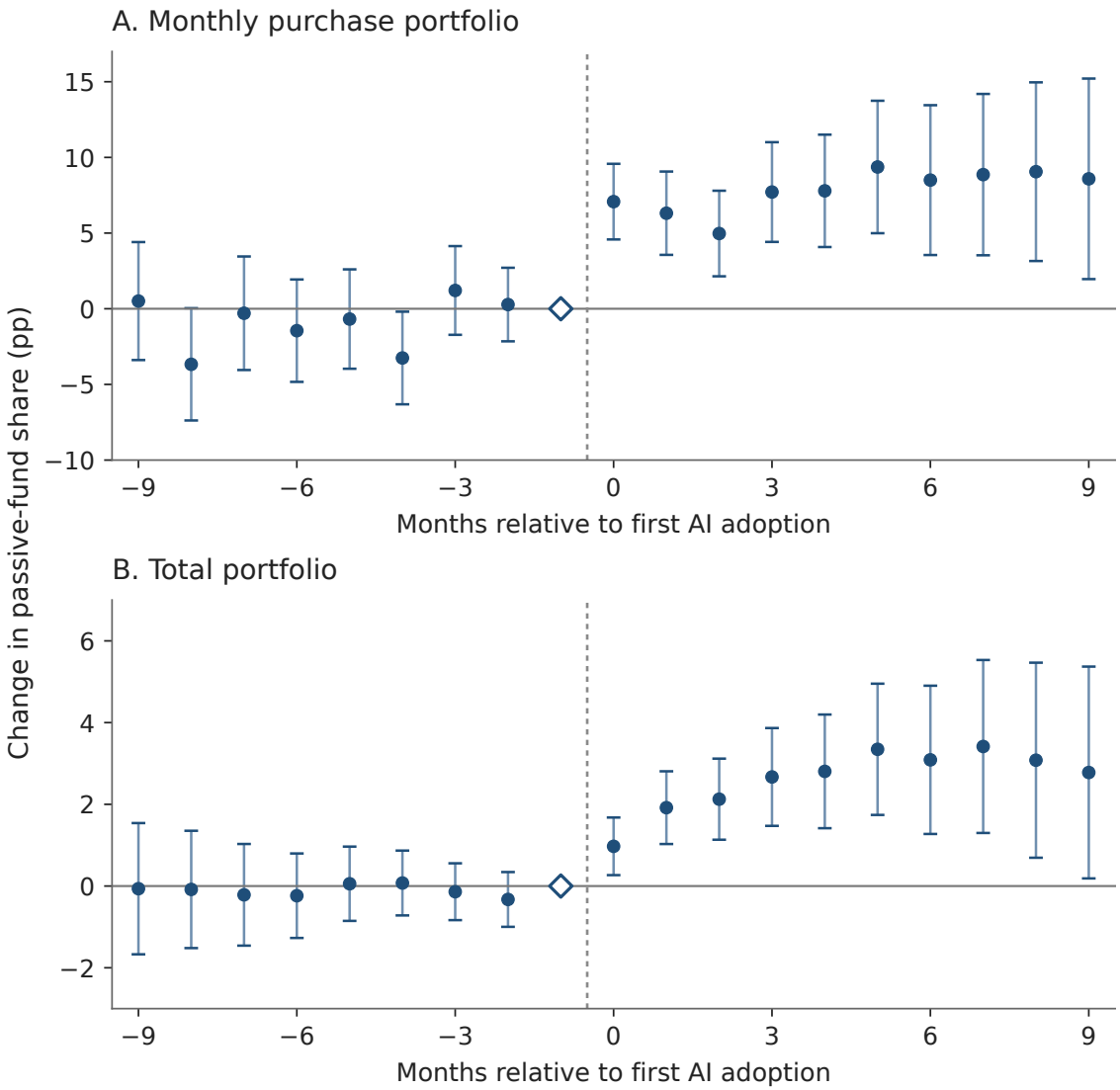


Figure 3. Passive fund shares after AI adoption

Notes: This figure plots the coefficients and 95% confidence intervals from event-study regressions of AI adoption on passive index fund shares in monthly purchases and portfolio holdings, using the Sun and Abraham (2021) interaction-weighted estimator. The dependent variable is the direct index fund share of the monthly purchase portfolio in Panel A and the direct index share of portfolio holdings at month end in Panel B. Event time is measured relative to AI adoption. Standard errors are clustered by both individual and calendar month.

AI adoption produces an economically sizable increase in the share of investment allocated to index funds. Table 4 reports the average effects on purchases and holdings. Column (1) reports that, controlling for calendar month and individual fixed effects, obtaining AI financial advice increases the share of monthly investments going to direct-index funds by 7.60 percentage points, or 22.8 percent relative to the dependent variable mean of 33.30 percent.

The reallocation toward funds is also evident after accounting for sales and in month-end holdings, alongside a decline in direct stock investment. Column (2) of Table 4 reports the estimated impact on the net passive fund share, defined as fund purchases plus individual stock sales minus fund sales and individual stock purchases, scaled by gross trading. The net passive fund share increases by 4.03 percentage points. Column (3) reports the effect on direct stock purchases: the individual stock share falls by 5.88 percentage points. Finally, columns (4) and (5) report the impacts on month-end equity holdings: the direct-index fund share increases by 2.73 percentage points, and the individual stock share falls by 2.53 percentage points.

Table 4. Effects of AI adoption on investment allocation

	Monthly purchase portfolio			Month-end portfolio holdings	
	Passive fund share	Net passive fund share	Individual stock share	Passive fund share	Individual stock share
	(1)	(2)	(3)	(4)	(5)
Post-adoption effect	7.60***	4.03**	−5.88***	2.73***	−2.53***
	(1.55)	(1.95)	(1.71)	(0.68)	(0.71)
DV mean	33.30	14.39	46.81	38.39	49.69
Observations	31,432	41,047	31,432	52,481	52,481
R^2	0.61	0.36	0.63	0.90	0.91
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Calendar month fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table presents the results of difference-in-differences regressions estimating the effect of AI adoption on investment allocation. Columns (1)–(3) report effects on the direct-index fund share of purchases, the selling-inclusive net passive fund share, and the individual stock share of purchases, respectively. Columns (4) and (5) report effects on the direct-index fund and individual stock shares of month-end equity holdings. All columns use the Sun and Abraham (2021) estimator and include individual and calendar month fixed effects. Observations are at the individual-month level. Standard errors are double-clustered by individual and calendar month and reported in parentheses. $^{*}p < .10$, $^{**}p < .05$, $^{***}p < .01$.

The increase in fund allocations leads to a complementary decrease in idiosyncratic volatility. This matters because AI adoption could increase the diversified part of the purchase portfolio while also increasing purchases of volatile individual stocks. The decline in purchase risk is robust across alternative measures of idiosyncratic volatility, and purchased funds also have lower fees. Columns (1)–(3) of Table D.1 report the alternative risk measures. Shifting from individual stocks to diversified funds lowers idiosyncratic volatility: it falls by 13.6 percent using a U.S. market benchmark, 11.0 percent using the MSCI World Index, and 11.5 percent using Welch's BSWA estimator. Column (4) reports the average volatility of stock-only purchases with no statistically significant change. Column (5) reports the expense ratio of funds purchased, which decreases by 3.57 basis points. Following Calvet, Campbell, and Sodini (2007), I also report the return loss of the combined monthly purchase portfolio. It falls by 39.9 percent.

Figure D.1 plots the distributions of purchase portfolio idiosyncratic volatility, total volatility, the passive fund share of purchases, and fund expense ratios before and after AI adoption. Figure D.2 reports the Sun and Abraham (2021) event-study estimates for the idiosyncratic volatility of purchases. Conditional on purchasing individual stocks, I find no statistically significant change in their concentration, total volatility, or idiosyncratic volatility (Table D.2). The estimated effects on fund expense ratios remain negative across alternative estimators (Table D.3). These results provide evidence that the idiosyncratic volatility of new purchases declines following AI adoption, consistent with the increase in fund investment. I find no evidence of an offsetting increase in the risk of individual stock purchases.

6.2 Results from System Prompt DiD

I next turn to the second identification strategy, which exploits the system prompt changes described in Section 5.2. First, Panel A of Table 5 reports changes in the share of AI answers recommending funds. Figure 2 shows the time series of the share of answers containing ticker buy recommendations, with vertical lines indicating the two prompt changes. The first prompt change increased the average fund recommendation rate from 18.9 percent to 34.2 percent, by 15.26 percentage points, as reported in column (2) of Panel A. The second prompt change decreased the rate from 34.2 percent to 4.5 percent, by 29.72 percentage points, as reported in column (4).

Second, Panel B of Table 5 reports the simple, reduced-form DiD estimates of the effects of the system prompt changes on fund investment. Following the December 2025 change, which increased the fund recommendation rate by 15.26 percentage points (Panel A), the fund share of new investment increases by 8.13 percentage points for AI users relative to nonusers, as reported in column (2) of Panel B. Following the January 2026 change discouraging ticker-level recommendations, this share falls by 7.67 percentage points relative to the preceding high-recommendation period, as reported in column (4) of Panel B. Both estimates are statistically significant at the 1 percent level.

Finally, Panel C reports the instrumented DiD estimates. This estimate divides the change in fund investment in Panel B, which is the reduced-form estimate of the effect of the prompt change, by the change in the fund recommendation rate in Panel A, which is the first-stage estimate. Inference uses the usual 2SLS standard errors. The resulting estimate of 0.533 in column (2) implies that an increase of one percentage point in the fund recommendation rate increases the fund share of new investment by 0.533 percentage points, and is significant at the 1 percent level. The joint F-statistic is 118.6, alleviating weak instrument concerns.

Figure 4 reports the corresponding event-study estimates around the increase in ticker-level recommendations in Panel A and the subsequent decrease in Panel B. Fund purchases rise when AI provides more

Table 5. System Prompt Changes, Fund Recommendations, and Fund Investment

	Increase in recommendations December 2025		Decrease in recommendations January 2026	
	(1)	(2)	(3)	(4)
Panel A: First-Stage DiD on Fund Recommendation Rates				
Post × AI user	15.26***	15.26***	−29.72***	−29.72***
	(3.30)	(3.31)	(3.23)	(3.24)
R^2	0.947	0.967	0.958	0.974
Panel B: Reduced-Form DiD on Fund Investment				
Post × AI user	8.13***	8.13***	−7.67***	−7.67***
	(1.97)	(1.98)	(2.23)	(2.24)
R^2	0.500	0.506	0.522	0.529
Panel C: Instrumented DiD on Fund Investment				
2SLS estimate	0.533***	0.533***	0.258***	0.258***
	(0.110)	(0.110)	(0.064)	(0.064)
R^2	0.499	0.505	0.523	0.529
Observations	11,434	11,434	7,824	7,824
Individual fixed effects	Yes	Yes	Yes	Yes
Calendar week fixed effects	No	Yes	No	Yes

Notes: Table reports difference-in-differences regressions of fund recommendation rates (Panel A) and fund investment (Panel B), with 2SLS estimates in Panel C. Fund investment is the percentage of new investment allocated to ETFs and funds. Observations are individual–calendar weeks. Post indicates weeks after the respective prompt change. Columns (1) and (2) report the effect of the December 2025 increase in recommendations relative to the preceding period. Columns (3) and (4) report the effect of the January 2026 decrease relative to the high recommendation period. All specifications include individual fixed effects, with recommendation period indicators in columns (1) and (3) and week fixed effects in columns (2) and (4). Standard errors are double-clustered by individual and calendar week and reported in parentheses. The joint first-stage F-statistic for both prompt changes, estimated over the full sample with individual and calendar week fixed effects, is 118.56. See de Chaisemartin and D'Haultfœuille (2018) and Hudson, Hull, and Liebersohn (2017) for the identifying conditions. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

ticker-level recommendations and fall when these recommendations become less frequent. These results complement the evidence from the staggered DiD and provide additional evidence that investors respond to AI recommendations that encourage fund investment.

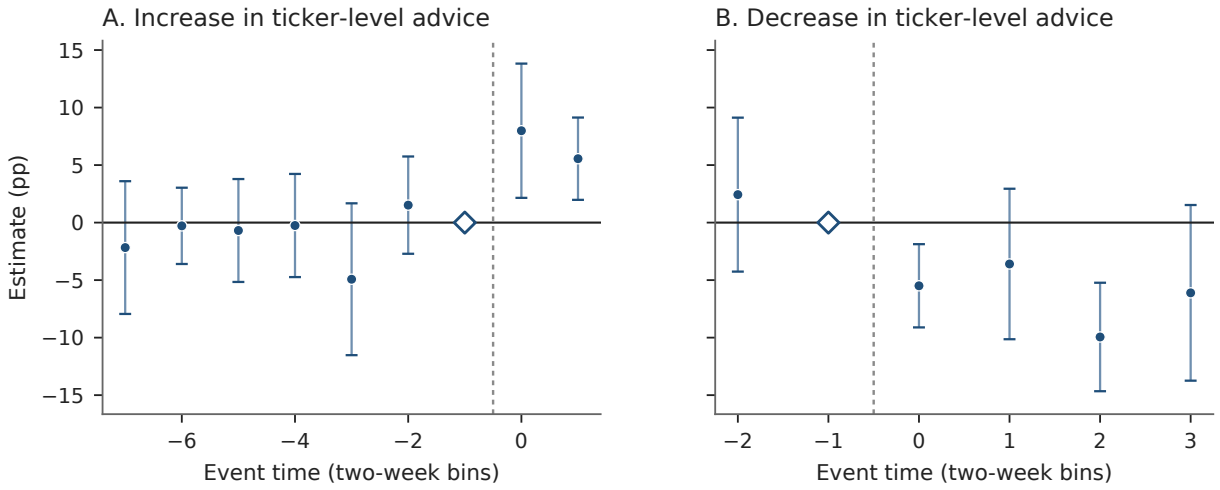


Figure 4. ETF/fund investment shares around system prompt changes

Notes: This figure plots the estimated coefficients and 95 percent confidence intervals from event-study regressions around the increase in ticker-level advice (Panel A) and the decrease in ticker-level advice (Panel B). The outcome is the share of new investment allocated to ETFs and funds. Event time is measured in two-week intervals. Regressions include individual and calendar week fixed effects. Standard errors are clustered by individual and calendar week.

An additional advantage of the instrumented DiD estimate is that it is more directly comparable to the main staggered DiD estimate after similarly scaling the fund investment response for the overall sample in the main staggered adoption design. The main staggered DiD estimate implies a 7.6 percentage point increase in fund investment after AI adoption. Across all adoption months, the weighted share of fund advice is 18 percent, naturally increasing from zero before AI adoption. Thus, using this increase to scale the main staggered DiD estimate yields $7.6/18 = 0.422$ percentage points. This scaled response therefore is comparable in magnitude to the main instrumented DiD estimate for the December increase, which implies a 0.533 percentage point increase in fund investment per percentage point increase in AI fund advice. Both empirical strategies thus yield comparable effect sizes when scaled by the respective changes in fund recommendation rates, providing causal evidence that AI advice shifts investment toward funds. Section 7 will discuss the implications of this shift for diversification and investment performance.

6.3 Heterogeneity by Portfolio Characteristics

The robo-advice literature provides evidence that investors with less diversified portfolios tend to benefit more from advice (D'Acunto, Prabhala, and Rossi, 2019; Rossi and Utkus, 2024). I explore whether this also applies to generative AI advice. I sort individuals into quintiles of pre-AI fund shares. Table 6 reports how the effect varies across quintiles. I find that the estimated effect of generative AI advice adoption on the index fund purchase share is strongest for the bottom two quintiles. This result is consistent with the

idea that individuals with lower initial fund shares have more scope to benefit from advice that encourages diversification. Section 7.1 examines how the gains from diversification vary with investors' pre-AI portfolio diversification.

Table 6. Heterogeneity by pre-AI fund share

	Q1	Q2	Q3	Q4	Q5
Index fund purchase share					
Post-adoption effect	7.45^{***}	14.35^{***}	6.82^{***}	3.99	1.04
	(1.81)	(2.53)	(2.49)	(2.68)	(1.98)

Notes: This table reports estimates for the effects of AI adoption on the index fund purchase share by pre-AI fund share quintile. Individuals are grouped into quintiles based on their pre-AI fund share, with group assignments held fixed. All columns report estimates using the Sun and Abraham (2021) estimator. Coefficients and standard errors are expressed in percentage points. Standard errors are clustered by both individual and calendar month and reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

6.4 Robustness

In this subsection, I conduct a number of robustness tests of the main results. First, I provide evidence that my main result is robust to using different staggered difference-in-differences estimators. Table D.4 reports the main estimates using Sun and Abraham (2021), stacked DiD, Callaway and Sant'Anna (2021), Borusyak et al. (2024), de Chaisemartin and D'Haultfœuille (2026), and traditional two-way fixed effects. The coefficients for the monthly purchase portfolio remain economically and statistically significant and of similar magnitude. The net passive fund share estimates are also positive and statistically significant in three of the four available specifications. The estimates for the month-end portfolio share remain positive but are smaller and less precisely estimated.

Second, I examine robustness to different passive fund definitions. Table D.5 reports estimates using the direct-index and CRSP B/D/E index fund definitions. The estimates for both the monthly purchase portfolio and the month-end portfolio remain positive and similar under the narrow and broad definitions.

Third, I examine an alternative to the staggered DiD estimation. Carlin et al. (2023) use the release of a mobile financial application in intention-to-treat and DiD specifications to study how easier access to financial information affects consumer financial decisions. I replicate this approach by using the November 2024 introduction of the AI assistant on Mezzi as the common start-of-treatment date. I compare changes in fund purchase shares between investors who begin using AI within nine months of launch and investors who do not. The plot in Appendix Figure D.3 indicates an increase in fund purchase shares among AI users relative to nonusers. Panel A plots the weekly difference in fund purchase shares between the two groups. Panel B reports the event-study estimates.

Fourth, I examine robustness to the minimum activity required in a month. Table D.6 reports estimates using every positive purchase month and purchase requirements of 0.1 and 0.5 percent of the opening portfolio. The estimated increase in the passive fund purchase share remains positive and statistically significant under each requirement.

Finally, I provide evidence that the average treatment effect is not skewed by a single adoption cohort. Table D.7 and Figure D.4 report treatment effects for early, middle, and late adopters under two cohort definitions. The first defines cohorts by dividing the months since November 2024 when the AI assistant was introduced into equal periods. This keeps the duration of the adoption windows fixed but allows the number of adopters to vary. The second defines cohorts by dividing into groups of approximatley equal numbers of adopters. This balances cohort size but allows the calendar span to vary. Although later cohorts exhibit somewhat larger point estimates, the estimates are positive across both groupings, and the equality tests do not reject the hypothesis that the three cohort effects are the same.

7 Welfare Implications

Section 6 provides evidence that investors shift both their monthly investments and portfolio allocations towards ETFs, particularly the diversified index funds most likely to be recommended by AI (Section 3). A natural next question is how this shift towards index funds benefits investors. In this section, I examine the implications of AI adoption for diversification and investment performance.

7.1 Gains from Diversification

The shift toward diversified index funds examined in Section 6 may improve investors' portfolios by reducing under-diversification, one of the investment mistakes discussed in household finance (Campbell, 2006; Campbell and Ramadorai, 2026). This section analyzes the effect of AI adoption on portfolio efficiency using measures introduced by Calvet, Campbell, and Sodini (2007).

I first consider the return loss, that is, the average return individuals lose by choosing their actual portfolio rather than a position combining the benchmark portfolio with cash to achieve the same risk level (see equation (10) in Calvet et al., 2007).[4] Graphically, return loss is the vertical distance from an investor's portfolio to the efficient frontier at the same level of risk. Before AI adoption, investors in my sample have an average annual return loss of 236 basis points, as reported in column (1) of Panel A in Table 7, comparable in order of magnitude to the 175 basis points reported for Swedish households' risky portfolios by Calvet et al. (2007) and 131 basis points reported for U.S. investors in 401(k) accounts by Tang et al. (2009).

Overall, the estimates reported in Table 7 provide evidence that AI adoption leads to more diversified portfolios and improved outcomes. The second column of Panel A in Table 7 reports the simple difference in this measure before and after AI adoption: annual return loss falls by 46.2 basis points. This means that investors forgo 19.5 percent less expected return relative to an efficient portfolio with the same level of risk. Panel B reports staggered DiD estimates, following Section 6. I find a negative and statistically significant effect of AI adoption on the return loss, of similar magnitude to the simple difference: column (2) of Panel B reports an overall reduction of 43.2 basis points, significant at the 5 percent level.

I report estimates of the effects of AI adoption on two other measures of portfolio risk, the Sharpe ratio and the portfolio idiosyncratic risk share. The portfolio idiosyncratic risk share is the share of portfolio variance that can be attributed to idiosyncratic risk, as opposed to systematic risk, and is calculated following equations (1) to (3) in Calvet, Campbell, and Sodini (2007).

To give a sense of the economic magnitudes of the effects, I compare the overall 43.2 basis point reduction in Table 7 with the return loss differences across financial numeracy groups reported by von Gaudecker (2015). Annual return losses are 44 basis points lower among investors at or above the median numeracy score than among those below it. Thus, the overall post-AI reduction in the return loss is comparable in magnitude to the difference between the low and high financial literacy groups.

There is significant heterogeneity underlying these results, with less diversified investors experiencing more diversification benefits. I define less diversified investors as those with annual return loss above the median in the last month before AI adoption. In terms of the main return loss measure, less diversified investors reduce their return loss by 101 basis points from a pre-adoption mean of 407.2 basis points (columns (3) and (4) of Panel B in Table 7). This is a 24.8 percent reduction, significant at the 1 percent level. By comparison, the effect for the more diversified half of investors is not statistically significant, as reported in column (6). Figure 5 additionally plots the Sun and Abraham (2021) event-study effects on return loss for the

[4] I follow Calvet, Campbell, and Sodini (2007) and assume that assets are priced according to an international CAPM, using the MSCI World Index as a proxy for the market portfolio. Returns are defined in excess of the return on a one-month U.S. Treasury bill. For ETFs and mutual funds, fees are subtracted before calculating returns. Return loss is $L_{it} = \sigma_{it}(S_M - S_{it})$, where σ_{it} is annual portfolio volatility, S_{it} is the portfolio Sharpe ratio, and S_M is the benchmark Sharpe ratio.

Table 7. Portfolio efficiency: raw changes and staggered DiD estimates

Measure	All investors		Less diversified		More diversified	
	Pre-AI mean	Change	Pre-AI mean	Change	Pre-AI mean	Change
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A. Raw differences (after minus before)						
Return loss (annual bp)	236.32	−46.20***	439.97	−100.24***	22.42	10.56*
		(10.59)		(20.94)		(5.08)
Annual volatility (%)	24.63	−1.14***	30.22	−2.36***	18.76	0.14
		(0.25)		(0.46)		(0.15)
Relative Sharpe loss (%)	11.65	−1.63***	20.80	−3.95***	2.03	0.81*
		(0.35)		(0.64)		(0.39)
Sharpe ratio	0.5079	0.0094***	0.4553	0.0227***	0.5631	−0.0047*
		(0.0020)		(0.0037)		(0.0023)
Idiosyncratic risk share (%)	19.95	−2.41***	35.12	−6.04***	4.01	1.41**
		(0.53)		(0.96)		(0.64)
HHI (0–100)	12.58	−2.63***	22.74	−5.87***	1.91	0.78
		(0.59)		(1.07)		(0.59)
Panel B. Staggered difference-in-differences						
Return loss (annual bp)	214.88	−43.18**	407.20	−101.06***	22.55	17.55
		(19.57)		(25.63)		(17.60)
Annual volatility (%)	24.18	−1.16**	29.56	−2.50***	18.81	0.28
		(0.47)		(0.57)		(0.45)
Relative Sharpe loss (%)	11.05	−1.18*	20.03	−3.58***	2.07	1.37**
		(0.62)		(0.74)		(0.62)
Sharpe ratio	0.5113	0.0068*	0.4597	0.0206***	0.5629	−0.0079**
		(0.0035)		(0.0043)		(0.0036)
Idiosyncratic risk share (%)	19.05	−1.79*	34.03	−5.52***	4.07	2.21**
		(0.91)		(1.07)		(0.93)
HHI (0–100)	11.67	−1.88*	21.65	−5.31***	1.69	1.74*
		(0.99)		(1.19)		(1.02)

Notes: Panel A of this table reports raw changes from the last month before AI adoption to the last observed month between 6 and 9 months after AI adoption. Panel B reports Sun and Abraham (2021) estimates. SEs are two-way clustered by investor and calendar month. Changes in percentage measures are percentage points; Sharpe ratios are unscaled. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

less and more diversified halves of investors, based on their diversification before AI adoption.

7.2 Generative AI vs Other Technological Innovations

Overall, the estimates reported in Section 7.1 provide evidence that AI adoption leads to more diversified portfolios and improved efficiency, especially among less diversified investors. I next compare the economic effects of generative AI adoption to those of other technological innovations studied in the literature. These technological innovations range from broadband internet and smartphones, which are general purpose

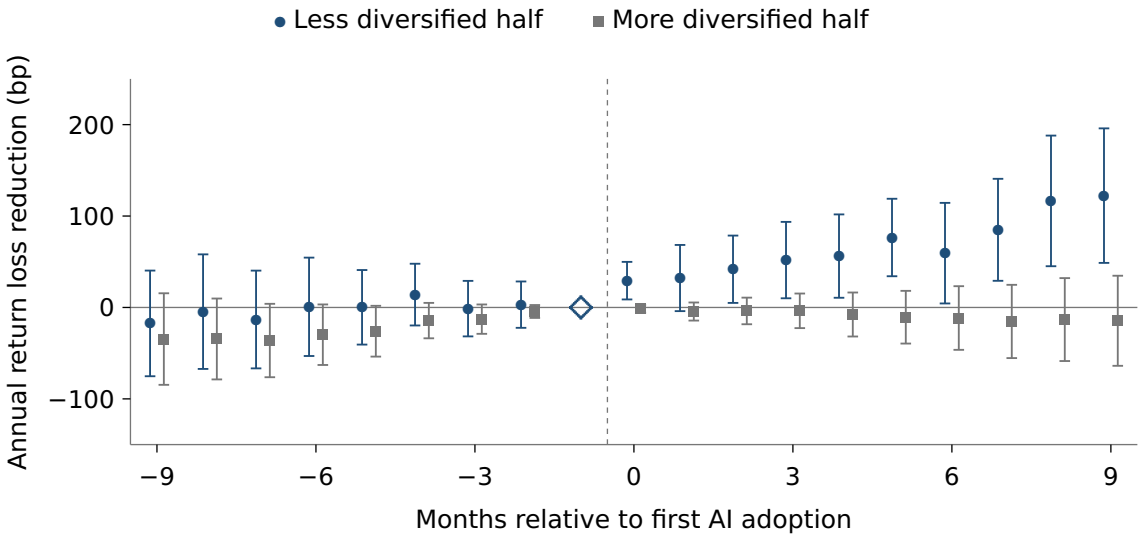


Figure 5. Effects of AI adoption on return loss by initial diversification

Notes: This figure plots event-study estimates for reductions in annual return loss following AI adoption, separately by initial diversification. Return loss is measured following Calvet, Campbell, and Sodini (2007), using the MSCI World as the benchmark. Groups use median annual return loss before AI adoption, as in Table 7. Bars indicate 95 percent confidence intervals. Standard errors are double-clustered by individual and time. Positive values indicate lower annual return loss.

technologies, to robo advisers explicitly built to optimize and outsource investment decision making.

The effects of general purpose technological innovations documented in the literature are mixed. Exploiting variation in broadband internet rollout in Norway, Hvide et al. (2024) find that internet access increases investment in equity funds and improves portfolio efficiency, which they argue may reflect easier access to information about sound portfolio allocation principles. For smartphones, on the other hand, Kalda et al. (2024) find more purchases of assets that do not improve diversification and lower Sharpe ratios, which they link to greater reliance on automatic and intuitive thinking.

The literature provides evidence that robo-advising, as a technology purpose-built to construct and maintain diversified portfolios, improves portfolio diversification and efficiency (see D'Acunto and Rossi (2023) for a literature review). The effects vary depending on the type of robo advisor. Rossi and Utkus (2024) study Vanguard's robo-advice service, where investors fully delegate portfolio management, and find a strong increase in the share of mutual fund holdings invested in index funds. By comparison, D'Acunto et al. (2019) study a portfolio optimizer in India that requires investor approval and find an increase in the number of stocks and lower portfolio volatility among initially less diversified investors.

Overall, my estimates for the effect of generative AI adoption on fund allocation and portfolio efficiency among initially less diversified investors lie between those reported for broadband internet and robo-advising. Appendix Table G.1 compares the main estimates with those found for broadband internet (Hvide et al., 2024), mobile phone trading (Kalda et al., 2024), and robo-advice (Rossi and Utkus, 2024). Among initially less

diversified investors, the index fund share of equity holdings increases by 4.8 percentage points, compared with 3.1 percentage points for the equity fund share following broadband expansion and 35 percentage points for the index fund share of mutual fund holdings following adoption of Vanguard's adviser, which fully rebalances investors' portfolios. The Sharpe ratio increase likewise exceeds that reported for broadband internet but is smaller than that reported for robo-advice.

Apart from effect size, another useful dimension along which to compare generative AI adoption is the extent of adoption. Generative AI, while still expanding, has been adopted faster than the internet and the personal computer (Bick et al., 2026; Lynch, 2026). As of 2026, surveys report that 54–70 percent of investors have used AI to obtain investment information or support financial decisions (Opinium, 2026; Amundi, 2026). While robo-advice can be highly effective, its use remains limited: in a 2026 Gallup survey, 3 percent of U.S. adults who sought financial guidance in the previous year used a robo adviser (Kemp, 2026). Overall, the results above suggest that the rapid adoption of generative AI may bring investment choices more in line with the advice from portfolio theory and improve diversification for retail investors who adopt this technology.

8 Personalization of AI Answers

8.1 Who Takes Advice, and Why?

Sections 6 and 7 provide evidence that investors shift their portfolios towards funds following AI adoption, thereby improving diversification, consistent with the emphasis on index fund investing and diversification in AI answers. Overall, the shift toward funds and improvement in diversification contrast with the limited changes in behavior documented in other settings: Bhattacharya et al. (2012) report that only about 5 percent of investors offered free, unbiased advice obtain it, and the average advisee does not improve portfolio diversification.

This is consistent with a wider behavioral literature providing evidence that it is difficult to change behavior. Fernandes, Lynch, and Netemeyer (2014) find that the effects of financial education on behavior are small, even for programs involving many hours of instruction. This raises the question of what features of generative AI advice lead to the immediate and persistent shift documented in Section 6. Section 8.1.1 examines the effect of personalization on implementation.

8.1.1 Personalization

Personalization is one of the key mechanisms explored in the nudge literature as a way of altering people's behavior (DellaVigna and Linos, 2022). At the same time, a key feature of generative AI chatbots is their

ability to tailor answers to the user's circumstances. In my setting, AI answers can be personalized using two sources of information about the investor: information provided directly through current or previous conversations and portfolio information supplied by the platform.

I begin by constructing two alternative measures of personalization. First, I use a quantitative measure that counts personalization details. Second, I use a qualitative measure in which an open-source LLM assigns an overall personalization score. The counted details cover portfolio state and life details. Portfolio state covers current financial facts, such as holdings, allocation, concentration, cash balances, and account structure. Life details describe personal circumstances, such as age, employment, dependents, or retirement status. The Online Appendix reports the classification prompt. Appendix H provides examples of both measures. The correlation between the detail counting measure and the LLM categorization measure is 0.75.

Personalization primarily draws on portfolio information of the kind supplied by the platform. Portfolio details account for 85.3 percent of the personalization details, as reported in Appendix Table H.1. Using these measures, I examine whether investors are more likely to implement recommendations in more personalized AI answers.

I start by estimating a simple linear probability model relating implementation to personalization:

$$I_{iqt} = \alpha_i + \beta P_{iqt} + \varepsilon_{iqt}, \tag{5}$$

where I_{iqt} denotes implementation of the recommendation in answer q received by investor i on date t, and P_{iqt} denotes personalization in that answer, expressed in standard deviation units. α_i denotes investor fixed effects. As in Section 4, implementation is a dummy variable equal to one if the investor buys at least one recommended ticker within ten trading days of the answer. The coefficient β measures the association between personalization and the probability that the investor implements the recommendation.

More personalized answers are more likely to be implemented under both measures of personalization, including after accounting for investor and month fixed effects. Appendix Table H.2 reports the regression estimates. The sample is restricted to ETF/fund recommendations, which account for 89% of ticker-level recommendations. Panel A reports results for the LLM-generated measure, and Panel B for the number of personalization details. In the pooled specification, a one-standard-deviation increase in personalization is associated with a 5.64-percentage-point higher implementation rate for the LLM score and a 3.50-percentage-point higher rate for the detail count. The positive association is robust to adding investor and month fixed effects. Investor fixed effects control for persistent differences in investors' tendencies to trade and implement recommendations, while month fixed effects control for conditions common to investors receiving answers in the same calendar month, including market conditions. With both sets of fixed effects, the corresponding

associations are 3.24 and 1.86 percentage points.

Figure 6 illustrates this relationship. The positive association remains after controlling for investor and month fixed effects: a one-point increase in the personalization score is associated with a 1.56-percentage-point increase in implementation (standard error 0.34).

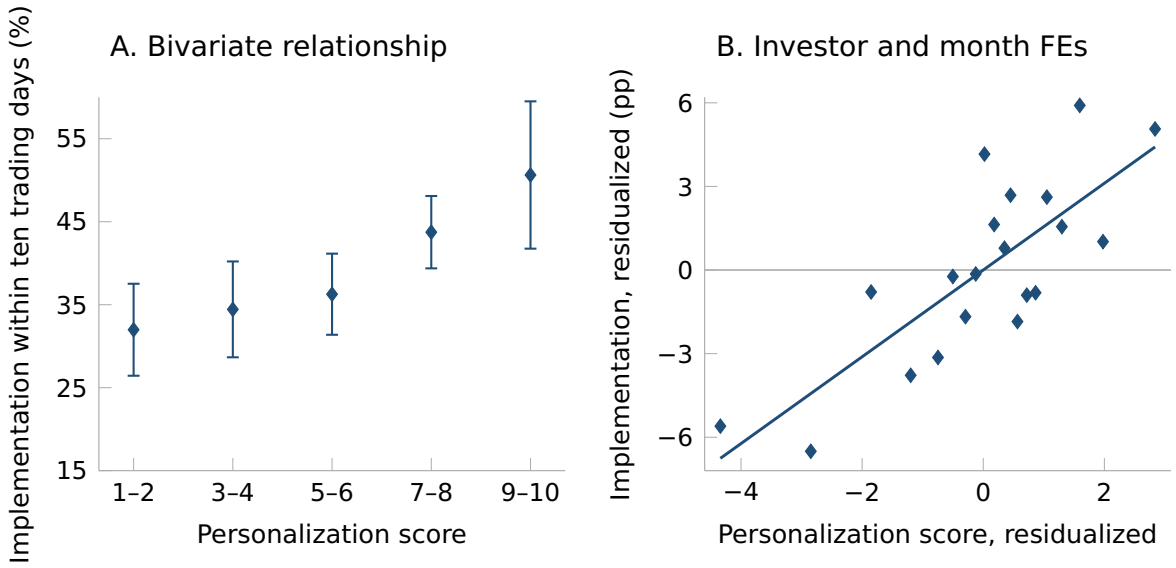


Figure 6. Personalization and implementation within ten trading days

Notes: Panel A plots average advice implementation rates within ten trading days by personalization score. Panel B plots the same relationship after controlling for investor and calendar month fixed effects.

There are several reasons for this regression to be biased. In particular, investors' skill at using AI, the types of questions they ask, and unobserved investment intentions may affect both the personalization of answers and whether advice is followed. While the descriptives provide evidence that most personalization draws on portfolio information of the kind supplied by the platform, investors' questions may still influence how this information is used in an answer. For instance, an investor who already intends to trade may provide more personal information or ask a more specific question, eliciting a more personalized answer.

To address these concerns, investor fixed effects account for stable differences across investors, including differences in general cognitive ability and the propensity to act on recommendations. To account for experience with AI, I also control for the investor's number of previous interactions. Question type controls distinguish information requests from advice and action requests, while topic controls account for changes in the financial subjects investors ask about. Additional controls include the log of the question token count and the number of named securities. I also include day-of-week fixed effects and control for purchases and sales during the 30 calendar days before the question.

The positive relationship between personalization and implementation remains after controlling for

question characteristics, AI experience, and prior trading. Column (2) of Table 8 reports the OLS estimates with these controls. With the full set of controls, a one-standard-deviation increase in personalization is associated with a 2.92-percentage-point higher implementation rate for the LLM measure and 1.83 percentage points for portfolio details. Both estimates remain positive after adding controls to column (1), which includes only investor fixed effects. Portfolio details remain positively associated with implementation when portfolio and life details enter jointly (Appendix Table H.3).

While it is reassuring that the personalization effect is driven by portfolio details, which are supplied by the app rather than by the user, the concern of skill-based differences is only partially addressed by controlling for previous AI usage. An ideal empirical setup to estimate the causal effect of personalization would be if investors were randomly assigned to LLMs giving different levels of personalized advice in order to obtain exogenous variation in personalization.

To approximate this setup, note that a significant improvement in LLM generations over my sample period concerns their ability to handle long contexts. This improvement can affect the level of portfolio personalization, since that relies on processing long financial context. Appendix Figure H.1 provides evidence that implementation rates have increased along with personalization.

I thus use assignment to different models, which depends on time, to instrument for portfolio personalization. I estimate the effect of portfolio personalization on implementation using two-stage least squares (2SLS). As in the dummy instrument formulation described by Goldsmith-Pinkham et al. (2026), 2SLS combines the model indicators into a single predictor, obtained from the first-stage regression of portfolio personalization on these indicators, controls, and investor fixed effects. The first-stage regression is

$$P_{iqt} = \sum_{m=2}^{M} \pi_m Z_{miqt} + \mathbf{X}'_{iqt}\delta + \mu_i + v_{iqt}, \tag{6}$$

where $Z_{miqt} = \mathbf{1}\{M_{iqt} = m\}$ indicates that model m generated the answer, and M_{iqt} is the model category. The model indicators include GPT-4, GPT-4.1, GPT-5, GPT-5.2, and the Gemini 3 and 3.1 model. $\mathbf{X}_{iqt}$ contains the controls described above, and μ_i denotes investor fixed effects. In the second stage, I use the fitted value $\widehat{P}_{iqt}$ in place of P_{iqt}. Both stages include the same controls and investor fixed effects and use the same observations.

Identifying Assumptions. Whether the model indicators can be used as valid instruments for variation in personalization hinges on the standard relevance and exclusion restriction assumptions of the IV framework. Evaluating the relevance assumption, Appendix Table H.4 reports first-stage estimates of a regression of personalization on different model indicators, i.e., each model is represented by an indicator. The model indicators strongly predict portfolio personalization, with a Montiel Olea and Pflueger (2013) effective

F-statistic of 44.23 in the baseline specification (Panel B of Table 8), exceeding the critical value of 18.29 and relieving concerns about weak instruments. The first-stage estimates change little when I include controls for question characteristics, previous AI interactions and pre-advice trading activity, and day-of-week fixed effects.

The exclusion restriction requires that model identity affect implementation only through personalization. While the exclusion restriction is not directly testable, the main concerns about it can be examined empirically in this case. In particular, the rich textual data allow for additional tests of other features of AI answers that may have changed across models. The key concerns are changes in the recommendations themselves (e.g., it is possible that models suggest tickers that are popular or get better at suggesting tickers that an investor is likely to buy) and substantive changes in the arguments for investments (e.g., an emphasis on diversification or simplicity), rather than in the qualitative nature of how the arguments are given in terms of personalization. Question characteristics predict actual personalization but do not jointly predict the assigned model's personalization tendency, conditional on investor fixed effects and prior activity.

Results. The IV comparison focuses on variation in personalization predicted by model assignment, while the OLS estimates provide a benchmark. Table 8 reports the results from estimating Equation (5). In columns (1)–(2), I present the OLS estimates for comparison purposes. Columns (3)–(4) report results when using the model indicators to instrument for portfolio personalization in the first-stage regression. Appendix Table H.4 presents first-stage results for the same sample.

Since column (4) accounts for stable differences across investors and observable differences in question characteristics, experience with AI and pre-advice trading activity, as well as day-of-week effects, it is my preferred specification. In this column, a one-standard-deviation increase in portfolio personalization (3.04 details) corresponds to an 11.87-percentage-point increase in implementation, with a standard error of 3.91, compared with an OLS estimate of 1.83 percentage points. The IV estimate is equivalent to 27.5 percent of the sample mean implementation rate of 43.14 percent.

Table 8. Personalization and implementation: OLS and IV

Dependent variable: implementation (percentage points)	OLS		2SLS	
	(1) Investor FE	(2) Full controls	(3) Investor FE	(4) Full controls
Panel A. Qualitative LLM measure				
Personalization	3.56***	2.92***	6.82**	8.01***
	(0.82)	(0.79)	(2.83)	(2.62)
R^2	0.26	0.27	0.25	0.26
First-stage F (clustered)	—	—	49.63	92.82
MOP effective F	—	—	95.85	165.94
Panel B. Detail counting measure				
Personalization	2.24***	1.83**	10.29**	11.87***
	(0.77)	(0.80)	(4.25)	(3.91)
R^2	0.25	0.27	0.23	0.24
First-stage F (clustered)	—	—	42.73	64.10
MOP effective F	—	—	35.74	44.23
Investor fixed effects	Yes	Yes	Yes	Yes
Topic fixed effects	No	Yes	No	Yes
Day-of-week fixed effects	No	Yes	No	Yes
Other question / experience controls	No	Yes	No	Yes
Prior 30-day purchases and sales	No	Yes	No	Yes
DV mean (%)	43.14	43.14	43.14	43.14

Notes: This table reports estimates for the relationship between personalization and advice implementation using the regression specification in Equation (5). Columns (1)–(2) report OLS estimates, and columns (3)–(4) report 2SLS estimates using model indicators as instruments for personalization. Panel A measures personalization using the qualitative LLM score, and Panel B uses the number of portfolio details. The dependent variable indicates whether the investor buys at least one recommended security within ten trading days of the answer. Coefficients are expressed in percentage points per standard deviation of personalization. All columns include investor fixed effects. Columns (2) and (4) include the additional controls listed in the table. MOP denotes the effective first-stage F-statistic of Montiel Olea and Pflueger (2013). Standard errors are clustered by investor and answer date and reported in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

8.2 Financial Vocabulary in Subsequent Prompts

In addition to personalization, I find suggestive evidence that investors learn from interacting with generative AI: financial terms used in AI answers are more likely to appear in investors' subsequent prompts, including in later conversations. Whether AI helps individuals acquire knowledge, beyond providing answers to their immediate questions, is a central issue in the literature on AI and learning (Brynjolfsson et al., 2025; Bastani et al., 2025). I examine one potential indication of learning: whether individuals subsequently use financial concepts introduced by generative AI in their own questions. To do so, I estimate the effect of AI introducing a financial concept on subsequent prompts by investors. I estimate the following regression at the individual–concept–conversation level:

$$Y_{ic,j+1} = \beta\, \text{AIIntro}_{icj} + \alpha_i + \tau_{w(j)} + \varepsilon_{icj}, \tag{7}$$

The outcome variable, $Y_{ic,j+1}$, equals one if individual i uses concept c in the first prompt of conversation

$j+1$. My main variable of interest is AIIntro_{icj}, an indicator that equals one if generative AI uses the concept in conversation j before the individual has used it in that or any earlier conversation. I include individual fixed effects, α_i, and calendar week fixed effects, $\tau_{w(j)}$, where $w(j)$ denotes the week of conversation j. Standard errors are double-clustered by individual and time.

For each individual and concept, the sample retains conversations until the first observed use by either the individual or AI. I group different forms of the same word under one concept. The main estimates pool four concepts: fund, allocation, diversification, and index. I add volatility, concentration, and risk for the concept-specific examples in Panel B of Appendix Table I.1.

Panel A of Appendix Table I.1 provides evidence that AI introduction is positively associated with subsequent use of financial concepts. After controlling for individual and calendar week fixed effects, the estimated coefficient is 1.98 percentage points (standard error 0.61), compared with an outcome mean of 4.67 percent. The coefficients decline with the inclusion of fixed effects, but remain positive and statistically significant.

The association between AI introduction and later vocabulary use varies across financial concepts. Panel B of Appendix Table I.1 reports the concept-specific estimates. The coefficients on fund, allocation, diversification, and risk are positive and statistically significant. In contrast, the coefficients on volatility and concentration are negative and statistically significant.

I also examine a systematically selected vocabulary to assess whether the analysis depends on the initial choice of financial concepts. Appendix Table I.2 reports this alternative selection. It retains only terms that appear at least three times as often in AI answers as in prompts and that appear in at least 10 percent of AI answers. Appendix Table I.3 reports the retained financial terms that meet these criteria. Panel A reports the prevalence of selected financial terms in AI answers and prompts.

9 AI Adoption and the Concentration of Investment Demand

Section 7 provides evidence that individual investors may benefit from the adoption of generative AI as a new information technology by democratizing access to advice consistent with portfolio theory. While AI adoption may thus improve individual investors' portfolios through a shift towards low-cost index investing, a separate question is how this adoption may change cross-sectional variation in investment behavior and what this may imply for asset markets. This question relates to the literature on the role of retail trades in financial markets, which has examined noise trading (De Long et al., 1990), herding (Barber et al., 2009), and attention-driven buying (Barber and Odean, 2008; Barber et al., 2022).

In this section, I examine the concentration of AI recommendations and its implications for investment

demand. I first compare the concentration of fund recommendations with the concentration of assets within Lipper categories. I then provide evidence that, after adoption, investors allocate a larger share of monthly investment to the most widely recommended funds, including within fund investment.

9.1 Concentration in investment recommendations

LLMs have been described as next token predictors or "stochastic parrots." Although these characterizations provide an incomplete account of their capabilities, it remains true that their responses can concentrate on a limited set of alternatives, even when many answers are possible. Post-training can reinforce this tendency while improving the quality of individual responses (Lin, 2026; Chen et al., 2026). As a result of this feature, LLMs tend to exhibit homogeneity when making recommendations, a feature that, e.g., Allouah et al. (2025) document in the context of simulated online shopping.

Homogeneity can be a desirable feature: factual questions often have one correct answer. Financial information may also be a setting where most questions have an answer that is either factually correct or correct in the sense that it is consistent with academic portfolio theory. However, the choice homogeneity documented in other product recommendation settings may have adverse implications in financial markets by increasing the concentration and correlation of demand. In the context of retail trading on Robinhood, Barber et al. (2022) document that the app's stock displays lead to correlated retail demand, which can cause price pressure and subsequent negative abnormal returns.

It is thus useful to start by documenting to what extent the product homogeneity is present in LLM answers to financial questions. To examine the concentration in ticker-level recommendations, I first identify the 10 tickers which were recommended to the most investors by each ChatGPT and Gemini model in my dataset from October 2024 to July 2026. For each of these, I calculate the share of investors receiving a recommendation to buy this ticker as the number of investors that received a given ticker-level recommendation divided by the total number of investors who received ticker-level recommendations during the time a given Gen AI model was in use. Appendix Figure J.1, Panel A, provides evidence that the recommendations are highly concentrated: across models, the top 10 most commonly recommended tickers are recommended to 71.5 percent of investors receiving ticker-level recommendations, on average. VTI is the most widely recommended ticker for six of the seven models when ranked by distinct investors, consistent with the broad market index focus of recommendations discussed in section 3.2.

A natural question to ask is how the concentration of recommendations on certain funds compares to how much market share those funds have within their Lipper categories, that is, can the concentration simply be explained by these tickers being the most popular within their category? To do so, I compare

recommendation concentration with the concentration of AUM within the ten largest Lipper categories by total fund AUM. Specifically, I define s_i as fund i's share of all distinct investor ticker recommendations or total AUM in its category and calculate HHI = $\sum_i s_i^2$ separately for each distribution. I also calculate top three and top ten shares alongside the HHIs, using investor ticker recommendations. I rank funds separately by recommendations and AUM within each category. Differences are recommendations minus AUM, expressed in percentage points for the top shares; HHIs and their differences are reported without scaling.

I find that recommendations are more concentrated than fund assets within Lipper categories. In terms of magnitudes, the mean HHI for recommendations is 0.3690, approximately 2.9 times the AUM index of 0.1263. The difference of 0.2427 is statistically significant at the 1 percent level. Recommendation concentration exceeds AUM concentration in all seven models. For example, within the Lipper Multi-Cap Core Funds category, VTI accounts for 53.9 percent of investor ticker recommendations, compared with 15.8 percent of category AUM.

The top three funds account, on average, for 82.4 percent of investor ticker recommendations within Lipper categories, compared with 49.1 percent of category AUM for the three largest funds. The corresponding top ten shares are 99.1 percent for recommendations and 75.1 percent for AUM. Both differences are statistically significant at the 1 percent level.

In summary, this analysis provides evidence that the feature of recommendation homogeneity extends to financial advice. Note that this homogeneity in ticker recommendations is particularly interesting given that these tickers are extracted from answers to a variety of questions asked by different investors, as described in section 3.1.

9.2 Concentration in recommendations and fund purchases

The previous section documents ticker-level homogeneity in LLMs' investment recommendations. I next investigate whether the recommendations translate into increased investment concentration in the AI-preferred tickers for AI adopters.

To do so, I first repeat the main staggered DiD specification restricting the purchase numerator to the top 20 fund recommendations (all of which are passive index funds). After AI adoption, the share of monthly investments in the top 20 AI recommended funds increases by 4.81 percentage points, as reported in column (2) of Table 9. For the top ten funds, the increase is 3.02 percentage points. Comparing the top twenty increase to the overall 7.6 percentage point increase in passive index investing discussed in section 6.1, this implies that the increase in only the top 20 funds accounts for 63.3% of the increase in passive index investing among AI adopters.

One concern with this comparison is that the increase in purchases of the top twenty recommended funds could reflect an overall shift toward fund investing without a change in the composition of the funds users buy. I thus additionally examine whether the share of the top 20 AI recommendations also increases within new monthly fund investments only. Indeed, this share rises by 7.19 percentage points, as reported in Table 9, column (3), indicating a shift toward these particular funds within fund investing. The corresponding increase for the top ten funds is 4.94 percentage points. Appendix Figure J.2 plots the changes around AI adoption in the top 20 share (Panel A) and their share within fund investment (Panel B).

Finally, I examine the HHI in terms of ticker-level investment, abstracting from the diversification benefits provided by funds, as discussed in Section 7, to investigate whether investment is concentrated in fewer securities. This is reported in column (4) of Table 9. The HHI increases by 2.98 points on a scale from 0 to 100 after AI adoption (standard error 1.31), indicating more concentrated purchases across tickers.

Overall, these results suggest that AI adoption may concentrate retail investment in AI recommended securities. Given widespread adoption, this may have implications for financial markets and the changing role of retail investment by both reinforcing the shift towards passive index fund investing and concentrating investment in a small set of commonly recommended funds.

Table 9. AI adoption and the composition and concentration of monthly investment

	Monthly investment		Monthly fund investment	Monthly investment
	Passive fund share	Top 20 fund share	Top 20 fund share	Ticker HHI (×100)
	(1)	(2)	(3)	(4)
Post-adoption effect	7.60***	4.81***	7.19***	2.98**
	(1.55)	(1.28)	(1.26)	(1.31)
DV mean	33.30	13.54	25.59	38.69
R^2	0.61	0.48	0.48	0.55
Individual fixed effects	Yes	Yes	Yes	Yes
Calendar month fixed effects	Yes	Yes	Yes	Yes

Notes: This table reports estimates for changes in the composition and concentration of monthly investment after AI adoption using Sun and Abraham (2021). Columns (1) and (2) report passive fund and top twenty fund shares of monthly investment. Column (3) reports the top twenty share within monthly fund investment. Column (4) reports the HHI of monthly investment multiplied by 100, treating each fund as one security. Investment shares are in percent. Standard errors are clustered by individual and calendar month and reported in parentheses. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

10 Conclusion

Generative AI has become one of the most rapidly adopted technologies. This rapid adoption extends to financial advice, with a recent Investing.com survey reporting that nearly two-thirds of U.S. retail investors have used generative AI tools to assist with investment decisions (Investing.com, 2026). This paper offers the first assessment, to my knowledge, of how generative AI financial advice affects observed investment decisions and portfolio allocations. Using linked records of investors' questions, AI answers, security-level recommendations, transactions, and holdings, I find that generative AI predominantly recommends funds in response to open-ended requests for investment ideas. After investors begin using generative AI, their purchases shift toward broad passive index funds, and this shift subsequently appears in portfolio holdings. The securities they purchase carry less idiosyncratic risk, and the funds they select charge lower fees. These portfolio changes address problems widely documented in household finance: retail investors often hold underdiversified portfolios and bear uncompensated idiosyncratic risk (Calvet et al., 2007; Badarinza et al., 2016), and they frequently select investment products with unnecessarily high fees or otherwise undesirable fund characteristics (Choi et al., 2010; Bailey et al., 2011). This shift in investment behavior is driven by AI advice consistently directing investors toward greater diversification through low-cost index investing, as prescribed by standard financial theory (Markowitz, 1952; Campbell and Viceira, 2002; Campbell, 2006). I find high rates of advice implementation. This is notable given extensive evidence that changing household financial behavior is difficult (Madrian and Shea, 2001; Bhattacharya et al., 2012; Fernandes et al., 2014; Andersen et al., 2020). Overall, my findings suggest that generative AI can improve household investment decisions by providing financial guidance that is consistent with standard financial theory and that investors act on.

References

Allouah, Amine, Omar Besbes, Josué D. Figueroa, Yash Kanoria, and Akshit Kumar. 2025. "What Is Your AI Agent Buying? Evaluation, Biases, Model Dependence, & Emerging Implications for Agentic E-Commerce." Working paper, arXiv:2508.02630, version 3, December 17. https://arxiv.org/abs/2508.02630.

Amundi. Decoding Investors: Retirement Planning Takes Center Stage. Survey report, 2026. Survey findings.

Andersen, Steffen, John Y. Campbell, Kasper Meisner Nielsen, and Tarun Ramadorai. Sources of inaction in household finance: Evidence from the Danish mortgage market. *American Economic Review*, 110(10): 3184–3230, 2020.

Appel, Ruth, Maxim Massenkoff, Peter McCrory, Miles McCain, Ryan Heller, Tyler Neylon, and Alex Tamkin. Anthropic Economic Index report: Economic primitives. Anthropic research report, January 2026.

Badarinza, Cristian, John Y. Campbell, and Tarun Ramadorai. International comparative household finance. *Annual Review of Economics*, 8:111–144, 2016.

Bailey, Warren, Alok Kumar, and David Ng. Behavioral biases of mutual fund investors. *Journal of Financial Economics*, 102(1):1–27, 2011.

Barber, Brad M. and Terrance Odean. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance*, 55(2):773–806, 2000.

Barber, Brad M. and Terrance Odean. Online investors: Do the slow die first? *Review of Financial Studies*, 15(2):455–488, 2002.

Barber, Brad M., and Terrance Odean. 2008. "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." *Review of Financial Studies* 21(2): 785–818. https://doi.org/10.1093/rfs/hhm079.

Barber, Brad M., Terrance Odean, and Ning Zhu. 2009. "Do Retail Trades Move Markets?" *Review of Financial Studies* 22(1): 151–186. https://doi.org/10.1093/rfs/hhn035.

Barber, Brad M., Xing Huang, Terrance Odean, and Christopher Schwarz. 2022. "Attention-Induced Trading and Returns: Evidence from Robinhood Users." *Journal of Finance* 77(6): 3141–3190. https://doi.org/10.1111/jofi.13183.

Bastani, Hamsa, Osbert Bastani, Alp Sungu, Haosen Ge, Özge Kabakcı, and Rei Mariman. Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26):e2422633122, 2025.

Battiston, Diego, Jordi Blanes i Vidal, Rafael Hortala-Vallve, and Dong Lou. The effect of advisors' incentives on clients' investments. *Journal of Finance*, 81(3):1701–1739, 2026.

Bhattacharya, Utpal, Andreas Hackethal, Simon Kaesler, Benjamin Loos, and Steffen Meyer. Is unbiased financial advice to retail investors sufficient? answers from a large field study. *Review of Financial Studies*, 25(4):975–1032, 2012.

Bhattacharya, Utpal, Benjamin Loos, Steffen Meyer, and Andreas Hackethal. Abusing etfs. *Review of Financial Studies*, 30(4):1289–1320, 2017.

Bianchi, Milo and Marie Brière. Human-robot interactions in investment decisions. *Management Science*, 72(1):14–31, 2026.

Bick, Alexander, Adam Blandin, and David J. Deming. The rapid adoption of generative AI. *Management Science*, published online January 20, 2026.

Bini, Pietro, Lin William Cong, Xing Huang, and Lawrence J. Jin. Behavioral economics of AI: LLM biases and corrections. NBER Working Paper 34745, 2026.

Blankespoor, Elizabeth, Jade Croom, and Stephanie M. Grant. Generative ai and investor processing of financial information. *Journal of Accounting and Economics*, page 101908, 2026.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6):3253–3285, 2024.

Brynjolfsson, Erik, Danielle Li, and Lindsey R. Raymond. Generative ai at work. *Quarterly Journal of Economics*, 140(2):889–942, 2025.

Callaway, Brantly and Pedro H. C. Sant'Anna. Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230, 2021.

Calvet, Laurent E., John Y. Campbell, and Paolo Sodini. Down or out: Assessing the welfare costs of household investment mistakes. *Journal of Political Economy*, 115(5):707–747, 2007.

Campbell, John Y. Household finance. *Journal of Finance*, 61(4):1553–1604, 2006.

Campbell, John Y. and Tarun Ramadorai. Household finance in retrospect and prospect. NBER Working Paper No. 34621, 2026.

Campbell, John Y. and Luis M. Viceira. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press, 2002.

Carlin, Bruce, Arna Olafsson, and Michaela Pagel. Mobile apps and financial decision making. *Review of Finance*, 27(3):977–996, 2023. doi: 10.1093/rof/rfac040.

Chalmers, John and Jonathan Reuter. Is conflicted investment advice better than no advice? *Journal of Financial Economics*, 138(2):366–387, 2020.

Chatterji, Aaron, Thomas Cunningham, David J. Deming, Zoe Hitzig, Christopher Ong, Carl Yan Shan, and Kevin Wadman. How people use ChatGPT. NBER Working Paper 34255, 2025.

Chen, Yilei, Souradip Chakraborty, Lorenz Wolf, Ioannis Paschalidis, and Aldo Pacchiano. 2026. "Post-training Large Language Models for Diverse High-Quality Responses." *International Conference on Learning Representations*. https://proceedings.iclr.cc/paper_files/paper/2026/hash/cfeb1fa78c5fb06be27c0980aa8ee581-Abstract-Conference.html.

Choi, James J., David Laibson, and Brigitte C. Madrian. Why does the law of one price fail? an experiment on index mutual funds. *Review of Financial Studies*, 23(4):1405–1432, 2010.

Choi, James J., David Laibson, and Andrew Metrick. How does the internet affect trading? Evidence from investor behavior in 401(k) plans. *Journal of Financial Economics*, 64(3):397–421, 2002.

Choi, James J. and Adriana Z. Robertson. What matters to individual investors? Evidence from the horse's mouth. *Journal of Finance*, 75(4):1965–2020, 2020.

Choukhmane, Taha, Tim de Silva, Weidong Lin, and Matthew Akuzawa. AI financial advice: Supply, demand, and life cycle implications. NBER Working Paper 35574, 2026.

D'Acunto, Francesco, Nagpurnanand Prabhala, and Alberto G. Rossi. The promises and pitfalls of robo-advising. *Review of Financial Studies*, 32(5):1983–2020, 2019.

D'Acunto, Francesco and Alberto G. Rossi. Robo-advice: Transforming households into rational economic agents. *Annual Review of Financial Economics*, 15:543–563, 2023.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2018. Fuzzy Differences-in-Differences. *Review of Economic Studies* 85(2): 999–1028.

de Chaisemartin, Clément and Xavier D'Haultfœuille. Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 108(4):863–880, 2026.

De Long, J. Bradford, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. 1990. "Noise Trader Risk in Financial Markets." *Journal of Political Economy* 98(4): 703–738. https://doi.org/10.1086/261703.

DellaVigna, Stefano, and Elizabeth Linos. RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica*, 90(1):81–116, 2022. doi:10.3982/ECTA18709.

Dimmock, Stephen G., Roy Kouwenberg, Olivia S. Mitchell, and Kim Peijnenburg. Household portfolio underdiversification and probability weighting: Evidence from the field. *Review of Financial Studies*, 34 (9):4524–4563, 2021.

Ecker, Frank, Xitong Li, Yilan Li, and Fan Wu. How stock market participants use generative artificial intelligence: Evidence from user-platform interaction data. *Journal of Accounting Research*, 64(3): 1375–1426, 2026.

Fedyk, Anastassia, Ali Kakhbod, Peiyao Li, and Ulrike Malmendier. AI and perception biases in investments:

An experimental study. *Journal of Financial Economics*, forthcoming, 2026.

Fernandes, Daniel, John G. Lynch, Jr., and Richard G. Netemeyer. Financial literacy, financial education, and downstream financial behaviors. *Management Science*, 60(8):1861–1883, 2014.

Fieberg, Christian, Lars Hornuf, Maximilian Meiler, and David J. Streich. Using large language models for financial advice. CESifo Working Paper 11666, 2025.

Foerster, Stephen, Juhani T. Linnainmaa, Brian T. Melzer, and Alessandro Previtero. Retail financial advice: Does one size fit all? *Journal of Finance*, 72(4):1441–1482, 2017.

Foltyn, Richard and Jonna Olsson. The worth of a "wo": Gender bias in financial advice from LLMs. CEPR Discussion Paper 21323, 2026.

Gennaioli, N., A. Shleifer, and R. Vishny. 2015. Money Doctors. *Journal of Finance* 70(1), 91–114. doi:10.1111/jofi.12188.

Giglio, Stefano, Matteo Maggiori, Johannes Stroebel, and Stephen Utkus. Five facts about beliefs and portfolios. *American Economic Review*, 111(5):1481–1522, 2021.

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár. Leniency designs: An operator's manual. Working paper, June 2026 version. arXiv:2511.03572. doi:10.48550/arXiv.2511.03572.

Google. More than 1 billion people are using the Gemini app every month. The Keyword, August 11, 2026.

Haddad, Valentin, and Tyler Muir. Market macrostructure: Institutions and asset prices. *Annual Review of Financial Economics*, 17:133–150, 2025.

Handa, Kunal, Alex Tamkin, Miles McCain, Saffron Huang, Esin Durmus, Sarah Heck, Jared Mueller, Jerry Hong, Stuart Ritchie, Tim Belonax, Kevin K. Troy, Dario Amodei, Jared Kaplan, Jack Clark, and Deep Ganguli. Which economic tasks are performed with AI? Evidence from millions of Claude conversations. Anthropic research paper, arXiv:2503.04761, 2025.

Hao, Yuzhi and Danyang Xie. A multi-LLM-agent-based framework for economic and public policy analysis. Working paper, arXiv:2502.16879, 2025.

Hirshleifer, D., and J. Sheng. 2022. Macro News and Micro News: Complements or Substitutes? *Journal of Financial Economics* 145, 1006–1024.

Hong, Claire Yurong, Xiaomeng Lu, and Jun Pan. Financial inclusion via FinTech: From digital payments to platform investments. *Management Science*, 72(8):6826–6850, 2026.

Horton, John J., Apostolos Filippas, and Benjamin S. Manning. Large language models as simulated economic agents: What can we learn from Homo silicus? NBER Working Paper 31122, 2023; revised February 2026.

HSBC Holdings plc. The Trust Threshold: AI Makes Investors Bolder, but They Want Human Judgement to

Make Decisions. Research commissioned from Ipsos, 2026. Global findings; U.S. findings, June 25, 2026.

Hudson, Sally, Peter Hull, and Jack Liebersohn. 2017. Interpreting Instrumented Difference-in-Differences. Working paper, September 4.

Hvide, Hans K., Tom G. Meling, Magne Mogstad, and Ola L. Vestad. Broadband internet and the stock market investments of individual investors. *Journal of Finance*, 79(3):2163–2194, 2024.

Investing.com. Survey: Nearly two-thirds of retail investors use ai to inform market decisions, April 2026. URL https://www.investing.com/news/stock-market-news/survey-nearly-twothirds-of-retail-investors-use-ai-to-inform-market-decisions-4598846.

J.D. Power. How Consumers Are Leveraging AI in Financial Services. Insights, September 2025. Report.

Kaiser, Tim, Annamaria Lusardi, Lukas Menkhoff, and Carly Urban. Financial education affects financial knowledge and downstream behaviors. *Journal of Financial Economics*, 145(2):255–272, 2022.

Kalda, Ankit, Benjamin Loos, Alessandro Previtero, and Andreas Hackethal. Smart(Phone) investing? Working paper, December 2024.

Karlsson, N., G. Loewenstein, and D. Seppi. 2009. The Ostrich Effect: Selective Attention to Information. *Journal of Risk and Uncertainty* 38(2), 95–115. doi:10.1007/s11166-009-9060-6.

Kemp, Andy. Where Americans and Canadians turn for financial guidance. Gallup, August 5, 2026.

Lin, Zhicheng. 2026. "Six Misconceptions about Large Language Models: A Minimal Model and Diagnostic Taxonomy." *PNAS Nexus* 5(7): pgag236. https://doi.org/10.1093/pnasnexus/pgag236.

Linnainmaa, Juhani T., Brian T. Melzer, and Alessandro Previtero. The misguided beliefs of financial advisors. *Journal of Finance*, 76(2):587–621, 2021.

Liu, H., L. Peng, and Y. Tang. 2023. Retail Attention, Institutional Attention. *Journal of Financial and Quantitative Analysis* 58, 1005–1038.

Lloyds Banking Group. Over 28 Million Adults Now Using AI Tools to Help Manage Their Money. Consumer Digital Index findings, November 3, 2025. Survey findings.

Loos, Benjamin, Pengfei Ma, Alessandro Previtero, and Andreas Hackethal. Robo-advisers and investor behavior. Working paper, May 2024.

Lynch, Shana. Inside the AI Index: 12 takeaways from the 2026 report. Stanford Institute for Human-Centered Artificial Intelligence, April 13, 2026.

Madrian, Brigitte C. and Dennis F. Shea. The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics*, 116(4):1149–1187, 2001.

Markowitz, Harry. Portfolio selection. *Journal of Finance*, 7(1):77–91, 1952.

Mo, Hongwei and Shumiao Ouyang. (Generative) AI in financial economics. *Journal of Chinese Economic and Business Studies*, 23:509–587, 2025.

Montiel Olea, José Luis and Carolin Pflueger. A robust test for weak instruments. *Journal of Business & Economic Statistics*, 31(3):358–369, 2013.

Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar. The market for financial advice: An audit study. NBER Working Paper 17929, 2012.

Ngangoue, Kathleen, Andrew Schotter, and Bill Wang. Life's big decisions in the age of AI. CESifo Working Paper 12515, 2026.

Noy, Shakked and Whitney Zhang. Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654):187–192, 2023.

OpenAI. A milestone in expanding access to AI. August 31, 2026.

Opinium. AI and Investing: Trust, Usage and Emerging Behaviours. Research in partnership with the Investment Association, July 30, 2026. Survey findings.

Ouyang, Shumiao, Hayong Yun, and Xingjian Zheng. AI as decision-maker: Ethics and risk preferences of LLMs. Working paper, revised December 2025.

Patwardhan, Tejal, Rachel Dias, Elizabeth Proehl, Grace Kim, Michele Wang, Olivia Watkins, Simón Posada Fishman, Marwan Aljubeh, Phoebe Thacker, Laurance Fauconnet, Natalie S. Kim, Patrick Chao, Samuel Miserendino, Gildas Chabot, David Li, Michael Sharman, Alexandra Barr, Amelia Glaese, and Jerry Tworek. GDPval: Evaluating AI model performance on real-world economically valuable tasks. arXiv Working Paper 2510.04374, 2025.

Polkovnichenko, Valery. Household portfolio diversification: A case for rank-dependent preferences. *Review of Financial Studies*, 18(4):1467–1502, 2005.

Quispe-Torreblanca, Edika, John Gathergood, George Loewenstein, and Neil Stewart. 2026. Attention Utility: Evidence from Individual Investors. *Review of Economic Studies* 93(1), 664–696. doi:10.1093/restud/rdaf028.

Reher, Michael and Stanislav Sokolinski. Robo advisors and access to wealth management. *Journal of Financial Economics*, 155:103829, 2024.

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

Rossi, Alberto G. and Stephen P. Utkus. The diversification and welfare effects of robo-advising. *Journal of Financial Economics*, 157:103869, 2024.

Ross, Jillian, Yoon Kim, and Andrew W. Lo. LLM economicus? Mapping the behavioral biases of LLMs via

utility theory. *Conference on Language Modeling*, 2024.

Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–2244, 2023.

Rumpf, Matthias, Michael Haliassos, Tetyana Kosyakova, and Thomas Otter. From humans to algorithms: How financial advice differs across professionals, peers, and LLMs. Working paper, 2026.

Sarkar, S. K. 2026. Economic Representations. Working paper, University of Chicago Booth School of Business, June 25. Author manuscript.

Sicherman, N., G. Loewenstein, D. Seppi, and S. Utkus. 2016. Financial Attention. *Review of Financial Studies* 29, 863–897.

Stolper, Oscar and Andreas Walter. Birds of a feather: The impact of homophily on the propensity to follow financial advice. *Review of Financial Studies*, 32(2):524–563, 2019.

Sun, Liyang and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199, 2021.

Takayanagi, Takehiro, Kiyoshi Izumi, Javier Sanz-Cruzado, Richard McCreadie, and Iadh Ounis. Are generative AI agents effective personalized financial advisors? *Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 286–295, 2025.

Tang, Ning, Olivia S. Mitchell, Gary R. Mottola, and Stephen P. Utkus. The efficiency of sponsor and participant portfolio choices in 401(k) plans. NBER Working Paper No. 15317, 2009.

van der Maaten, L., and G. Hinton. 2008. Visualizing Data using t-SNE. *Journal of Machine Learning Research* 9, 2579–2605. Published article.

Van Nieuwerburgh, Stijn and Laura Veldkamp. Information acquisition and under-diversification. *Review of Economic Studies*, 77(2):779–805, 2010.

Vanguard. Smart technology. Powerful advice. Vanguard advice resources, accessed October 25, 2026.

Veldkamp, Laura. Valuing data as an asset. *Review of Finance*, 27(5):1545–1562, 2023.

Verstyuk, Sergiy and Michael R. Douglas. Consumption and savings with large language model agents. Working paper, December 2025; revised January 2026.

von Gaudecker, Hans-Martin. How does household portfolio diversification vary with financial literacy and financial advice? *Journal of Finance*, 70(2):489–507, 2015.

Wang, Chenhao, Ting Zhang, and Shanyi Zhu. Big data recommendations and portfolio diversification: Evidence from account-level data. *Finance Research Letters*, 92:109541, 2026.

Welch, Ivo. Simply better market betas. *Critical Finance Review*, 11(1):37–64, 2022.

Appendix

A Data and Advice

Table A.1. Summary statistics: Investors in my sample and the 2022 Survey of Consumer Finances

	Investors in my sample			2022 SCF		
Variable	Mean	SD	Median	Mean	SD	Median
Panel A. Sample and Demographic Characteristics						
Equity-owning sample, *N*	1,854	–	–	2,966	–	–
Age (years)	52.9	12.9	55.0	52.1	16.8	52.0
Married (%)	79.6			67.3		
Any children/dependents (%)	54.0			39.7		
Number of children/dependents	0.91	0.99	1.00	0.71	1.05	0.00
Panel B. Portfolio Characteristics						
Hold > 1 direct stock (%)	75.0			28.2		
Portfolio value ($000)	1,908.9	2,925.8	555.7	845.3	5,986.9	155.0
Retirement share of investable wealth (%)	33.9	36.4	19.3	71.6	37.3	96.8
Portfolio composition						
Direct share of total equity (%)	44.9	32.6	41.0	45.4	38.5	35.9
Number of direct stocks	39.3	62.7	13.0	10.4	20.0	4.0
Total equity share of portfolio (%)	70.4	23.4	75.3	46.1	29.3	44.0
Indirect equity share of portfolio (%)	42.5	25.5	43.4	41.7	28.2	37.8
Direct equity share of portfolio (%)	33.6	29.4	25.3	20.8	24.1	10.9

Notes: This table reports summary statistics for investors in my sample and equity holders in the 2022 Survey of Consumer Finances (SCF). For binary measures, the Mean column reports the percentage satisfying the condition; SD and Median are left blank. The portfolio measures are defined following Polkovnichenko (2005). I measure SCF shares relative to financial assets as defined in that paper, and portfolio shares for investors in my sample relative to observed connected portfolio value. Indirect equity includes funds, retirement accounts, annuities, trusts, and managed accounts. Retirement share is the ratio of retirement account wealth to investable wealth.

Table A.2. Monthly portfolio turnover

Sample	Account coverage	Mean (%)	Median (%)
Portfolio tracking app	All linked accounts	8.39	2.09
Portfolio tracking app	Largest account only	7.50	1.55
Barber and Odean (2000)	One brokerage	6.49	2.67
Giglio et al. (2021)	Vanguard only	2.30	0.91

Notes: This table reports mean and median monthly turnover for investors using the portfolio tracking app and for the samples in Barber and Odean (2000) and Giglio et al. (2021). The app estimates cover all linked accounts or the largest account only. The figures from Barber and Odean (2000) refer to purchase turnover. All values are percentages.

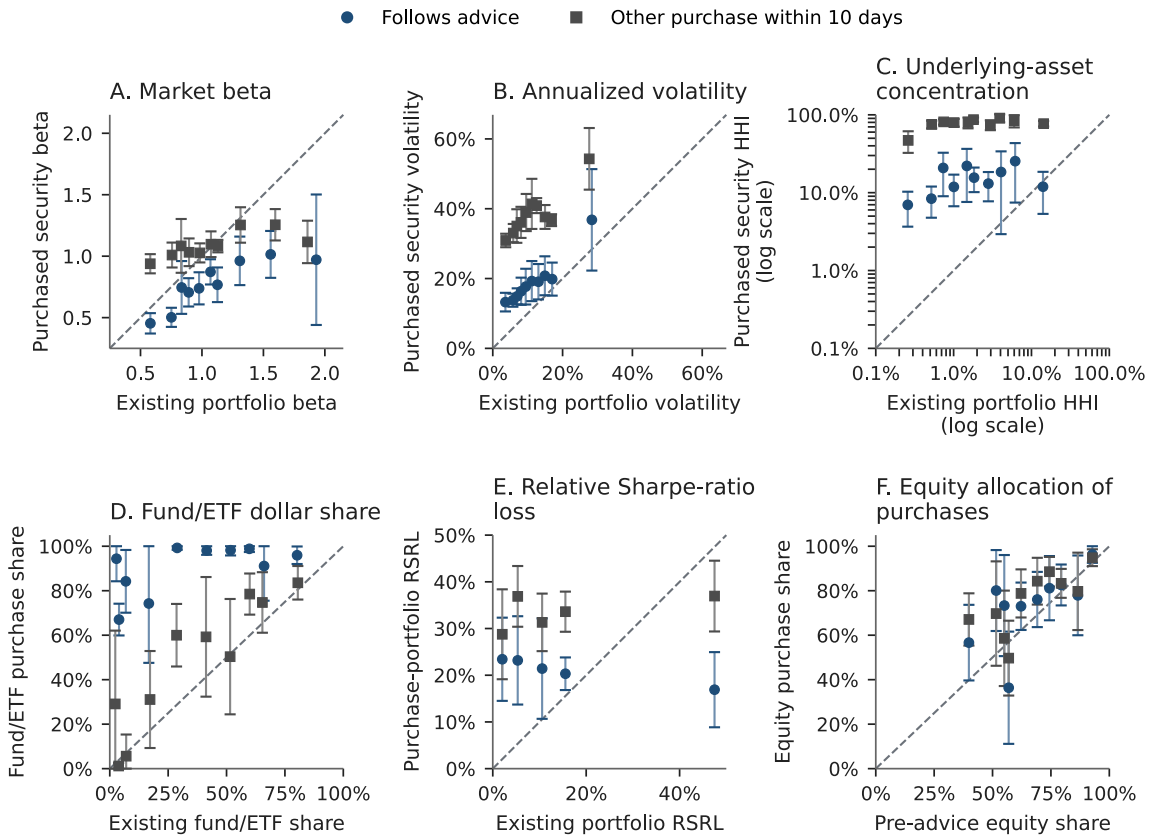


Figure A.1. Advice-following trades and existing portfolio characteristics

Notes: This figure plots means within five groups of the corresponding pre-advice characteristic. The error bars indicate 95 percent confidence intervals. Standard errors are double-clustered by individual and time. Relative Sharpe ratio loss is computed following Calvet et al. (2007).

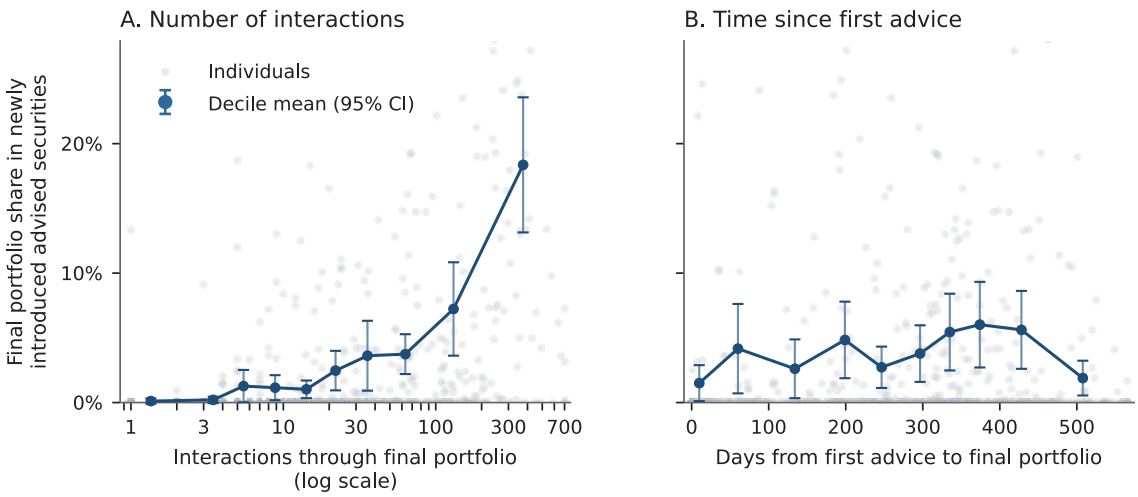


Figure A.2. Accumulation of newly introduced advised investments in investor portfolios

Notes: This figure plots the relationship between advice exposure and the share of the final portfolio invested in newly introduced AI-advised tickers. The horizontal axis measures cumulative classified answer interactions on a logarithmic scale in Panel A and days since first advice in Panel B. Gray points represent individuals; blue points report means within deciles, with 95 percent confidence intervals.

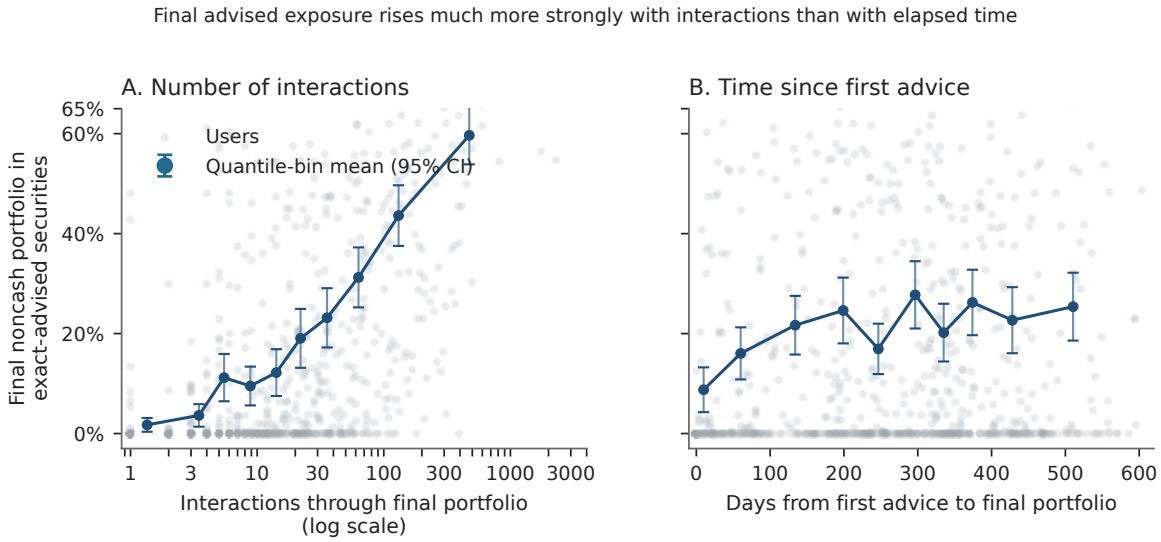


Figure A.3. AI-recommended securities in final portfolios

Notes: This figure plots the share of investors' final portfolios invested in securities previously recommended by AI. Panel A plots this share against the number of AI interactions on a log scale. Panel B reports the share invested against the days from the first AI recommendation. Both panels sort into deciles.

B Outcome Definitions

B.1 Portfolio Construction

I reconstruct daily portfolios from transactions, transfers, reinvestments, and dividend receipts. For positions that predate the available transaction history, the number of shares is calculated as the recorded dividend or fund distribution divided by the corresponding CRSP distribution per share. Purchases, sales, reinvestments, and split adjustments update each position.

The market value of each position is calculated as the shares held times the daily closing price per share. I calculate total portfolio value by summing positions across linked accounts. I compare the reconstructed holdings with periodic holdings snapshots. From these data, I construct daily measures of market value, asset class composition, concentration, and returns.

B.2 Variable Definitions

I define the following portfolio shares:

$$\text{Equity share} = \frac{\text{Equity holdings}}{\text{Total portfolio value}} \times 100,$$

$$\text{Passive fund portfolio share} = \frac{\text{Passive index fund holdings}}{\text{Total portfolio value}} \times 100,$$

$$\text{Individual stock portfolio share} = \frac{\text{Individual stock holdings}}{\text{Total portfolio value}} \times 100.$$

Total portfolio value is the value of all holdings across linked accounts. For the regression outcomes, I use the following purchase and month-end equity shares:

$$\text{Passive fund purchase share} = \frac{\text{Passive index fund purchases}}{\text{Purchases}} \times 100,$$

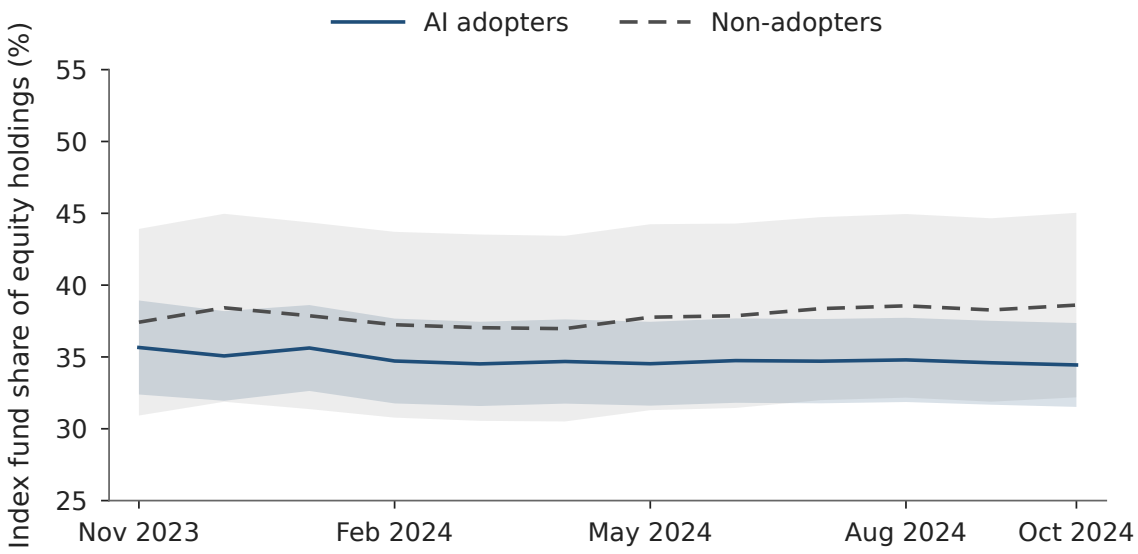
$$\text{Individual stock purchase share} = \frac{\text{Individual stock purchases}}{\text{Purchases}} \times 100,$$

$$\text{Passive fund holdings share} = \frac{\text{Passive index fund holdings}}{\text{Equity holdings}} \times 100.$$

Purchases and holdings are measured in dollars.

C Empirical Strategy

Figure C.1. Portfolio allocations before the introduction of the AI financial assistant



Notes: This figure plots the mean monthly share of passive index ETFs and funds in investors' portfolios from November 2023 to October 2024. Adopters are users who start using the AI financial assistant by the end of my sample in July 2026; never adopters are users with no observed AI use by that date. Bands indicate pointwise 95% confidence intervals. A t-test of equal means gives a p-value of 0.389.

D Additional Results and Robustness Checks

Table D.1. Effects of AI Adoption on Purchase Characteristics

	Full purchase portfolio			Stock-only purchases	Fund expense ratio
	U.S. market	World market	Welch BSWA (Welch, 2022)	U.S. market	Basis points
	(1)	(2)	(3)	(4)	(5)
Post-adoption effect	-0.15^{***}	-0.12^{***}	-0.12^{***}	0.01	-3.57^{**}
	(0.04)	(0.03)	(0.04)	(0.02)	(1.41)
DV mean	2.67	2.77	2.27	3.57	29.05
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
Calendar month fixed effects	Yes	Yes	Yes	Yes	Yes

Notes: This table reports estimates for the effects of AI adoption on purchase characteristics. The dependent variable in columns (1)–(4) is log idiosyncratic volatility under the indicated market model. Columns (1)–(3) use the combined monthly purchase portfolio, while column (4) is restricted to individual stock purchases. In column (5), the dependent variable is the purchase-dollar-weighted expense ratio of funds bought that month. Standard errors are double-clustered by individual and calendar month and reported in parentheses. Separately, CCS return loss (Calvet, Campbell, and Sodini, 2007) falls by 39.9 percent for the combined purchase portfolio. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

D.1 Purchase Risk and Fund Fee Distributions

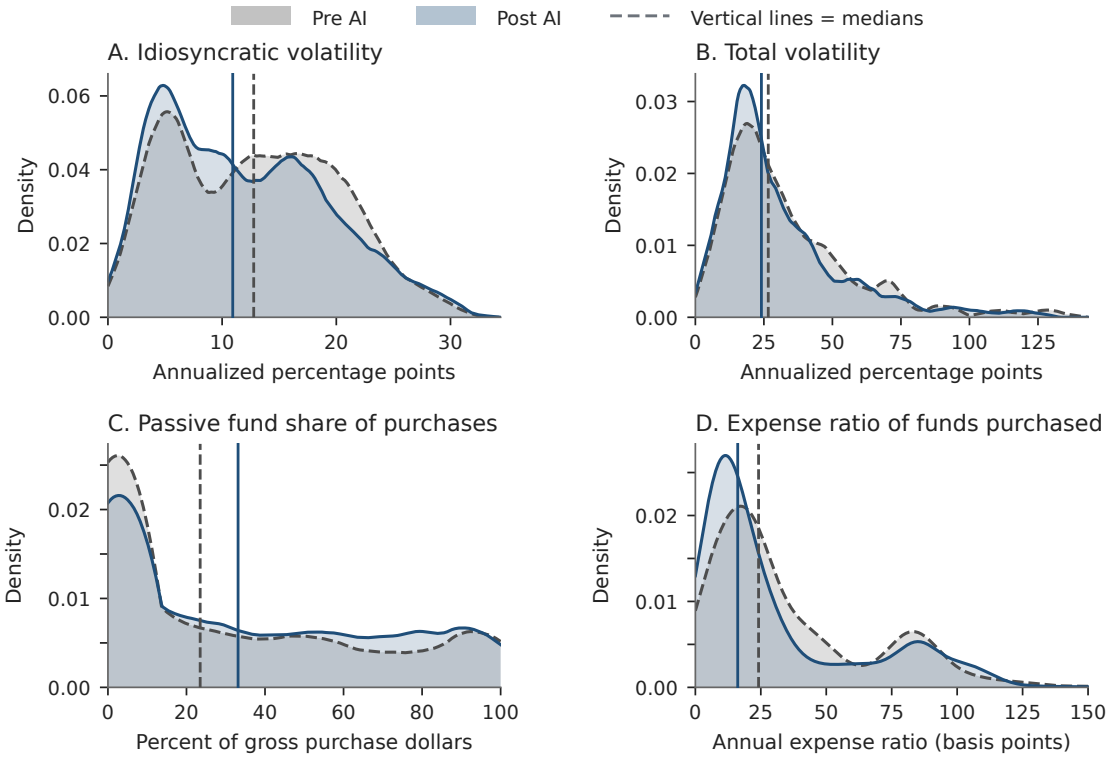


Figure D.1. Purchase portfolio distributions before and after AI adoption

Notes: This figure plots the distributions of purchase portfolio characteristics before and after AI adoption. Panels A and B measure purchase-dollar-weighted annualized volatility. Panel C measures strict passive index ETF and mutual fund purchases as a share of purchases. Panel D reports monthly fund expense ratios. The curves depict kernel densities, and the vertical lines indicate medians.

Panel	Pre median	Post median	Change	p-value
A. Idiosyncratic volatility (%)	12.77	10.93	-1.84	< 0.01
B. Total volatility (%)	26.60	24.10	-2.50	< 0.01
C. Passive fund purchase share (%)	23.49	33.14	+9.65	< 0.01
D. Fund expense ratio (bp)	24.14	16.18	-7.96	0.03

Changes equal post minus pre medians, in percentage points for Panels A–C and basis points for Panel D. Two-sided p-values test equality of the pre/post distributions.

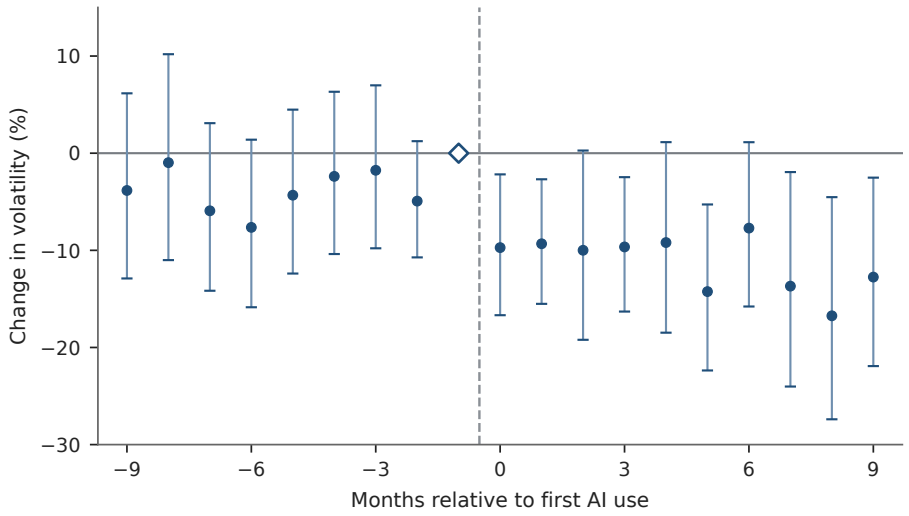


Figure D.2. Idiosyncratic volatility of purchases following AI adoption

Notes: This figure plots the coefficients and 95% confidence intervals from event-study regressions of AI adoption on the idiosyncratic volatility of purchases, using the Sun and Abraham (2021) interaction-weighted estimator. The dependent variable is the log of annualized residual volatility of purchased securities, estimated using an MSCI World raw return OLS market model. Event time is measured relative to AI adoption. All regressions include individual and calendar month fixed effects. Standard errors are clustered by both individual and calendar month.

Table D.2. Concentration and volatility of individual stock purchases

	Stock purchase HHI (1)	Log total volatility (2)	Log idiosyncratic volatility (3)
AI adoption	2.28	0.48	1.11
	(1.94)	(2.20)	(2.36)
R^2	0.58	0.59	0.59
Calendar month FE	Yes	Yes	Yes
Individual FE	Yes	Yes	Yes
DV mean	44.25	40.36	33.30

Notes: This table reports estimates for changes in the concentration and volatility of individual stock purchases following AI adoption using Sun and Abraham (2021). The dependent variables in columns (1)–(3) are 100 times the sum of squared stock-purchase-dollar weights, log total volatility, and log idiosyncratic volatility, respectively. Column (3) uses CRSP value-weighted market model residual volatility. Log coefficients are multiplied by 100 and approximate percentage changes; DV means are in levels. All specifications include individual and calendar month fixed effects. Standard errors are double-clustered by individual and calendar month and reported in parentheses. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

D.2 Fund Purchase Fees

Table D.3. Robustness of fund purchase fee estimates

	Weighted stacked DiD (1)	Two-way fixed effects (2)	Cohort-specific event study (3)	Group-time ATT (4)	Any fund purchase (5)
AI adoption	-2.18	-3.43***	-3.57**	-3.73	-2.20**
	(2.29)	(1.30)	(1.41)	(2.38)	(0.91)

Notes: This table reports estimates for changes in fund purchase expenses following AI adoption across alternative specifications. The dependent variable is the purchase-dollar-weighted annual expense ratio of fee-covered ETF and mutual fund purchases, measured in basis points. Column (5) uses the cohort-specific event-study estimator. In column (1), the event-specific datasets are stacked together. Standard errors are double-clustered by individual and time. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

D.3 Robustness

Table D.4. Passive Fund Shares across Difference-in-Differences Estimators

	Point estimate	Standard error	Lower bound	Upper bound
Panel A. Monthly Purchase Portfolio				
Sun and Abraham (2021)	7.60***	1.55	4.55	10.65
Stacked DiD	7.73***	2.20	3.42	12.05
Callaway and Sant'Anna (2021)	6.11**	2.81	0.61	11.61
Borusyak et al. (2024)	5.32***	1.93	1.53	9.12
de Chaisemartin and D'Haultfœuille (2026)	6.03***	2.18	1.76	10.29
TWFE–OLS	4.98***	1.61	1.82	8.13
Panel B. Selling-inclusive net share, 0.5 percent activity requirement				
Sun and Abraham (2021)	7.53***	2.23	3.16	11.90
Stacked DiD	11.18**	4.34	2.67	19.69
Callaway and Sant'Anna (2021)	6.91	4.25	−1.42	15.23
TWFE–OLS	6.08**	2.78	0.64	11.52
Panel C. Total Portfolio				
Sun and Abraham (2021)	2.73***	0.68	1.41	4.06
Stacked DiD	1.38	0.91	−0.41	3.17
Callaway and Sant'Anna (2021)	2.40**	1.13	0.18	4.61
TWFE–OLS	1.57*	0.86	−0.12	3.26

Notes: This table reports estimates for changes in passive fund shares following AI adoption using alternative difference-in-differences estimators. The dependent variables use the direct-index definition, and coefficients are expressed in percentage points. Standard errors are double-clustered by individual and time. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.5. Alternative Passive Fund Definitions

Specification	Point estimate	Standard error
Panel A. Monthly Purchase Portfolio		
Direct-index funds	7.60^{***}	1.55
CRSP B/D/E index funds	6.38^{***}	1.46
Panel B. Month-End Portfolio		
Direct-index funds	2.73^{***}	0.68
CRSP B/D/E index funds	2.50^{***}	0.68

Notes: This table reports estimates for changes in passive fund shares following AI adoption using alternative passive fund definitions and the Sun and Abraham (2021) estimator. The dependent variable is the passive fund share of the monthly purchase portfolio in Panel A and the passive fund share of month-end equity holdings in Panel B. All specifications include individual and calendar month fixed effects. Standard errors are double-clustered by individual and calendar month and reported in parentheses. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

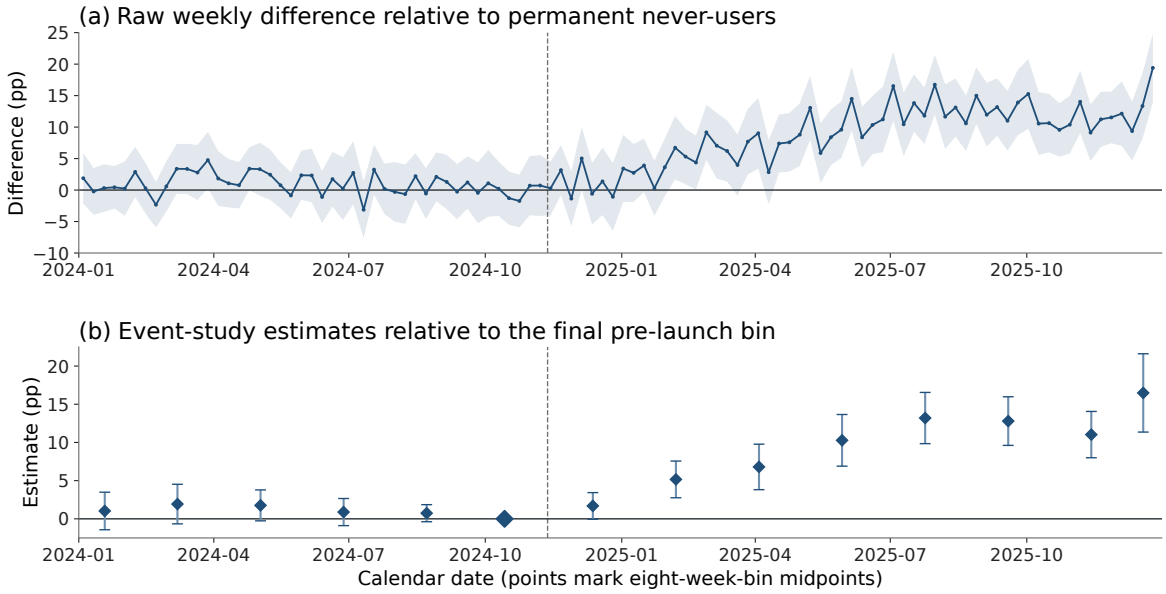


Figure D.3. AI introduction

Notes: This figure plots weekly differences in fund purchase shares and corresponding event-study estimates using the sample underlying the DiD specification. November 2024, when the AI assistant was first introduced, serves as the common treatment date. Panel A plots weekly differences in fund purchase shares between the two groups, and Panel B reports the corresponding event-study estimates in eight-week bins. The event-study specification includes investor and calendar-week fixed effects. Shaded bands and error bars indicate 95 percent confidence intervals. Standard errors are two-way clustered.

Table D.6. Alternative Monthly Activity Requirements

Monthly purchase requirement	Point estimate	Standard error
Any positive purchase	6.80^{***}	1.16
At least 0.1% of opening portfolio	5.78^{***}	1.43
At least 0.5% of opening portfolio	7.60^{***}	1.55

Notes: This table reports estimates for changes in the index fund purchase share following AI adoption using Sun and Abraham (2021) and the activity restrictions indicated in each column. The dependent variable is the direct-index fund share of the monthly purchase portfolio. All specifications include individual and calendar month fixed effects. Standard errors are double-clustered by individual and calendar month and reported in parentheses. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table D.7. Monthly Purchase Portfolio Effects across Adoption Cohorts

	Early	Middle	Late
Panel A. Equal calendar time periods			
Post-adoption effect	4.83	7.70***	7.72***
	(4.96)	(2.36)	(2.58)
Adoption dates	Oct. 2024– Mar. 2025	Apr.– Sept. 2025	Oct. 2025– Mar. 2026
Panel B. Equal-number adopter groups			
Post-adoption effect	6.09**	7.78***	8.17***
	(2.45)	(3.01)	(2.89)
Adoption dates	Nov. 13, 2024– July 17, 2025	July 17– Nov. 7, 2025	Nov. 8, 2025– Mar. 29, 2026

Notes: This table reports estimates for changes in the index fund purchase share following AI adoption, separately by adoption cohort, using stacked difference-in-differences. The dependent variable is the direct-index fund share of the Monthly Purchase Portfolio. Standard errors are double-clustered by individual and calendar month and reported in parentheses. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

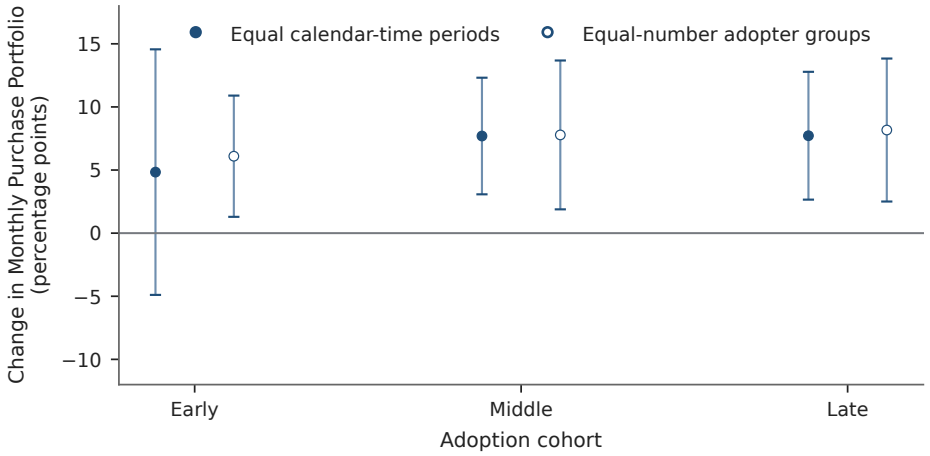


Figure D.4. Monthly Purchase Portfolio Effects across Adoption Cohorts

Notes: This figure plots the estimated coefficients and 95 percent confidence intervals from Table D.7. The dependent variable is the direct-index fund share of the Monthly Purchase Portfolio. Results are reported separately by adoption cohort.

E Additional Implementation Evidence

Table E.1. Advice implementation

	Implementation within (trading days)			
	1	2	5	10
Panel A. Answer-level implementation				
Constant (%)	16.6***	23.8***	32.0***	36.2***
	(2.15)	(2.52)	(2.59)	(2.43)
R^2	0.00	0.00	0.00	0.00
Panel B. Ticker-level implementation				
Constant (%)	9.0***	12.6***	16.8***	18.5***
	(2.29)	(2.60)	(2.70)	(2.46)
R^2	0.00	0.00	0.00	0.00
Individual fixed effects	No	No	No	No
Answer date fixed effects	No	No	No	No
Panel C. Ten-day regressions with fixed effects				
	Pooled	Individual FE	Date FE	Both FE
	(1)	(2)	(3)	(4)
Answer-level DV mean (%)	36.2	36.2	36.2	36.2
Answer-level R^2	0.00	0.26	0.14	0.34
Ticker-level DV mean (%)	18.5	18.5	18.5	18.5
Ticker-level R^2	0.00	0.22	0.15	0.28
Individual fixed effects	No	Yes	No	Yes
Answer date fixed effects	No	No	Yes	Yes

Notes: This table reports estimates for advice implementation rates across trading-day windows. Panels A and B report cumulative implementation rates (percent) from constant-only regressions. Implementation requires a purchase of any recommended security in an answer (A) or the specific ticker (B). Panel C adds the indicated fixed effects to the ten-day regressions. Standard errors are double-clustered by individual and time. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table E.2. Investment in recommended securities after advice

	10 trading days		30 trading days	
Panel A. Overall share of purchase dollars				
Share invested in recommended securities (%)	24.65		22.29	
Panel B. Per advice answer				
	Mean	Median	Mean	Median
Total purchases ($)	337,010	51,183	635,355	168,099
Purchases of recommended securities ($)	33,173	0	48,978	0
Share invested in recommended securities (%)	9.13	0.00	8.18	0.00
Panel C. Distribution of purchase outcomes (%)				
No purchases	8.32		3.24	
Only nonrecommended securities	53.56		48.76	
Recommended and nonrecommended securities	37.33		47.63	
Only recommended securities	0.80		0.37	

Notes: This table reports investment in recommended securities after AI advice. Panel A divides total recommended purchase dollars by total purchase dollars. Recommended purchases match any buy recommendation issued on or before their date whose window includes the purchase. Panels B and C match purchases to individual answers. Panel B's shares are conditional on positive purchases; Panel C reports percentages of answers.

E.1 Recommended Portfolios Versus Alternative Purchases

Table E.3. Historical diversification and risk: recommendations versus purchases

	Main estimate	Winsorized 1/99	Investors weighted equally
Reduction in relative Sharpe ratio loss	0.11***	0.11***	0.11***
	(0.01)	(0.01)	(0.01)
Reduction in market beta	0.23***	0.24***	0.24***
	(0.03)	(0.03)	(0.03)
Reduction in annualized volatility (pp)	11.27***	10.80***	12.22***
	(1.37)	(1.18)	(1.38)
Reduction in maximum peak-to-trough decline (pp)	10.35***	10.33***	10.31***
	(1.05)	(1.03)	(1.06)
Compared within the same advice event	Yes	Yes	Yes
Standard errors	Individual and time	Individual and time	Individual and time

Notes: This table reports estimates for differences in historical diversification and risk between recommended investments and purchases. Relative Sharpe ratio loss follows Calvet et al. (2007): $1 - S_h/S_B$, where S_h is the relevant constructed portfolio Sharpe ratio and S_B is market Sharpe. Under the zero-alpha one-factor CAPM, it is the share of market Sharpe lost to imperfect diversification. The row reports the reduction in this loss, so positive values favor the recommendation. Standard errors are double-clustered by individual and time. Maximum drawdown is the largest decline from a preceding peak. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table E.4. Characteristics of purchases that follow advice

	Follows advice vs. no recent advice	Follows advice vs. other post-advice purchase	Security named vs. other post-advice purchase
Diversified fund (pp)	31.56***	36.34***	24.88***
	(3.03)	(3.54)	(4.57)
Individual stock (pp)	-33.10***	-40.86***	-29.44***
	(3.23)	(3.66)	(5.76)
Broad passive equity (pp)	9.05***	10.66***	12.34***
	(1.32)	(1.49)	(1.59)
Expense ratio (bp)	-6.17***	-9.09***	-7.64***
	(0.82)	(1.15)	(1.27)
Historical volatility (pp)	-8.85***	-10.26***	-7.01***
	(0.74)	(0.81)	(1.13)
Historical beta	-0.19***	-0.22***	-0.10***
	(0.02)	(0.02)	(0.03)
Reduction in maximum peak-to-trough decline (pp)	8.50***	9.74***	6.68***
	(0.70)	(0.76)	(1.08)
Individual FE	Yes	Yes	Yes
Date FE	Yes	Yes	Yes

Notes: This table reports estimates for differences in the characteristics of purchases that follow advice and other purchases. Columns (1) and (2) compare purchases following advice with purchases without recent advice and other post-advice purchases, respectively. Column (3) compares named securities with other post-advice purchases. All specifications include investor and purchase date fixed effects. The coefficients are measured in percentage points for composition and volatility, and basis points for fees. Standard errors are clustered by investor and calendar week and reported in parentheses. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

F Investor Use: Supporting Evidence

Passive fund versus stock buying advice

Table F.1. Passive Fund and Stock Buying Advice by Financial Subject and Request Type

	Information seeking		Advice seeking	
Financial subject	Passive funds	Stocks	Passive funds	Stocks
Panel A. Raw shares (%)				
Securities and investment products	11.9	2.7	32.2	12.6
Portfolio allocation and risk	10.8	1.0	42.1	6.4
Existing holdings and performance	3.2	1.2	6.5	1.2
Retirement and financial planning	2.7	0.2	11.5	0.8
Taxes	5.6	0.6	24.5	4.0
Panel B. Adjusted probabilities (%)				
Securities and investment products	12.6	2.3	31.6	10.9
	(0.9)	(0.6)	(1.2)	(0.8)
Portfolio allocation and risk	10.6	3.0	39.7	6.4
	(1.9)	(0.7)	(0.9)	(0.5)
Existing holdings and performance	5.4	1.4	8.9	1.0
	(0.8)	(0.4)	(1.1)	(0.5)
Retirement and financial planning	4.0	1.6	12.5	1.9
	(0.9)	(0.3)	(1.2)	(0.4)
Taxes	6.4	0.9	23.2	4.2
	(1.2)	(0.5)	(1.4)	(0.6)

Notes: This table reports the likelihood that an AI answer recommends buying passive funds or individual stocks across financial subjects and request types. Panel A reports raw answer shares. Panel B reports OLS estimates (i.e., linear probability models) of these probabilities with individual, calendar month, and day of week fixed effects. Results are expressed as percentages. Standard errors are clustered by individual and shown in parentheses in Panel B.

G Additional Welfare Results

Table G.1. Fund allocation and portfolio efficiency: comparisons with other technologies

Paper / investor group	Fund holdings share (pp)	Return loss (annual bp)	Sharpe ratio change	Idiosyncratic volatility (%)
Generative AI: difference-in-differences				
All investors	2.73	−43.2	0.00680	−6.99
Less diversified	4.80	−101.1	0.02057	−19.83
More diversified	0.37	17.6	−0.00788	9.39
Other investment technologies				
Broadband: Hvide et al. (2024)	3.08	−34.0	0.00250	—
Smartphones: Kalda et al. (2024)	—	—	−0.14200	—
Vanguard: Rossi and Utkus (2024)	35.00	—	0.05300	−16.05

Notes: This table reports estimates for changes in fund allocation and portfolio efficiency following AI adoption alongside estimates for other technological innovations from the literature. Hvide et al. (2024) exploit variation in broadband rollout in Norway to estimate the effects of internet access on investment behavior and portfolio efficiency. Kalda et al. (2024) compare purchases made by the same investor in the same month via smartphones and other platforms at two German banks. Rossi and Utkus (2024) compare investors' portfolios before and after adopting Vanguard's robo-advice service.

Dashes indicate no comparable estimate.

H Personalization measures

The prompt reported in the Online Appendix assigns the overall LLM score and extracts personalization details from the question and AI answer.

Table H.1. Summary statistics for personalization

	Mean	SD	Median	Answers with any detail (%)	Share of all details (%)
Panel A. Overall measures					
Original LLM personalization score	7.04	2.08	8	—	—
Portfolio and life details	5.53	3.47	5	94.2	100.0
Panel B. Detail counts by type					
Portfolio state	4.72	3.04	5	93.4	85.3
Life details	0.81	1.13	0	46.0	14.7

Notes: This table reports summary statistics for personalization in AI answers. Panel A reports the LLM personalization score and the total number of portfolio and life details. Panel B reports portfolio and life details separately.

Table H.2. Personalization measures and advice implementation

	Advice implementation within ten trading days		
	(1)	(2)	(3)
Panel A. LLM personalization score			
LLM personalization score	2.71***	1.71***	1.56***
	(0.66)	(0.38)	(0.34)
One-standard-deviation increase (pp)	5.64***	3.56***	3.24***
	(1.38)	(0.79)	(0.70)
R^2	0.01	0.26	0.26
Panel B. Number of personalization details			
Number of personalization details	1.01**	0.61***	0.54**
	(0.41)	(0.23)	(0.21)
One-standard-deviation increase (pp)	3.50**	2.12***	1.86**
	(1.43)	(0.79)	(0.73)
R^2	0.01	0.25	0.26
DV mean (%)	43.06	43.06	43.06
Individual fixed effects	No	Yes	Yes
Month fixed effects	No	No	Yes

Notes: This table reports OLS estimates (i.e., a linear probability model) of the likelihood that an investor implements AI advice. The dependent variable is an indicator for purchasing at least one recommended security within ten trading days of the answer. The sample includes answers containing ETF or fund recommendations. Personalization is measured by the LLM's overall score from 1 to 10 in Panel A and the number of portfolio and life details in Panel B. The correlation between the two measures is 0.75. Column (1) includes no fixed effects, column (2) adds individual fixed effects, and column (3) additionally includes calendar month fixed effects. Coefficients are in percentage points per additional score point or detail, with corresponding estimates for an increase of one standard deviation. Standard errors are double-clustered by individual and time and shown in parentheses. Statistical significance at the 1%, 5%, and 10% levels is indicated with ***, **, and *, respectively.

H.1 Portfolio and life details

I next distinguish between life details and portfolio details. Life details describe personal circumstances that investors disclose through current or previous conversations. Portfolio information is supplied directly by the app to the LLM as context and can therefore be used to personalize an answer even when the investor does not mention it in the question. I consider the two types jointly to examine whether portfolio details are associated with implementation after accounting for life details.

Table H.3 reports this joint regression. With investor and month fixed effects, a one-standard-deviation increase in portfolio details is associated with a 2.04-percentage-point higher implementation rate (standard error 0.74). The corresponding estimate for life details is −0.36 percentage points (standard error 0.69) and is not statistically distinguishable from zero.

Table H.3. Portfolio and life details and implementation

	Implementation within ten trading days (×100)			
	(1)	(2)	(3)	(4)
Portfolio details (1 SD)	3.81***	2.26***	2.10**	2.04***
	(1.35)	(0.73)	(0.95)	(0.74)
Life details (1 SD)	-0.22	-0.20	-1.70	-0.36
	(1.40)	(0.85)	(1.41)	(0.69)
R^2	0.01	0.25	0.03	0.26
DV mean (%)	43.06	43.06	43.06	43.06
Individual fixed effects	No	Yes	No	Yes
Month fixed effects	No	No	Yes	Yes

Notes: This table reports estimates for the relationship between portfolio and life details and advice implementation. Each observation is an AI answer. Implementation equals one if the investor buys at least one recommended security within ten trading days of the answer. Portfolio and life detail counts enter jointly; coefficients are percentage points per standard deviation (3.04 portfolio details and 1.13 life details). Columns add the indicated fixed effects. Standard errors are double-clustered by individual and time. With both sets of fixed effects, the portfolio association is 2.04 percentage points, while the life detail estimate is indistinguishable from zero. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

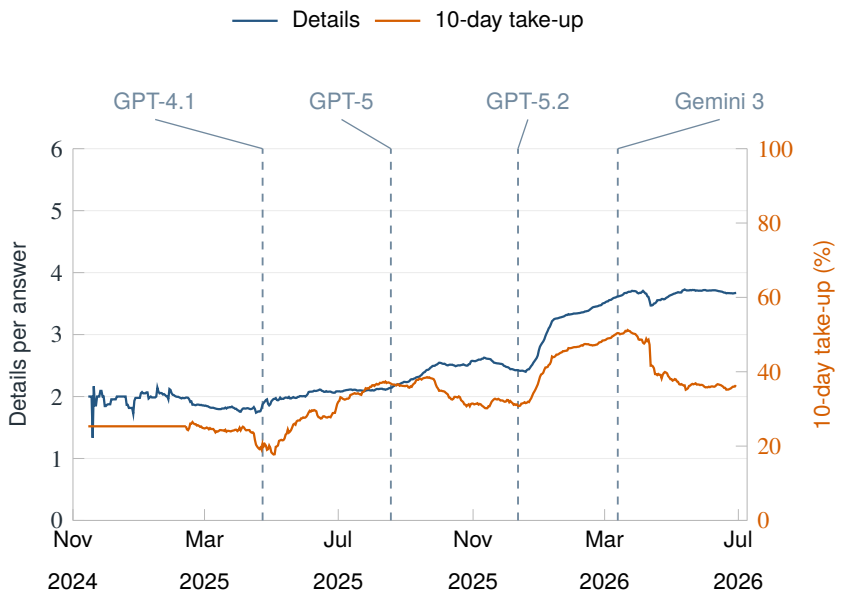


Figure H.1. Portfolio personalization and implementation over time

Notes: This figure plots portfolio personalization and advice implementation over time. The blue line (left axis) plots the average number of distinct portfolio detail categories used to personalize an answer. The orange line (right axis) plots the percentage of answers followed by a purchase of at least one recommended security within ten trading days of the answer, excluding the answer date. Both series are trailing 90-calendar-day moving averages. Vertical lines indicate the dates at which the default model changed.

H.2 First-stage regressions

Table H.4. Model indicators: first-stage regressions

	First stage Personalization (SD)	
	(1) Investor FE	(2) Full controls
GPT-4.1	0.20***	0.33***
	(0.07)	(0.08)
GPT-5	0.90***	0.93***
	(0.10)	(0.11)
GPT-5.2	0.78***	0.96***
	(0.15)	(0.15)
Gemini 3 Flash	0.99***	1.16***
	(0.08)	(0.08)
Gemini 3 Pro	0.61***	0.80***
	(0.12)	(0.10)
Gemini 3.1 Pro	0.63***	0.87***
	(0.10)	(0.10)
Investor fixed effects	Yes	Yes
Topic fixed effects	No	Yes
Day-of-week fixed effects	No	Yes
Other question / experience controls	No	Yes
Prior 30-day purchases and sales	No	Yes
Within R^2	0.08	0.24
First-stage F (clustered)	42.73	64.10
Effective first-stage F	35.74	44.23
Weak-IV critical value	17.86	18.29
First-stage partial R^2	0.08	0.09

Notes: This table reports estimates for the relationship between model indicators and portfolio personalization from first-stage regressions. Personalization is measured in standard deviations; one SD equals 3.04 portfolio details. GPT-4 is the reference category. The controls match Table 8. Standard errors in parentheses are clustered by investor and answer date. The first-stage F-statistic tests all model indicators jointly. Effective F-statistics and critical values follow Montiel Olea and Pflueger (2013), using a 5% test size and a 10% worst-case bias benchmark. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

I Financial Vocabulary

Table I.1. AI introduction and subsequent use of financial concepts

	Individual uses concept in first prompt of next conversation		
	(1)	(2)	(3)
Panel A. Pooled estimate			
AI introduces concept	3.19***	2.07***	1.98***
	(0.54)	(0.60)	(0.61)
DV mean	4.67	4.67	4.67
Total R^2	0.01	0.20	0.20
Individual fixed effects	No	Yes	Yes
Calendar week fixed effects	No	No	Yes

Panel B. Examples of concept-specific estimates

Concept	Estimate	Concept	Estimate
Fund	4.43***	Allocation	8.16***
	(0.93)		(1.17)
Diversification	4.52***	Volatility	-3.36***
	(0.98)		(0.36)
Concentration	-2.46***	Risk	11.21***
	(0.49)		(1.34)

Notes: This table reports estimates for investors' subsequent use of financial concepts introduced by AI. The unit of observation is an individual–concept–conversation. The dependent variable is an indicator equal to one if the individual uses the concept in the first prompt of the next conversation, within 90 days. The main regressor is an indicator equal to one if generative AI first introduces the concept. Panel A pools fund, allocation, diversification, and index. Panel B adds volatility, concentration, and risk and reports six examples from the seven-concept model with individual and calendar week fixed effects. Different forms of the same word are grouped together. Coefficients and means are expressed in percentage points. Column (1) includes no fixed effects; column (2) adds individual fixed effects; column (3) adds calendar week fixed effects. Standard errors are double-clustered by individual and time. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table I.2. Algorithmically selected financial terms and subsequent individual use

Panel A. Prevalence-based term selection			
Term	AI answers (%)	Individual prompts (%)	Ratio
Exposure	37.6	6.0	6.3×
Allocation	31.3	6.1	5.1×
IRA	30.2	5.0	6.1×
Yield	21.4	1.6	13.1×
Volatility	20.6	0.6	36.4×
Diversification	17.7	2.5	7.0×
Liquidity	13.5	0.6	24.5×
Roth	13.2	0.9	15.6×
Panel B. Subsequent use in the next conversation			
	(1)	(2)	(3)
AI introduces term	2.94^{***}	2.52^{***}	2.50^{***}
	(0.31)	(0.33)	(0.33)
DV mean	2.08	2.08	2.08
Total R^2	0.01	0.07	0.07
Individual fixed effects	No	Yes	Yes
Calendar week fixed effects	No	No	Yes

Notes: This table reports the prevalence of selected financial terms and estimates for investors' subsequent use of those terms. Panel A reports the prevalence of selected financial terms that meet the screen. I retain exact words or phrases appearing in at least 10 percent of AI answers and at least three times as often in answers as in individual prompts. I measure prevalence as the share of individual-written prompt–answer events containing the term. Panel B repeats the outcome in Table I.1 for the eight terms shown in Panel A. Coefficients and means are expressed in percentage points. Column (1) includes no fixed effects; column (2) adds individual fixed effects; column (3) adds calendar week fixed effects. Standard errors are double-clustered by individual and time. Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table I.3. Financial terms retained by the prevalence screen

Term	AI answers (%)	Individual prompts (%)	Ratio
Tax	47.4	9.4	5.1×
Cash	42.6	7.6	5.6×
Growth	42.0	3.5	12.0×
Risk	38.9	6.8	5.7×
Exposure	37.6	6.0	6.3×
Taxable	37.2	4.5	8.3×
Value	35.8	2.1	17.3×
Fund	33.9	4.2	8.0×
Allocation	31.3	6.1	5.1×
IRA	30.2	5.0	6.1×
Stock	28.0	4.6	6.1×
Equity	27.8	1.7	16.6×
Retirement	27.6	5.3	5.2×
ETF	27.5	2.5	10.8×
Sector	24.8	4.2	5.9×
Yield	21.4	1.6	13.1×
Volatility	20.6	0.6	36.4×
Dividend	20.1	4.2	4.8×
Index	20.1	0.9	23.1×
Bond	19.3	1.6	12.0×
Return	19.1	2.6	7.4×
Concentration	18.1	1.4	13.3×
Diversification	17.7	2.5	7.0×
Liquidity	13.5	0.6	24.5×
Roth	13.2	0.9	15.6×

Notes: This table reports the frequencies of financial terms retained by the prevalence screen. This list only contains terms that appear in at least 10 percent of AI answers and at least three times as often in answers as in prompts. Frequencies refer to the word shown; terms are ordered by frequency in AI answers.

J Concentration of Investment Recommendations

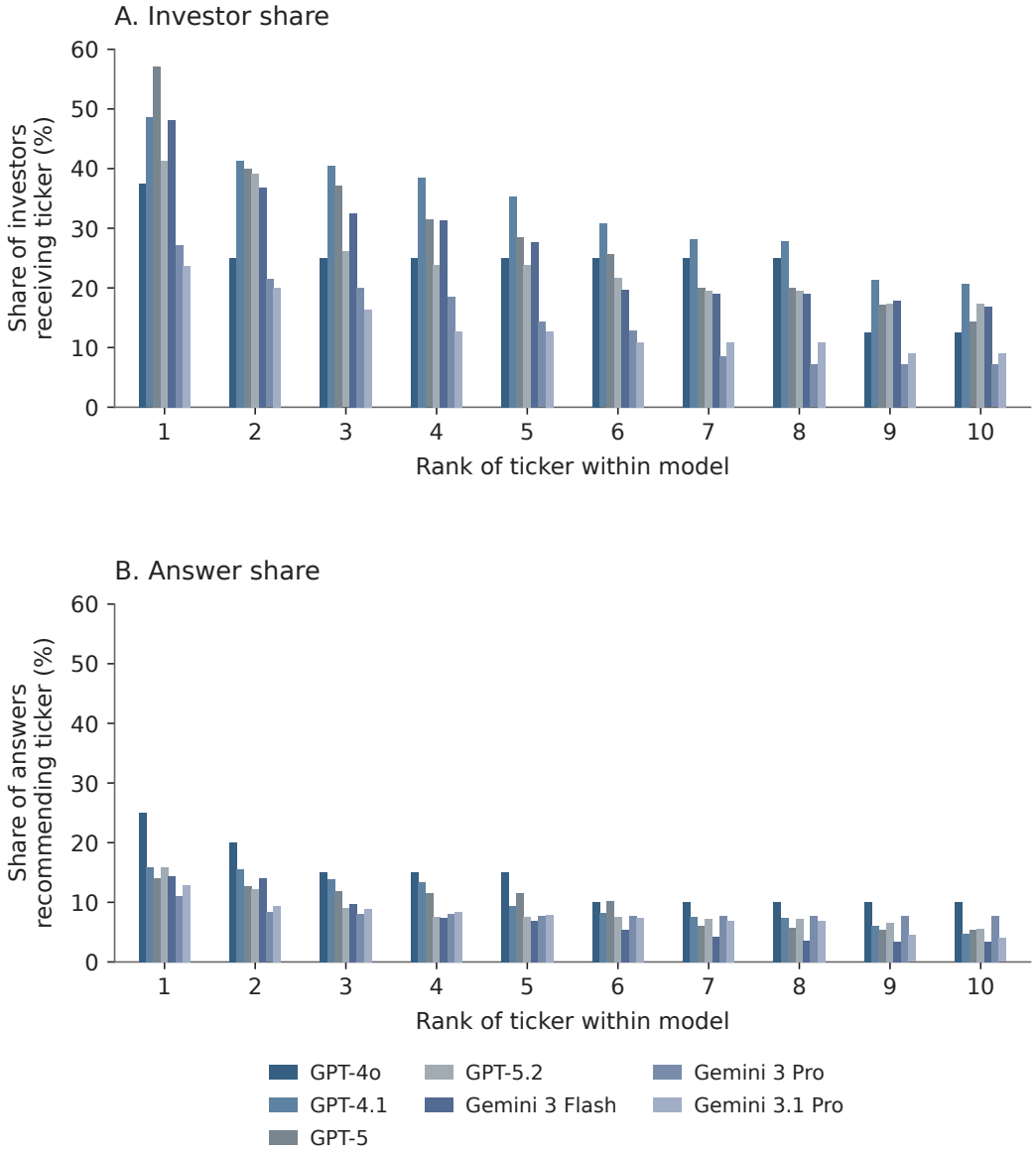


Figure J.1. Concentration of investment recommendations across models

Notes: This figure plots each model's recommendation shares for the top 10 most recommended tickers. Tickers are ranked separately within each model. Panel A reports the share of investors receiving a recommendation for each ticker, and Panel B reports the share of answers recommending each ticker.

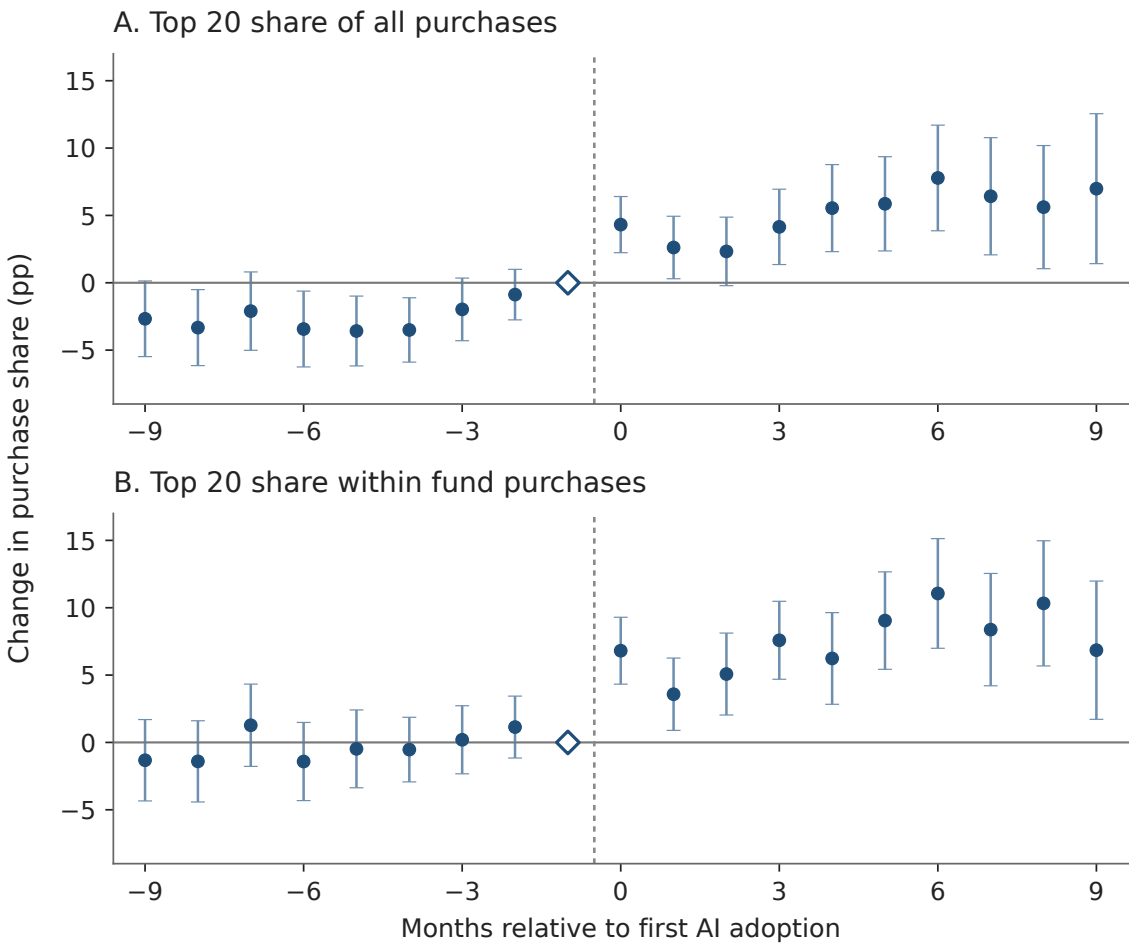


Figure J.2. Allocation of monthly investment around AI adoption

Notes: This figure plots event-study estimates for changes in the allocation of monthly investment around AI adoption. The top 20 are the funds recommended to the largest number of distinct investors, pooling recommendations across models. Points represent Sun and Abraham (2021) event-study estimates with 95 percent confidence intervals. All specifications include investor and calendar month fixed effects, with standard errors clustered by investor and calendar month.